PART IV

Macroeconomics

CHAPTER

13

Macroeconomic Concepts: GNP and Welfare

Microeconomics focuses on how markets function. It is useful for analyzing markets, with the aim of ensuring that they operate efficiently. Beyond that, however, it steers clear of policy recommendations. In essence, it assumes that the best policy is to let the market do its thing without interference. Macroeconomics looks at the economy as a whole, at the national or global level. In contrast to microeconomics, macroeconomics more often recognizes the importance of policy interventions, especially fiscal policy (government spending and taxation) and monetary policy (money supply and interest rates).[1] These policy interventions are important. However, the making of policy implies a goal. The traditional goal for macroeconomic policy is stable market-driven economic growth without limit, and to a lesser extent full employment. But unlimited economic growth is impossible. Many of the scarcest resources are nonmarket goods and services, and many of the most serious problems we now face extend beyond the borders of the nation-state.

In ecological economics, optimal scale replaces growth as a goal, followed by fair distribution. Traditional macroeconomics generally leaves allocation to market forces at the microeconomic level. Ecological economics more often recognizes that markets are inadequate for allocating

[1]There are many schools of macroeconomics, some of which do not call for policy interventions. For example, new classical macroeconomics, also known as rational expectations theory, argues that policy interventions are ineffective. The monetarists argue that policy interventions can be counterproductive. Both schools are quite conservative, favoring small, weak government. In practice, however, policy makers do use macroeconomic policy in efforts to attain policy objectives.

many scarce resources, and policy interventions are necessary to supply adequate quantities of these nonmarket goods. These different goals of ecological economics will favor different uses of traditional policies, and also suggest an array of alternative policy interventions.

In Chapters 13–16, we provide a brief introduction to some of the concepts, issues, and policy tools of mainstream macroeconomics and apply them to the policy goals of ecological economics. Remembering the circular flow diagram, we recall that macroeconomics deals with the aggregate flows of national product and income (the real sector). It also deals with the aggregate money supply, and with interest rates (the monetary sector). Following a discussion of the relationship of macroeconomics to microeconomics, this chapter will first look at the aggregate measure of the real sector, gross national product. Then we will have a look at money and the aggregate monetary sector, followed by a consideration of welfare indices other than GNP.

In Chapter 14, we will look at the medium by which wealth is measured in conventional economics: money. We then focus on distribution in Chapter 15. In Chapter 16, we will present the basic macroeconomic model for combining the two sectors into a simple general equilibrium model of the economy. This model (the IS-LM model) shows how the behavior of savers and investors in the real sector interacts with the behavior of the monetary authority (usually a national central bank, such as the Federal Reserve bank in the U.S.) and the money-holding public to determine the interest rate and the level of national income and employment. We will show how the goals of ecological economics lead to different policy recommendations than those supported by the mainstream. We will then discuss the possibilities for extending the IS-LM model to incorporate ecological constraints.

A Troubled Marriage

Microeconomics developed historically prior to macroeconomics. Indeed, if we understand the behavior of the decision-making units—firms and households—and how competitive markets work, then we will come to accept Adam Smith's "invisible hand"—that individuals seeking only their own benefit will automatically serve the common good. The microeconomic search for maximum private benefit will automatically result in the large-scale consequence of the greatest public welfare of all, or so it was thought. There was no need for special consideration of the macroeconomic picture, since the invisible hand would guarantee that if the microeconomics are right, then the macroeconomic picture will be right. Monopoly can ruin this nice result (as we showed in Chapter 8), so mar-

kets must be kept competitive, but that is about the only collective action needed.

In Chapter 2, we met Say's Law (supply creates its own demand), which guarantees that production always generates sufficient aggregate income to purchase aggregate production. Therefore, there can be no general glut of all products—at worst an imbalance in the mix of products, a misallocation, too much of something, too little of something else. That misallocation will soon be corrected by relative price changes in competitive markets. The same applies to the labor market—if there is unemployment (a surplus of labor), it simply means real wages are too high, not that there is a problem on a larger scale. If unemployment persists, you just need to let wages fall some more.

This view lasted well into the Great Depression of the 1930s and still has its adherents today. But under the leadership of John Maynard Keynes, economists began to think that prolonged unemployment, though theoretically impossible, was sufficiently real to warrant rethinking the theory. This rethinking led to the discovery of the leakages and injections from the circular flow, and the problem of making sure that total injections equal total leakages. We considered this in Chapter 2.

In addition, economists remembered the **fallacy of composition**, the false belief that whatever is true for the part must be true for the whole, or vice versa. For example, one spectator in a football stadium can get a better view by standing up. But all spectators cannot. If all stand on tiptoe, then no one has a better view than when everyone was comfortably seated. Similarly, one country can have a surplus or deficit in its balance of payments. But for the world as a whole, neither surplus nor deficit is possible because the sum of all exports must identically equal the sum of all imports. One worker may gain employment by being willing to work for a lower wage, but all workers probably could not, because lower wages for everyone means less income for the majority of the people, which means less spending on goods and services, and less demand for labor even at the lower wage. Reduced spending leads to reduced investment, which further lowers aggregate demand. In addition, any individual can easily convert his money holdings into real assets, but the community as a whole cannot, because when everyone tries to exchange money for real assets, someone has to end up holding the money.

There were, in sum, ample reasons to begin to develop a "macroeconomics" to deal with aggregate phenomena, especially unemployment and inflation. But contrary to what one might expect, the new macroeconomics did not build on the foundations of microeconomics. The macroeconomy is the aggregate of all the micro units, but macroeconomics is not just microeconomics aggregated. If it were, then we would be back to the invisible hand and the conclusion that macroeconomics was not necessary.

The entire economy described in microeconomic terms is the general equilibrium model. In it, all supply-and-demand relations in all markets are presented as one great interdependent system of simultaneous equations—say, a million equations in a million unknowns. This system of equations is solved by the market—a giant social computer that works by trial-and-error iteration. Economists have devoted much effort to counting equations and unknowns and making sure they were equal so that the general system, at least theoretically, could be solved. While the general equilibrium model is enlightening and conceptually satisfying, it is not very helpful from a policy perspective simply to know that everything depends on everything else. Policy needs a few leverage points at which to influence the gross behavior of the big system in its most important aspects. That is what macroeconomics has sought—simple models of the economy in terms of key aggregate variables, such as the money supply, aggregate price level, the interest rate, aggregate consumption and investment, exports and imports. And, of course, the biggest goal and leverage point of all—the rate of growth of GNP.

Ecological economics challenges today's standard emphasis on growth. Growth, yesterday's panacea, is rapidly becoming today's pandemic. Growth was a panacea because it was thought to be the solution to the macroeconomic problems of overpopulation, inequitable distribution, and involuntary unemployment. Microeconomists do not have much to say about growth, although not many would oppose it. Microeconomics is dominated by the concept of optimum and its associated "when to stop rule." As we argued in Chapter 2, if the macroeconomy is a part rather than the whole, then the logic of microeconomic optimizing applies, and at some point people trained in microeconomics will have to ask the macroeconomist, What is the optimal scale beyond which this economic subsystem should not grow? And when growth becomes uneconomic, as it must once we are at the optimum, then how are we going to deal with overpopulation, inequitable distribution, and involuntary unemployment?

It is the job of ecological economists to think about that: What happens after we reach the optimal scale, and how do we return there if we accidentally surpass it?

The whole is something more than the sum of its parts. In recognizing this, ecological economics bridges micro- and macroeconomics, though the exact relation between the two remains a bit mysterious. For our purposes, the relation between macro and micro is that shown in the circular flow diagram, Figure 2.4, repeated here as Figure 13.1.

The firms and households are our focus of attention in microeconomics. The firm as producing unit decides a supply plan for goods and a de-

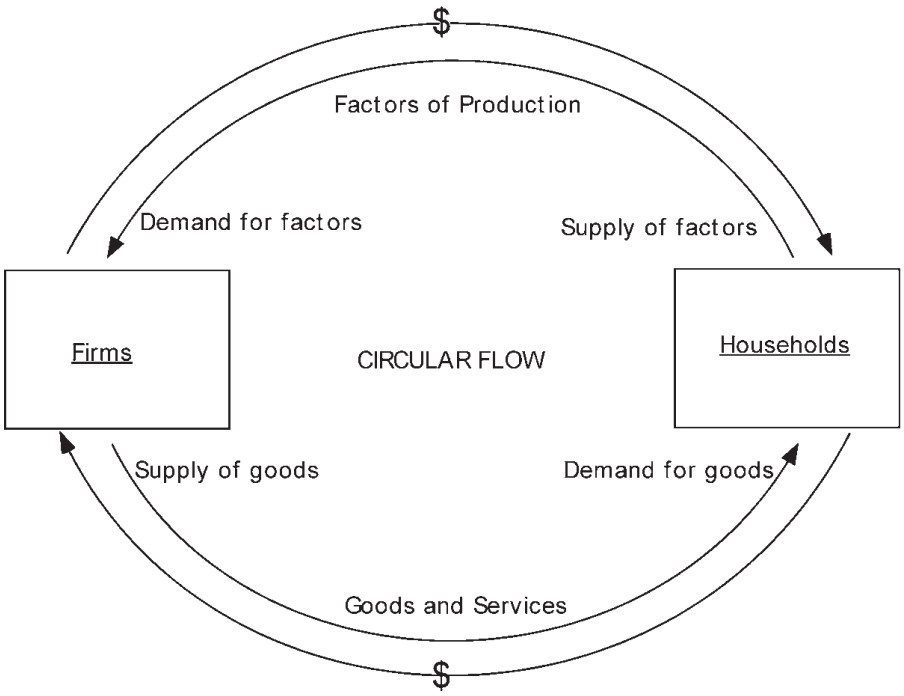

Figure 13.1 • The circular flow of the economy.

mand plan for factors. The household as consuming unit decides a demand plan for goods and a supply plan for factors. Microeconomics deals with these supply-and-demand decisions and their interactions in markets to determine the price and quantity of goods and factors exchanged in the markets. Because of its focus on prices, microeconomics is often called "price theory." Macroeconomics deals with the total volume of aggregate goods and services flowing through the goods market (national product), and the total volume of factors flowing through the factors market (national income). Because of its focus on aggregate income, macroeconomics is sometimes called "income theory."

Although we will speak of production and consumption because these terms are well established, it is important to remember (from earlier chapters), that in a *physical* sense, there is neither production nor consumption—only transformation. Raw materials are transformed into useful things (and waste) by "production." Useful things are transformed into waste by "consumption." What we are producing and consuming are "utilities"—useful temporary arrangements of matter and energy that serve our purposes. The throughput remains fundamental in both micro- and macroeconomics, even though it is not explicit in the accounts of firms and households, or in the aggregate accounts of nations. And the throughput is governed by the First and Second Laws of Thermodynamics, not by circular flow accounting conventions.

Gross National Product

Because economic growth is the paramount goal of nations, it is important to know just how it is measured. Growth in *what*, exactly? Economic growth is measured as growth in GNP, gross national product (or GDP, gross domestic product).[2]

As previously discussed in terms of the circular flow diagram, we have two measures of the aggregate circular flow that give the same number—national product and national income. Sometimes they are called national product at consumer goods prices (lower loop in Figure 13.1) and national product at factor prices (upper loop in Figure 13.1), or national income. Let's focus first on the lower loop, national product proper at consumer goods prices.

Gross national product (GNP) is the market value of final goods and services purchased by households, by government, and by foreigners (net of what we purchase from them), in the current year.

In this measure, **gross national product (GNP)** is the market value of final goods and services purchased by households, by government, and by foreigners (net of what we purchase from them), in the current year. With a few exceptions, anything not purchased this year is not counted.[3] Household production for the household itself is not sold and thus not counted—cooking, cleaning, childcare, and so on are omitted, unless done by a paid domestic helper. Intermediate transactions among firms are not counted. Only the sale of the final product to the household is counted. The wheat sold by the farmer to the miller is not counted, the flour sold by the miller to the baker is not counted; only the bread sold by the baker to the household for final consumption is counted. The value of the bread is the sum of the values added by the farmer, by the miller, and by the baker. Values added to what? To the basic natural resource—the wheat seed, the soil, the rain, the sunlight, and so on. The basic natural resources in most cases are considered to be free. Therefore, GNP is the sum of value added. It does not include any attribution of value to that to which the value was added. What is it that adds value to free natural resources? The transforming services of labor and capital funds.

Note that these accounting conventions are consistent with the neoclassical production function discussed in Chapter 9—namely, that production is a function of labor and capital only.[4] The exchange of existing

[2]The difference, not significant for our purposes, is that GNP counts production by all U.S. citizens whether at home or abroad. GDP counts all production within the geographic borders of the U.S., whether by citizens or by foreigners.

[3]E.g., annual rent is imputed to measure the current service of owner-occupied houses. The owner is thought of as renting his house from himself in the current year. Yet the owners of automobiles are not thought of as renting their cars to themselves.

[4]One might object that natural resources are not really free. A ton of coal does cost money on the market, but the money price is equal to the labor and capital cost of finding and extracting the coal. Coal in the ground, or *in situ*, as the resource economists say, is considered a free gift of nature. A particularly rich and accessible coal mine will require less labor and capital per ton of coal

assets is not counted because it is not current-year production. The value of a used car bought this year is not counted because it is a transfer of an existing asset. But the commission of the used car salesman will be counted as a service rendered this year. And of course the total value of a new car will be counted this year. The same holds for trading stocks on the stock market.

Total GNP is often divided by the population and stated as per-capita GNP. This is a simple mean and tells us nothing about the distribution of per-capita GNP of individuals about the mean. The mean may or may not reflect a representative central tendency in the distribution. Often modal or median per-capita income is a better measure of central tendency.[5]

GNP is measured in units of "dollar's worth". Dollar's worth of what? Of final goods and services traded in the market in the current year. It is the quantity of all such goods and services, times their price, all summed up. Changes in GNP over time can reflect price changes or quantity changes. To eliminate the effect of price level changes (inflation or deflation), economists correct the dollar figure by converting current dollars into dollars of constant purchasing power. This conversion is done by dividing nominal GNP by a price index that measures the rate of inflation. Suppose that there has been 20% inflation between 1990 and 2000. To convert year 2000 nominal GNP into real GNP, measured in dollars of 1990 purchasing power, we divide GNP in 2000 by 1.20; this is the price index that in the base year of 1990 would have been 1.00, but because of 20% inflation rose to 1.20 in 2000. This gives "real GNP" or rather GNP measured in dollars of constant purchasing power as of a base year.

Changes in real GNP are due to changes in quantities, not price levels. So real GNP, although measured in value units, is an index of quantities of something physical, and is therefore considered a better measure of economic growth than nominal GNP. Just as a dollar's worth of gasoline corresponds to a definite physical quantity of gasoline, so a dollar's worth of real GNP corresponds to some aggregate of physical goods and services. But because different goods and services have differing material and energy intensities, there is not a tight one-to-one relationship between real

---

than a marginal mine. Will its coal sell for less than that of the marginal mine? No, and this gives rise to producer surplus or differential rent. The more accessible mine earns a rent, which results from saved labor and capital relative to the marginal mine. Coal *in situ* is still a free gift of nature, but some free gifts are nicer than others, and differential rent takes that into account. The rent is attributed to the value of labor and capital saved in extraction, not to any original value of the coal in the ground.

[5]The mode is the income category that has the most members. The median is the per-capita income number for which there are as many members above as below. As students of statistics will know, for a normal distribution, the mean, median, and mode will coincide, all giving the same measure of central tendency.

GNP and physical throughput, as there is in the case of dollar's worth of gasoline and the throughput it represents.[6]

The point to emphasize is that although GNP is measured in value terms and cannot be reduced to a simple physical magnitude, it is nevertheless an index of an aggregate of things that all have irreducible physical dimensions. The relationship between real GNP and throughput is not fixed, but neither is its variability unlimited. And, to the extent that one believes that GNP growth can be uncoupled from throughput growth, then all the more reason to be willing to accept limits on throughput growth. If the environmental protection achieved by limiting throughput costs little or nothing in terms of reduced GNP growth, then no one should oppose it. If GNP could grow forever with a constant throughput, then ecological economists would have no objection.

GNP and Total Welfare

GNP is a measure of economic activity, not a measure of welfare. It tells us how fast the wheels are turning, not where the car is going. Economists all say that. Yet in the absence of a true measure of welfare, most policy makers look to the GNP as a trustworthy index of the general direction of change of welfare, based on the following:

Total welfare = economic welfare + noneconomic welfare

The faith-based assumption is that economic welfare and total welfare move in the same direction. But the increase in economic welfare could induce a more than offsetting decline in noneconomic welfare. For example, GNP goes up as labor becomes more mobile. But the welfare of being close to family and friends gets sacrificed as people have to move. Also, the extra income and job satisfaction of two-earner households raise economic welfare, but the stress of lost leisure and the extra financial burden and lost satisfactions resulting from external childcare reduce noneconomic welfare. Pollution-induced illnesses constitute an enormous loss of noneconomic welfare. Because the category "noneconomic welfare" is unmeasured while economic welfare has a numerical measure, we tend to overestimate the importance of the latter and underestimate the importance of the former. In Figure 13.2, the MDU curve, traditionally missing in economic analysis, represents the loss of "noneconomic welfare."

---

[6]But even here, economists try to keep the aggregate mix constant in calculating the price index. They assume a given basket of goods and given relative prices of goods in the basket in order to calculate a weighted average price of the basket and its change over time. This average price is not supposed to reflect either changes in relative composition of the basket of goods or changes in *relative* prices of the goods in the basket. Since relative prices inexorably do change over time, as does the composition of the representative basket of goods consumed, price level indexes inevitably "wear out" over time and have to be recalculated. Therefore, real GNP figures lose comparability over longer time periods.

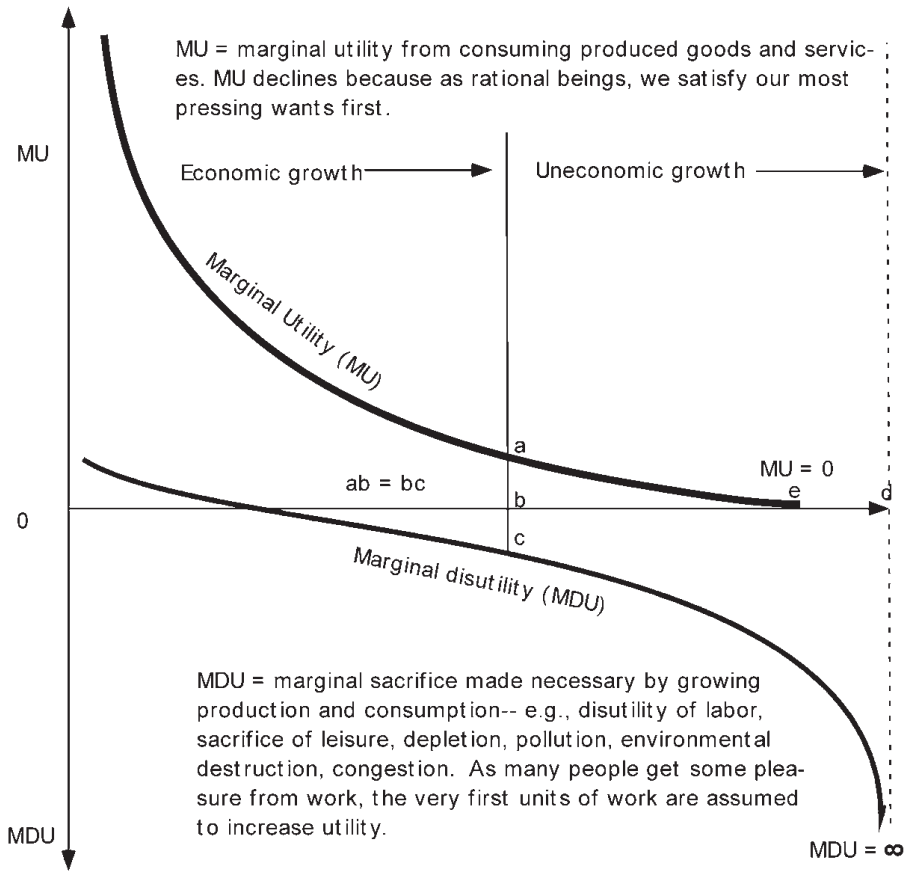

Figure 13.2 • Limits to growth of the macroeconomy. Point b = economic limit or optimal scale, where marginal utility (MU) = marginal disutility (MDU) (maximum net positive utility); e = futility limit, where MU = zero (consumer satiation), d = catastrophe limit, where MDU = infinity. At point d, we have gone beyond sustainable scale.

Defensive Expenditures and the Depletion of Natural Capital

Two other categories are problematic in national income and product accounts: regrettably necessary defensive expenditures and the depletion of natural capital. Let's have a look at each.

Regrettably necessary defensive expenditures, or "defensive expenditures" for short, are those expenditures that we have to make to protect ourselves from the unwanted consequences of the production and consumption of other goods by other people—for example, extra thick walls and windows to block out the sound of living near an airport or busy street or medical services resulting from pollution-induced asthma. In the sense of just measuring activity, these are freely chosen expenditures that people make in order to be better off in their concrete circumstances, and therefore should be counted—they are if not "goods," at least "anti-bads." In another sense, they are really involuntary intermediate costs of

Regrettably necessary defensive expenditures, or "defensive expenditures," are those expenditures that we have to make to protect ourselves from the unwanted consequences of the production and consumption of other goods by other people.

production that should not count as welfare to the final consumer, or as final consumption. This category could be broadly or narrowly defined. The examples just given reflect a narrow definition. Some would include all costs of global warming, and extra legal and law-enforcement costs resulting from a general breakdown in trust and increases in complexity attributed to economic growth. Exactly where to draw the line is a matter of judgment.

The depletion of natural capital is a more clear-cut category. GNP is *gross* national product. It is gross of depreciation of capital. If we deduct depreciation of manmade capital, we get net national product (NNP), which is a closer approximation to what we can consume without eventual impoverishment. But even in calculating NNP, there is no deduction for the depreciation and depletion of natural capital. Even NNP is gross of natural capital consumption (as well as gross of defensive expenditures). What's more, manmade capital is not a perfect substitute for natural capital for the simple reason that the former cannot exist without the latter. The two are complements. Putting a dollar value on the depreciation of both manmade capital and natural capital implicitly assumes that both types of capital are perfect substitutes, and that we can accept the loss of natural capital as long as manmade capital grows by a compensating amount. In reality, less natural capital makes our manmade capital less valuable as well. Of what use is a car if there is no gas to put in it?

Sustainable Income

The true definition of **income**, implicitly stated above, is the maximum that a community can consume in a given time period without causing itself to have to consume less in future time periods.[7] In other words, income is the maximum you can consume this year without reducing your capacity to produce and consume the same amount next year, and the year after—without reducing future productive capacity, that is, without consuming capital. Strictly speaking it is redundant to say "sustainable income" because income by definition is sustainable. Yet this feature of income has been so overlooked that a bit of redundancy for the sake of emphasis seems useful. If it's not sustainable it is, at least in part, capital consumption, not income.

The whole idea of income accounting is the prudent concern to avoid inadvertent impoverishment by consuming capital. Of course, there are times when we may choose to consume capital—for example, using a nest egg during retirement or liquidating the inventory of a store going out of business. Most of us, however, prefer not to run our national economy

[7] J. Hicks, *Value and Capital*, 2nd ed. Oxford, England: Clarendon, 1948.

and ecosystem as if it were a business in liquidation. Certainly you may choose to consume capital and voluntarily become impoverished. The income accountant's job is to make sure you know what you're doing, not to tell you what to do. But if the accountant does not deduct the consumption of natural capital in calculating income, then she has failed at her professional duty.

To be concrete, if you cut only this year's net growth of a forest, that's income because you can do the same thing again next year. If you cut down the whole forest, you cannot do it again next year, and the value of the cut forest is mostly capital consumption, not income. Yet in GNP, we count the whole amount as this year's income. The same is true for overexploited fisheries and croplands, and for depleted mines, wells, and aquifers.[8] Some neoclassical economists have come to realize that nature's services are a huge infrastructure to the economy, and we are failing to maintain that infrastructure.

Why do our national accountants fail to subtract natural capital consumption in calculating income? Neoclassical economics does not count natural capital consumption as a cost because in its preanalytic vision of the world, nature is not scarce. The reason natural funds and resource flows are absent from the usual neoclassical production function is also the reason there is no deduction for natural capital consumption in national income accounting.

Alternative Measures of Welfare: MEW and ISEW

In the early 1970s, there was considerable criticism of GNP growth as an adequate national goal—so much so that economists felt obliged to reply. The best reply came from William Nordhaus and James Tobin.[9] They questioned whether growth was obsolete as a measure of welfare, and thus as a proper guiding objective of policy. To answer their question, they developed a direct index of welfare, called Measured Economic Welfare (MEW), and tested its correlation with GNP over the period 1929–1965. They found that, for the period as a whole, GNP and MEW were indeed positively correlated; for every six units of increase in GNP, there was, on average, a four-unit increase in MEW. Economists breathed a sigh of relief, forgot about MEW, and concentrated again on GNP. Although GNP was not designed as a measure of welfare, it was, and still is, thought to be

[8]The running down of renewable stocks or funds of natural capital is *depreciation*, analogous to the depreciation of a machine. The running down of nonrenewable natural capital is *liquidation*, analogous to the liquidation of an inventory. Both represent capital consumption.

[9]W. Hordhaus and J. Tobin, "Is Growth Obsolete?" In *Economic Growth*, National Bureau of Economic Research, New York: Columbia University Press, 1972.

sufficiently well correlated with welfare to serve as a practical guide for policy.

Some 20 years later, Daly and Cobb revisited the issue and began to develop an Index of Sustainable Economic Welfare (ISEW) with a review of the Nordhaus and Tobin MEW. They discovered that if one takes only the latter half of the Nordhaus-Tobin time series (i.e., the 18 years from 1947 to 1965), the positive correlation between GNP and MEW *falls* dramatically. In this most recent half of the total period—surely the more relevant half for projections into the future—a six-unit increase in GNP yielded on average only a one-unit increase in MEW. This suggests that GNP growth at this stage in U.S. history may be quite an inefficient way of improving economic welfare—certainly less efficient than in the past.

The ISEW was then developed to replace MEW, since the latter omitted any correction for environmental costs, did not correct for distributional changes, and included leisure, which both dominated the MEW and introduced many arbitrary valuations.[10] The ISEW, like the MEW, though less so, was positively correlated with GNP up to a point (around 1980), beyond which the correlation turned slightly negative.[11] Figure 13.3 shows estimates of GNP and ISEW for seven different countries.

Measures of welfare are difficult and subject to many arbitrary judgments, so sweeping conclusions should be resisted. However, it seems fair to say that for the United States since 1947, the empirical evidence that GNP growth has increased welfare is weak, and since 1980 probably nonexistent. Consequently, any impact on welfare via policies that increase GNP growth would also be weak or nonexistent. In other words, the "great benefit," habitually used to justify sacrifices of the environment, community standards, and industrial peace, appears, on closer inspection, likely not even to exist.[12] Certainly if economic growth is to be the

[10]The concept of leisure is an important part of welfare, but the problems of valuing leisure are difficult. Is the leisure chosen or unchosen? Should sleep time count as leisure? Is commuting time leisure or "time cost of working"? Should we use the wage rate? The minimum wage? Should the "leisure" of mom taking care of children be valued at her opportunity cost if she's a doctor, or at the cost of avoided daycare? Such difficult choices have a big effect on the index.

[11]Neither the MEW nor the ISEW considered the effect of individual country GNP growth on the *global* environment, and consequently on welfare at geographic levels other than the nation. Nor was there any deduction for legal harmful products, such as tobacco or alcohol, or illegal harmful products, such as drugs. No deduction was made for overall diminishing marginal utility of income resulting from GNP growth over time (although a distributional correction for lower marginal utility of extra income to the rich was included). Such considerations would further weaken the correlation between GNP and welfare. Also, GNP, MEW, and ISEW all begin with personal consumption. Since all three measures have in common their largest single category, there is a significant autocorrelation bias, which makes the poor correlations between GNP and the two welfare measures all the more surprising.

[12]For further evidence from other countries, see M. Max-Neef, Economic Growth and Quality of Life: A Threshold Hypothesis, *Ecological Economics* 15: 115–118 (1995).

Figure 13.3 • Indices of GNP (solid) and ISEW (dashed) for seven countries. 1970 = 100 in all cases. (Source: R. Costanza, J. Farley, and P. Templet, "Quality of Life and the Distribution of Wealth and Resources." In R. Costanza and S. E. Jørgensen (eds.), *Understanding and Solving Environmental Problems in the 21st Century: Toward a New, Integrated Hard Problem Science*, Amsterdam: Elsevier, 2002.

number-one goal of nations and the central organizing principle of society, then citizens have a right to expect that the index by which we measure growth, GNP, ought to reflect general welfare more accurately than it does. Continued use of GNP as a proxy for welfare reminds us of the quote often attributed to Yogi Berra: "We may be lost, but we're making great time."

The objective, accurate scientific measurement of national costs and national benefits is not a realistic goal. Both costs and benefits of economic growth are spread out over time, and how we treat costs and benefits that affect future generations is an ethical issue, not a scientific one. The use of a particular discount rate to address intertemporal distribution is, for example, clearly a value-laden decision. Ecosystem change and evolution are not predictable, and how we treat the resulting uncertainty is also an ethical issue. Even using monetary measures of

Box 13-1 Gross National Cost?

Years ago, Kenneth Boulding suggested that the GNP be relabeled GNC for gross national *cost*. While we are sympathetic to Boulding's tongue-in-cheek plea, we can't help feeling a certain nostalgia for the good old days when newscasters regaled us with quarterly changes in the GNP. Now we are subjected to hammer-banging, gong-clanging reports of hourly changes in the Dow Jones and Nasdaq stock price indices—numbers that are an order of magnitude further removed from either welfare or income than GNP is! This is because stock market values are forward-looking, based on expectations of future earnings (even on speculators estimates of the expectations of others). By contrast, GNP is backward-looking, a historical record of what has already happened. Since the past is better known than the future, GNP is inherently a more trustworthy number than stock market values.

What should be done about GNP? Boulding's suggestion to count all of it as cost was an exaggeration, but no more so than the current practice of counting it all as benefit. A more sensible procedure would be to disaggregate GNP into two separate accounts: a national benefits account and a national costs account. As the scale of the economy grows, both benefits and costs will increase. We should be comparing those benefit and cost increases at the margin to find the optimal scale (see Figure 13.2).[a] It makes absolutely no sense to add them together!

[a] *For an effort in this direction for Australia, see P. A. Lawn,* Toward Sustainable Development: An Ecological Economics Approach, *Boca Raton, FL: Lewis Publishers, 2001.*

market goods is not objective; markets will yield different monetary values depending on the initial distribution of the wealth, and what comprises a desirable initial distribution is an ethical judgment. Monetary values for a given resource also vary depending on the amount of the resource society is using—for example, the price of oil depends primarily on current rates of extraction of oil. Oil is such an important input into so many economic processes that all prices are affected by how much oil we are using. Using prices determined by resource use in this period to decide the appropriate amount of a resource to use is therefore a case of circular reasoning; you can't do it on a computer spreadsheet, and you can't do it in real life. Efforts to put monetary values on nonmarket goods such as ecosystem services not only compounds these ethical issues with serious methodological problems, but also implicitly assumes that natural capital and manmade capital are perfect substitutes, a position that most ecological economists strongly reject.

Beyond Consumption-Based Indicators of Welfare

Personal consumption is not an end in itself, but merely one means toward achieving the end of enhancing human welfare. GNP is not only inadequate as a proxy for income, but income is only one element among many that provides human welfare. For example, the ecosystem services that increasing GNP inevitably encroaches upon are at least as important as GNP in providing welfare.[13]

Relative Wealth and Well-Being

While in the past, consumption was arguably one of the most limiting factors in enhancing welfare, abundant evidence suggests that this is no longer the case. Above a certain level, resource consumption and wealth may be "positional"—that is, our welfare is determined by comparing our position with that of others. Social psychology has long recognized that a person's sense of well-being is in part determined by his or her sense of relative affluence or deprivation. In this view, we are engaged in a never-ending wealth and consumption race, where greater consumption by our reference group demands greater consumption on our part simply to maintain the same relative position. If this is true, then doubling GNP with no change in distribution might have no positive impact on welfare. Moreover, a profoundly negative impact may result if the doubling of GNP comes at the expense of vital ecosystem services.

In reality, with current economic growth patterns leading to greater concentration of the wealth in the hands of the few, the majority of the population may be falling behind in this race. The wealthy generally compare themselves with each other and not with the poor, and therefore they may not be achieving greater welfare either. To the contrary, the blind pursuit of positional wealth and consumption places substantial demands on our time and resources, and leaves us with a decreasing ability to meet our other human needs.[14]

Human Needs and Welfare

Do other factors not yet discussed contribute to our welfare? It is reasonable to assume that welfare is determined by the ability to satisfy one's

[13]See, e.g., R. Costanza et al. The Value of the World's Ecosystem Services and Natural Capital, *Nature* 6630: 253–260 (1997), in which the value of global ecosystem goods and services is found to outweigh global GNP. While this paper does put monetary values on natural capital for purposes of comparison with manmade capital, it also explicitly discusses many of the problems with this approach.

[14]E.g., F. Hirsh, *The Social Limits to Growth*, Cambridge, MA: Harvard University Press, 1976; R. Frank, *Luxury Fever: Why Money Fails to Satisfy in an Era of Excess*. New York: Free Press, 1999.

needs and wants. What are our needs? Absolute needs are those required for survival and are biologically determined. Some 1.2 billion individuals globally and 28% of the population in the Third World currently live in extreme poverty and have difficulty meeting even these absolute needs.[15] For this group, greater consumption is probably very closely correlated to greater welfare.

Once absolute needs have been met, as is the case for the remaining four-fifths of the world's population, then welfare is determined by the satisfaction of a whole suite of primary human needs. Numerous researchers have proposed a variety of human needs, typically claiming that they are pursued in hierarchical order—Maslow's hierarchy (1954) (in which consumption only meets the lowest rung on the needs ladder) being the most famous. The hierarchical ordering, though generally not seen as rigid by these researchers, still leaves something to be desired. Even the 1.2 billion people living in absolute poverty seek to fulfill other needs than mere subsistence.

Manfred Max-Neef[16] has summarized and organized human needs into nonhierarchical axiological[17] and existential categories (Table 13.1). In this **matrix of human needs**, needs are interrelated and interactive—many needs are complementary, and different needs can be pursued simultaneously. This is a better reflection of reality than a strict hierarchy in which we only pursue higher needs after lower ones have been fulfilled. Also important in Max-Neef's conception, needs are both few and finite. This stands in stark contrast to the assumption of infinite wants, or the nonsatiety axiom in standard economics.

If we are to evaluate the success of economic policies both now and in the future (assuming that providing a high level of welfare for humans for the indefinite future is our economic goal), then we must develop measurable indicators that serve as suitable proxies for needs fulfillment and welfare.

To state the obvious, we cannot precisely measure welfare, which in the present context is equivalent to quality of life (QOL). In the words of Clifford Cobb:[18]

> *the most important fact to understand about QOL indicators is that all measures of quality are proxies—indirect measures of the true condition we are*

---

[15]D. E. Bloom, D. Canning, B. Graham, and J. Sevilla, "Out of Poverty: On the Feasibility of Halving Global Poverty by 2015" (Discussion Paper No. 52). Consulting Assistance on Economic Reform (CAER II), World Bank, 2001. Social Capital for Development: What Is Social Capital. Online: http://www.worldbank.org/poverty/scapital/whatsc.htm. World Bank.

[16]M. Max-Neef, Development and Human Needs. In P. Ekins and M. Max-Neef, *Real-life Economics: Understanding Wealth Creation*. London: Routledge, pp. 197–213, 1992.

[17]Axiology is the study of the nature of values and value judgments.

[18]C. W. Cobb, *Measurement Tools and the Quality of Life: Redefining Progress*, Oakland, CA. Online: http://www.rprogress.org/pubs/pdf/measure_qol.pdf.

Table 13.1

Max-Neef's Matrix of Human Needs

Axiological Categories	Existential Categories			
	Being	**Having**	**Doing**	**Interacting**
Subsistence	Physical health, mental health, equilibrium, sense of humor, adaptability	Food, shelter, work	Feed, procreate, rest, work	Living environment, social setting
Protection	Care, adaptability, autonomy, equilibrium, solidarity	Insurance systems, savings, social security, health systems, rights, family, work	Cooperate, prevent, plan take care of, cure, help	Living space, social environment, dwelling
Affection	Self-esteem, solidarity, respect, tolerance, generosity, receptiveness, passion, determination, sensuality, sense of humor	Friendships, family, partnerships with nature	Make love, caress, express, emotions, share, take care of, cultivate, appreciate	Privacy, intimacy, home, space of togetherness
Understanding	Critical conscience, receptiveness, curiosity, astonishment, discipline, intuition, rationality	Literature, teachers, method, educational policies, communication policies	Investigate, study, experiment, educate, analyze, meditate	Settings of formative interaction, schools, universities, academies, groups, communities, family
Participation	Adaptability, receptiveness, solidarity, willingness, determination, dedication, respect, passion, sense of humor	Rights, responsibilities, duties, privileges, work	Become affiliated, cooperate, propose, share, dissent, obey, interact, agree on, express opinions	Setting of participative interaction, parties, associations, churches, communities, neighborhoods, family
Idleness	Curiosity, receptiveness, imagination, recklessness, sense of humor, tranquility, sensuality	Games, spectacles, clubs, parties, peace of mind	Daydream, brood, dream, recall old times, give way to fantasies, remember, relax, have fun, play	Privacy, intimacy, space of closeness, free time, surroundings, landscapes
Creation	Passion, determination, intuition, imagination, boldness, rationality, autonomy, inventiveness, curiosity	Abilities, skills, method, work	Work, invent, build, design, interpret	Productive and feedback settings, workshops, cultural groups, audiences, spaces for expressions, temporal freedom

Continued

Table 13.1

Max-Neef's Matrix of Human Needs (Continued)

Axiological Categories	Existential Categories			
	Being	**Having**	**Doing**	**Interacting**
Identity	Sense of belonging, consistency, differentiation, self-esteem, assertiveness	Symbols, language, religion, habits, customs, reference groups, sexuality, values, norms, historical, memory, work	Commit oneself, integrate oneself, confront, decide on, get to know oneself, recognize oneself, actualize oneself, grow	Social rhythms, everyday settings, settings which one belongs, maturation stages
Freedom	Autonomy, self-esteem, determination, passion, assertiveness, open-mindedness, boldness, rebelliousness, tolerance	Equal rights	Dissent, choose, be different, run risks, develop awareness, commit oneself, disobey	Ability to come in contact with different people at different times in different places

The column of Being registers attributes, personal or collective, that are expressed as nouns. The column of Having registers institutions, norms, mechanisms, tools (not in material sense), laws, etc. that can be expressed in one or few words. The column of Doing registers locations and milieus (as time and spaces). It stands for the Spanish estar *or the German* befinden, *in the sense of time and space. As there is no corresponding word in English, Interacting was chosen for lack of something better.*

Source: M. Max-Neef, "Development and Human Needs." In P. Ekins and M. Max-Neef, Real-Life Economics: Understanding Wealth Creation. *London: Routledge, 1992, pp. 197–213.*

> *seeking to judge. If quality could be quantified, it would cease to be quality. Instead, it would be quantity. Quantitative measures should not be judged as true or false, but only in terms of their adequacy in bringing us closer to an unattainable goal. They can never directly ascertain quality. (p. 5)*

Objective Measures

Numerous efforts have been made to objectively measure welfare. The problem is that these studies have found only weak relationships between objective measures of welfare and the subjective assessments of the same by the subjects concerned.[19] However, both these studies and the various types of national accounts seem to include a relatively narrow range of objective indicators, often placing what we consider to be an excessive emphasis on consumption. Quite possibly the problem is that welfare is too rich a gumbo for us to recapture its flavor with so few ingredients. An important research agenda in economics is to develop a methodology for

[19]B. Haas, A Multidisciplinary Concept Analysis of Quality of Life. *Western Journal of Nursing Research* 21(6):728–743 (1999).

measuring access to "satisfiers" (the means by which we satisfy a given need) for Max-Neef's axiological and existential categories of human needs as indicators of welfare. With sufficient ingredients, we can produce something reasonably close to the flavor of welfare.

Max-Neef's human needs matrix as the basis of a welfare measure is a dramatic departure from existing national accounts, as well as from most of the proposed alternatives, differing even in its theoretical underpinnings. Neoclassical economics and GNP are explicitly utilitarian. Within utilitarian philosophy, individual welfare is determined by the degree to which individuals can satisfy their desires, and it is generally accepted that the goal of society is to provide the maximum amount of utility for its citizens. As utilitarian philosophy has been operationalized by NCE, citizens are the best able to determine what provides utility. Because it is extremely difficult to measure utility directly, economists have taken to using revealed preferences as a proxy. Preferences are revealed by people's objectively measurable choices in the market. In the market economy, preferences are revealed through market decisions, and market decisions can only be made with money. Under this conception of utilitarianism, the philosophy values only end-states and requires only "having" such things as possessions and experiences. Sustainable income accounting and measurements of economic welfare are basically just extensions of this philosophy, and they similarly value only "having."[20]

In Max-Neef's framework, having things is important, but it is just one of the elements required to meet our needs. Thus, a benevolent dictator with the resources to provide us with all the physical things we require for happiness would fail to meet our existential needs for being, doing, and interacting, as well as our axiological needs for creation, participation, and freedom. Also, within Max-Neef's conception, people are not always best able to determine what contributes to their quality of life—for example, advertising may falsely convince people that consumption satisfies their need for affection, freedom, or participation.

This approach, which values human actions independently of their outcomes, has been dubbed the "human development" approach to welfare. Its main proponents include Nobel Prize–winning economist Amartya Sen and Martha Nussbaum. In a similar tone to Max-Neef, they argue that "capabilities" and "functionings" are critical to welfare.[21] Roughly speaking, "functionings" correspond to human needs, while

[20]C. W. Cobb, Measurement Tools and the Quality of Life (Redefining Progress, Oakland, CA). (2000). Online: http://www.rprogress.org/pubs/pdf/measure_qol.pdf.

[21]Ibid.; M. Nussbaum, Aristotelian Social Democracy. In R. B. Douglass, G. M. Mara, and H. S. Richardson (eds.), *Liberalism and the Good*, New York: Routledge, pp. 203–252, 1990; R. Sugden, Welfare, Resources, and Capabilities: A Review of Inequality Reexamined by Amartya Sen. *Journal of Economic Literature* 31 (December): 1947–1962.

"capabilities" include both states of being and opportunities for doing, and are therefore analogous to access to satisfiers for these needs in Max-Neef's matrix (see Table 13.1). In utilitarian theory, we might have several different options, of which we choose one. If all options but that one were eliminated, it would not affect our welfare. In the human development approach, losing options restricts our capabilities and would therefore affect our welfare. The human development approach is less concerned with the actual choices that people make than with the options they are free to choose from, and the marketplace is only one of many spheres in which choice is important.

Operationalizing Human Needs Assessment as a Measure of Welfare

Measuring the extent to which human needs are satisfied is, of course, an exceptionally difficult task and highly subjective. Following the lead of Sen and Nussbaum, it would be most useful to measure capabilities, that is, the extent to which individuals have access to satisfiers. However, as noted by Max-Neef, specific satisfiers may vary by culture, and the difference in satisfiers required to meet a human need may indeed be one of the key elements that defines a culture. This means that objective "welfare accounts" must be very culture-specific. Second, some satisfiers might help fulfill several human needs, while other needs require several satisfiers. Further complicating matters, satisfiers may change through time. And humans are social creatures who inhabit a complex environment; needs are satisfied not only in regard to the individual, but also in regard to the social group and environment.[22] Furthermore, while needs are different and distinct, they are also interactive and may complement each other, and therefore may not be additive. Abundant access to satisfiers for one set of needs does not compensate for a lack of satisfiers for another set of needs. This suggests that separate "accounts" should be kept for access to satisfiers to different needs.

In developing welfare accounts based on **human needs assessment (HNA)**, it would be useful to test measurements of satisfiers empirically in studies comparing these objective measures against subjective assessments of welfare to determine their effectiveness. These empirical tests, as well as efforts to operationalize HNA accounts, must involve people in dialogues to confirm or refute the validity of the needs Max-Neef specifies, as well as the validity of the satisfiers we use to assess the degree to which needs are met. Such dialogues would almost certainly elicit additions and alternatives to the generic satisfiers, the entries in the columns of Table

[22]Op. cit., Max-Neef (1992).

13.1.[23] While the average person may not always know exactly what satisfiers will best meet their needs, interactive discussion with people is nonetheless essential to select and test appropriate indicators. We would also need to develop group-based methodologies to determine the effectiveness of our indicators in a social setting.

It is clear that Max-Neef's approach is very difficult to operationalize, even if his concept is theoretically more compelling than GNP or even ISEW. The debate over which approach to take to national accounting—theoretically sound measures or ease of accounting—is old. As Irving Fisher argued back in 1906, the appropriate measure, even of income, is one that captures the psychic flux of service (i.e., satisfaction of needs and wants) and not simply the final costs of goods and services.[24] And at the time Fisher wrote, the absence of suitable data for calculating either psychic flux of service or final costs no doubt led many to ignore the debate as entirely academic. The widespread use of GNP indicates that in practice, Fisher lost this earlier debate. However, measures such as the ISEW suggest that the GNP is becoming increasingly incapable of measuring economic welfare, much less general human welfare. Even if we can never quantify access to satisfiers as precisely as we currently quantify GNP, as Sen suggests, perhaps it is better to be vaguely right than precisely wrong.[25]

Accepting Max-Neef's human needs matrix as a framework for the specific elements of human welfare, and access to satisfiers as potentially the best objective indicator of welfare, has profound implications with respect to scale, distribution, and allocation. First, most of the possible indicators suggested by Max-Neef require few, if any, material resources beyond those required to sustain human life, and hence are not subject to physical exhaustion. Thus, for most elements of human welfare, increases for one person or one generation do not leave less for others. Second, explicitly accepting that there is a limit to material needs implies that we can limit consumption greatly with little, if any, sacrifice of welfare. This result is critical, because the laws of thermodynamics make it impossible to uncouple physical consumption from resource use and waste production. Abundant evidence suggests that current levels of consumption could not

[23]E.g., food and shelter are specific dimensions of "having" that are satisfiers of the need for "subsistence." How we actually meet our needs for food and shelter are culture-specific. A traditional Inuit might be satisfied with walrus blubber and an igloo, while a New Yorker would require hamburgers and a high-rise apartment.

[24]H. Daly and J. Cobb, *For the Common Good: Redirecting the Economy Toward Community, the Environment, and a Sustainable Future*. Boston: Beacon Press, 1989.

[25]D. Crocker, "Functioning and Capability: The Foundations of Sen's and Nussbaum's Development Ethic, Part 2." In M. Nussbaum and J. Glober, eds. *Women, Culture, and Development: A Study in Human Capabilities*, Oxford, England: Oxford University Press, 1995.

be sustainably met with renewable resources alone, and we must therefore limit consumption or else threaten the welfare of future generations.

The difficulty of operationalizing Max-Neef's framework may actually be a point in its favor. Why do we want to measure welfare in the first place? It's not just to track its rise or fall, but to help us create policies to improve it. Simply providing statistical data on welfare doesn't help us achieve this end. However, applying Max-Neef's framework would require extensive surveys asking people to think deeply about what their needs really are, and how they can satisfy them. Ultimately, improving welfare falls to decisions by political, cultural, and religious groups about what they want and how they want to achieve their goals, and making the correct decisions will require people to think deeply about what it is they ultimately desire.

BIG IDEAS to remember

- Fallacy of composition
- General equilibrium model versus aggregate macroeconomics
- Optimal scale of macroeconomy
- Gross national (or domestic) product
- Total welfare = economic welfare + noneconomic welfare
- Defensive expenditures or "anti-bads"
- Natural capital consumption
- Sustainable income
- MEW and ISEW
- Gross national cost
- Relative wealth and welfare
- Human needs and welfare
- Matrix of human needs (Max-Neef)
- Human needs assessment (HNA)

CHAPTER

14

Money

Money ranks with the wheel and fire as ancient inventions without which the modern world could not function. Probably more people today are "run over and burned" by out-of-control money than by out-of-control wheels and fires. Money is mysterious. Unlike matter and energy, it *can* be created and destroyed, evading the laws of thermodynamics. Private citizens (counterfeiters) are sent to jail for making even small amounts of it, yet private commercial banks make almost all of it, and we pay them for it! Sometimes money is a costly commodity (gold) and sometimes a costless token (paper notes). It is easily transferable into real assets by individuals, but the community as a whole cannot exchange its money into real assets at all, since someone in the community ends up holding the money. Some economists think the money supply should be determined by fixed rules, others think it should be manipulated by public authorities. And some people think the love of money is the root of all evil! Anyone who is not confused by money probably hasn't thought about it very much.

Money functions as a medium of exchange, a unit of account, and a store of value. The functions are interrelated but worth considering separately. To measure exchange value, we need a unit—call it a dollar, a peso, a franc, or a yen. If the unit is stable over time (no inflation or deflation), then money automatically serves as a store of exchange value. To function as a medium of exchange and let us escape the inconvenience of barter, money must hold its value at least long enough to effect both sides of the transaction, which in barter, of course, are simultaneous. A moment's reflection shows how tremendously inefficient barter is, and consequently how efficient money is. In barter there must be a coincidence of wants—it is not enough that I want what you have to trade; you also have to want whatever it is that I have to trade, and we have to find each other. Money

provides a common denominator that everyone wants simply because everyone else is willing to accept it. It is a standard, well-defined commodity (or later a token) that breaks the two sides of a difficult barter arrangement into two separate and easy transactions.

Karl Marx analyzed transactions as follows. First we have simple barter, which he denoted as:

$$C—C^*$$

Commodity C is exchanged for commodity C*. You have C and prefer C*; I have C* and prefer C. We are both better off after the transaction. We both increase the use value of what we own. Exchange value is not separated from use value. No money is needed, but we were lucky to have found each other.

Next for Marx comes "simple commodity production":

$$C—M—C^*$$

Now we have money functioning as a medium of exchange. Exchange value, the sum of money, M, is entirely instrumental to bringing about an increase in use values by facilitating the exchange. The process begins and ends with commodity use values. The goal is to increase use value, not exchange value.

For Marx the critical change comes in the historical shift from simple commodity production to "capitalist circulation," which he symbolized as:

$$M—C—M^*$$

The capitalist starts with a sum of money capital, M, uses it to make commodity C, and then sells C for the amount M*, presumably greater than M. Thus:

$$M^*—M = \Delta M$$

ΔM is profit, or surplus value in Marxist terms. For us the important thing is not Marx's notion of surplus value, which is tied up with his very problematic labor theory of value, but the simple observation that in moving from C—M—C* to M—C—M* the driving motive has shifted from increasing use value to increasing exchange value.

Use value arises from the actual use of commodities, it is concrete and physically embodied. **Exchange value** is abstract and inheres in money. It has no necessary physical embodiment.[1] Real wealth—commodities—obey the laws of thermodynamics. Money, a mere symbolic unit of account, can be created out of nothing and destroyed into nothing. There is a physical limit to the accumulation of use values. There is no obvious

[1]Though, of course, exchange value is only real if something exists for which money can actually be exchanged.

limit to the accumulation of exchange value. Fifty hammers are not much better than two (one and a spare) as far as use values are concerned. But in terms of exchange value, fifty hammers are much better than two, and better yet in the form of fifty hammers' worth of fungible money that can be spent on anything, anywhere, and at any time.

Box 14-1 Diamonds-Water Paradox

The distinction between use value and exchange value goes back to Aristotle and was used to "resolve" the "diamonds-water paradox"—the paradox that although water is a necessity it has a low price, while diamonds are practically useless but have a high price. Economists dealt with this conundrum by declaring that there are two basic kinds of value, use value and exchange value, and one has nothing to do with the other. In the late 1800s the marginalist revolution in economic thinking resolved the paradox as follows: Exchange value is determined by marginal utility and use value is determined by total utility; that is, exchange value equals marginal use value. Water has enormous total utility, but it is so plentiful that at the margin we use it for trivial satisfactions. This marginal utility determines exchange value. How do we know that? If you want to buy a gallon of water from me, what determines how much you will have to give me in exchange? If I give you a gallon of water, I won't stop drinking and go thirsty, nor will I stop bathing and be dirty. I'll probably water my petunias less often. The petunias are my least important use value, my marginal utility of water, my opportunity cost for a gallon of water. Since the *marginal* utility of water is what I will sacrifice by trading away a gallon, that's what determines the exchange value of water. Exchange value is determined by the least important use value, the value sacrificed. Water is abundant so its marginal utility is very small; diamonds are scarce, so their marginal utility is still high.

A hoard of hammers takes up space and is subject to rust, termites, fire, and theft. Fifty hammers' worth of money is not subject to rust, rot, and entropy, and far from costing a storage fee will earn interest from whomever gains the privilege of "storing" it for you. Production for use value is self-limiting. Production for the sake of exchange value is not self-limiting. Since there is no limit to the accumulation of abstract exchange value, and since abstract exchange value is convertible into concrete use value, we seem to have concluded that there must not be any limit to concrete use values either. This has perhaps led to the notion that exponential growth, the law of money growing in the bank at compound interest, is also the law of growth of the real, or material, economy.

Virtual Wealth

Frederick Soddy summarized all this by carefully distinguishing wealth from debt.[2] He noted that "a weight, although it is measured by what it will pull up, is nevertheless a pull down. The whole idea of balancing one thing against another in order to measure its quantity involves equating the quantity measured against an equal and opposite quantity. Wealth is the positive quantity to be measured and money as the claim to wealth is a debt" (p. 103).[3] Monetary debt, the measure of wealth, is negative wealth, say minus two pigs. It obeys the laws of mathematics, but not of physics. Wealth, on the other hand, plus two pigs, obeys the laws of thermodynamics as well as mathematics. Positive pigs die, have to be fed, and cannot reproduce faster than their gestation period allows. Negative pigs are hyper-fecund and can multiply mathematically without limit. As Soddy put it, "you cannot permanently pit an absurd human convention, such as the spontaneous increment of debt (compound interest), against the natural law of the spontaneous decrement of wealth (entropy)" (p. 30).

The holding of token money by the public to avoid the inconvenience of barter gives rise to the curious phenomenon that Soddy called **virtual wealth**, which he defined as the aggregate value of the real assets that the community voluntarily abstains from holding in order to hold money instead. Individuals can always convert their money holdings into real assets, but they choose not to in order to avoid the inconvenience of barter. This raises the question of whether money should be counted as a part of the real wealth of the community? Yes, if money is a commodity like gold that circulates at its commodity value. No, if it is token money like a dollar bill whose commodity value is nil but whose exchange value is significant. Even though each person can at an instant convert his money into real assets, it is impossible for the community as a whole to do this, as we have previously noted.

Money, therefore, represents not real wealth but, in Soddy's term, "virtual wealth." More exactly, it is the magnitude of virtual wealth that determines the value of money. What happens if the government puts into circulation more money than people currently want to hold? People will exchange money for real assets and drive up the price of real assets. As the price of real assets rises, the real value of money falls until it again coincides with the virtual wealth of the community. If there is too little money, people will exchange real assets for money, thereby

[2]F. Soddy, *Wealth, Virtual Wealth, and Debt*, London: George Allen & Unwin, 1926.

[3]When banks create money by providing someone with a loan (see below), they actually create a debt as the first step. On the assets side of the accounting books, the banker enters a debt for the amount of money borrowed (to be paid off with interest). This borrowed money is then placed in a bank account, which is listed in the bank's books as a liability.

driving down the price of real assets. As the price of real assets fall, the value of money increases until it again equals the virtual wealth of the community. The value of a dollar, then, is the virtual wealth of the community divided by however many dollars are in circulation. It follows that the value of a unit of token money is not determined by the total wealth of a community, nor by its annual GNP, but by its virtual wealth relative to the money supply.

Box 14-2 Virtual Wealth and Fiduciary Issue

Nobel laureate economist James Tobin comes very close to Soddy's concept of virtual wealth in his explanation of the "fiduciary issue":

> *The community's wealth now has two components: the real goods accumulated through past real investment and fiduciary or paper "goods" manufactured by the government from thin air. Of course, the nonhuman wealth of such a nation "really" consists only of its tangible capital. But as viewed by the inhabitants of the nation individually, wealth exceeds the tangible capital stock by the size of what we might term the fiduciary issue. This is an illusion, but only one of the many fallacies of composition which are basic to any economy or society. The illusion can be maintained unimpaired as long as society does not actually try to convert all its paper wealth into goods."*[a]

[a] *J. Tobin, "Money and Economic Growth,"* Econometrica *(October 1965), p. 676.*

Seigniorage

Who owns the virtual wealth? Since it does not really exist, we might say that no one owns it. It is a collective illusion. Yet individuals voluntarily hold money instead of real assets, and they behave as if money were a real part of their individual wealth, even if they understand that collectively it is only "virtual" or illusory. Every member of the community who holds money had to give up a real asset to get it—except for the issuer of money. The one who creates the money and is the first to spend it gets a real asset in exchange for a paper token. The difference between the monetary value and the negligible commodity value of the token, the profit to the issuer of money, was and still is called **seigniorage**, in recognition of the lordly nature of this privilege. Who is this fortunate person? Historically it was the feudal lord, or the king, the sovereign, who issued money within his domain. One might expect that this privilege would have been passed on to the sovereign's legitimate heir, the democratic state. To some extent this is the case, because only governments can issue currency or legal tender. However, over 90% of our money supply today is not currency but demand deposits created by the private commercial banking

system.[4] They are created out of nothing and loaned into existence by the private commercial banks under rules set up by the government. Who gets the seigniorage? Seigniorage from currency goes to the government. Seigniorage from demand deposits goes to the private sector, initially to commercial banks. To the extent there is competition among banks for savings, they will redistribute some of the seigniorage to depositors. Sectors of society too poor to save will receive nothing.

What does money consist of in our economy? A further mystery of money is that it has several definitions. The most restrictive is "currency plus demand deposits in the hands of the nonbank public." More expansive definitions include savings deposits and even credit card debt. Most of our money supply bears interest as a condition of its existence. Whoever borrowed it into existence must pay back what he borrowed plus interest. Thus, a requirement for growth (or else inflation) is built into the very existence of our money supply. Moreover, the money supply, *ceteris paribus*, expands during boom times when everyone wants to borrow and invest, and contracts during recessions when loans are foreclosed, thereby aggravating cyclical instabilities.

On learning for the first time that private banks create money out of nothing and lend it at interest, many people find it hard to believe. Indeed, according to Joseph Schumpeter, as late as the 1920s, 99 out of 100 economists believed that banks could no more create money than cloakrooms could create coats. Yet now every economics textbook explains how banks create money. We will explain how it works in a minute, but first we'll let the strangeness of it sink in. Of course, this is not the only way to create money. Nonetheless, most economists today accept this situation as normal. But the leading economists of the early part of the twentieth century, Irving Fisher and Frank Knight, thought it was an abomination. And so did Frederick Soddy.

Box 14-3 Local Currencies and Local Exchange Trading Systems

Currencies are created not exclusively by governments. A variety of non-government legal currencies exist in countries throughout the world, and a closer look at local currencies can provide important insights into money. There are three ways to design a currency system. Most national currencies are created by fiat. There is nothing to back up fiat money but faith that someone else will accept it in exchange for goods ("in God we Trust"

[4]Demand deposits are ordinary checking accounts from which money is payable "on demand" to the bearer of your check.

or as the Ithaca HOUR says "in Ithaca we trust"). Second, a currency can be valued in terms of a commodity, and may or may not be redeemable in terms of that commodity. For example, the Constant, one of the earliest alternative currencies and a forerunner of today's local currencies, was introduced in the 1970s on an experimental basis in Exeter, New Hampshire. The Constant was designed to maintain a constant value against a basket of 30 different commodities. Finally, a currency can be backed by a commodity, which means it can freely be exchanged for that commodity. Such was the case for U.S. currency in the nineteenth century when money holders could theoretically exchange gold-backed dollars for gold at any time, and the necessary gold reserves were physically available to do this.[a]

The city of Ithaca, New York, has one of the best-developed local currency systems in the world. The currency is known as Ithaca HOURS. An individual can participate in the HOURS system simply by agreeing to accept HOURS in exchange for the goods or services she produces. New money must be issued to chase this greater supply of goods and services. Where does this new money come from?

Published backers of the HOUR directory, which is considered a service provided to Ithaca HOURS, are paid 2 HOURS (the equivalent of approximately $20 US) on first participation and again when they renew their commitment. Technically speaking, the participant is being paid for publicly backing HOURS, but one could also say that in the HOURS system, the person who agrees to generate new goods and services earns the right to seigniorage. While at first glance it may seem strange that one would be entitled to money for simply agreeing to accept money, new money must clearly come from somewhere, and it's reasonable for part of it to go to the person responsible for creating the new wealth.

Theoretically, the amount of new money created times the velocity with which the money circulates should equal the amount of new goods and services being offered. So far it seems that new participants have on average offered more than enough goods and services to use up their 2 new HOURS. Several mechanisms are used to increase the money supply and prevent deflation. Residents of Ithaca may request interest-free loans of HOURS, organizations may request grants of HOURS, employees of member businesses can accept HOURS as a regular part of their pay, and people may purchase hours into circulation with dollars, from the HOUR bank. Additional money is created to finance administrative costs of the system. The circulation committee of Ithaca HOURS is responsible for deciding how many HOURS to create. So far, Ithaca HOURS are holding their own against the U.S. dollar, and they continue to trade at a ratio of 1 HOUR to 10 U.S. dollars.[b]

---

[a]R. Swann and S. Witt, Local Currencies: Catalysts for Sustainable Regional Economies. Revised 1988 Schumacher lecture, 1995/2001. Online: http://www.schumachersociety.org/currencypiece.html (E. F. Schumacher Society).

[b]See also http://www.ithacahours.com. Paul Glover, the founder of Ithaca HOURS, was also very helpful in providing information.

The Fractional Reserve System

What allows banks to create money is our fractional reserve system. If banks had to keep 100% reserves against the demand deposits they create, then there would be no creation of money. Hence, the reform called for by Soddy, Fisher, Knight, and others was for a 100% reserve requirement. Banks would still provide the convenience of checks and safekeeping, and they would charge for these services. They could still lend other people's money for them and make a profit. But they could not create money any longer.

Exactly how does the fractional reserve system enable banks to create money? Suppose the law required banks to keep 10% reserves against their demand deposits (actually, it is much less). Reserves are either cash or deposits with the Federal Reserve Bank owned by the commercial bank. The bank needs reserves only to settle the difference between daily deposits and withdrawals, which nearly always balance to within a few percent. Therefore, the bank feels that keeping 100% reserves is excessively cautious. It can keep only 10% reserves and meet all imbalances that are statistically likely to ever happen. The "excess reserves" can be loaned at interest, thereby increasing the bank's profits. The government has concurred in this practice and made it legal; it is known as **fractional reserve banking**. It works as long as all depositors do not demand their money at once, as happens in a bank panic (when depositors doubt the solvency of the bank and all rush to get their money out at the same time). To avoid panics, the government set up the Federal Deposit Insurance Corporation (FDIC). If depositors are insured against loss when a bank fails, then they will be less likely to panic and cause the very failure they feared. (They will also be less likely to demand prudence from their bank, but that's another story we leave for later.)

How do banks actually create money? Let's first consider a monopoly commercial bank. Because it is the only bank, it knows that any check drawn against it in one branch will be deposited with it in another branch. When it clears its own check, there is no transfer of money, of reserves, to another bank. Therefore, if it has a new cash deposit of $100 that counts as reserves, and the reserve requirement is 10%, it can lend out in newly created demand deposits an amount of $900. People and businesses borrow only what they intend to spend, so it is certain that this $900 will be spent. Its total additional demand deposits are $100 in exchange for the new cash deposit, plus $900 new loans, giving $1000 in new demand deposits backed by $100 in new reserves, thus satisfying the 10% reserve requirement. Net addition to the money supply is $900 worth of demand deposits.

Now let's consider a competitive banking system rather than a single

monopoly bank. Suppose Bank A receives a new cash deposit of $100. Unlike the monopoly bank, Bank A cannot lend out $900 because nearly all of the checks written on that amount of new demand deposits will be deposited in other banks, not Bank A. Clearing will necessitate a transfer of reserves to other banks. If it had lent out $900, it would surely soon have to transfer almost that amount to other banks. But it only has $100 in new reserves and thus will not be able to meet its legal reserve requirement of 10%.

So how much can Bank A lend as a result of a new cash deposit of $100? If it safely assumes that all checks drawn on its loans will be deposited in other banks, it can only lend out $90. Therefore, it still creates money—$90 in new demand deposits above the $100 demand deposit in exchange for the $100 cash. But the process does not stop here. The $90 of excess reserves safely lent by Bank A end up being transferred to Bank B, which can now safely lend 90% of that, or 0.9 ($90) = $81. So now the money supply has gone up by $90 + $81 = $171. But then the new $81 excess reserves of Bank B get transferred to Bank C, which can create new deposits of 0.9 ($81) = $72.90. And the process continues in an infinite series, the sum of which turns out to be—can you guess it? Exactly $900 of new money, as with the monopoly bank, or $1000 of new demand deposits, remembering the exchange of $100 cash for a $100 demand deposit that started the whole process.[5] The whole process works in reverse when someone withdraws cash (reserves) from the bank. The net result of simultaneous processes of money creation and destruction determines the net growth of the money supply.

Money as a Public Good

Money is a collective phenomenon, not a privately owned resource. In a peculiar but very real way, money is a true public good. You might think that if you own money, you can exclude others from using it, but if you did so completely, your money would have no value whatsoever. Money only has value if everyone can use it. And money is certainly nonrival, in that my spending a dollar in no way decreases the value of that dollar for the next person. Since money is a public good, one would expect seigniorage to be public revenue, not private. The virtual wealth of the community could be treated as a publicly owned resource, like the atmosphere or electromagnetic spectrum. But that is not the case. The money supply is privately loaned into existence at interest. The fact that most of our money

[5]If r is the required reserve ratio, then the demand deposit multiplier is this infinite series:

$$1 + (1 - r) + (1 - r)^2 + (1 - r)^3 \ldots\ldots + (1 - r)^n = 1/r$$

was loaned into existence and must be paid back at interest imparts a strong growth bias, as well as cyclical instability, to our economy. There is no economic reason why the monetary system must be linked with the private commercial activity of lending and borrowing.

What are the alternatives? Soddy offered three reforms. His first proposal was to gradually raise the reserve requirement to 100%. That would put the private banks out of the money creation business and back into the business of borrowing and lending other people's real money, providing checking services, and so on. Control of the money supply would then belong to the government. How, then, would the government regulate the money supply? Soddy's second policy suggested an automatic rule, based on a price level index. If the price level index is falling, the government should finance its own activities by simply printing new money and spending it into existence. If the price level is rising, the government should cease printing money and tax more than it spends, that is, run a surplus. This would suffice for a closed economy, but for an open economy, one that engages in international trade, the domestic money supply can be increased or decreased by international payments balances. Soddy's third proposal (back in 1926, under the gold standard) was freely fluctuating exchange rates. Currencies would trade freely and directly against each other; an equilibrium exchange rate would eliminate any surplus or shortage (deficit) in the balance of payments, and consequently any international effect on the domestic money supply. Remember our discussion of surplus and shortage in Chapter 9.

Of course, this is not what we have now.

The gold standard has been abandoned, and fixed exchange rate regimes have given way to flexible exchange rates, but not to freely floating exchange rates, which are thought (rightly or wrongly) to be too volatile and disruptive of international trade. (We return to the topic of exchange rates in Chapter 19.) The money supply is largely determined by the commercial banking system, subject to some manipulation, but not control, by the Federal Reserve (the Fed). The **Federal Reserve System** is a coordinated system of district central banks in the U.S. that influences interest rates and money supply.

The Federal Reserve has three tools for manipulating the money supply. First, the Fed can set the reserve requirements, within limits prescribed by law, and thus reduce or expand the supply of money created by banks, as explained above. This tool is used infrequently, because it has large impacts on the financial sector. Second, the Fed can change the interest rate it charges to lend reserves to the commercial banks (known as the discount rate), thus making it more or less profitable for the commercial banks to lend to their customers, and in doing so expand (or limit the expansion of) the money supply. Third, the Fed can conduct open market

operations, directly increasing or decreasing the money supply by buying and selling government securities in the open market. When the Fed buys government securities, it does so by crediting the bank account (at Reserve Banks) of securities dealers. This directly increases the available supply of money by the amount of the purchase. The deposit also increases the bank's reserves, allowing the bank to make more loans and create even more money. When the Fed sells government securities, the money supply contracts.

Money and Thermodynamics

Frederick Soddy was a Nobel Prize winner in chemistry, and a great believer that science should be used to benefit humankind. He doubted that this would happen, however, and even predicted back in 1926 the development of the atomic bomb. Why are the fruits of science often badly used? Because, thought Soddy, we have a flawed and irrational economic system. Unless we reform that system, scientific progress will serve to help us destroy the world faster. Soddy spent the second half of his 80-year life studying the economic system. He understood thermodynamics and entropy and the biophysical basis of economics, and forcefully called attention to this interdependence. But he focused his attention mainly on money. Why? Because money was the one thing that did not obey the laws of thermodynamics; it could be created and destroyed. And yet this undisciplined, imaginary magnitude was used as a symbol and counter for real wealth, which has an irreducible physical dimension, and cannot be created or annihilated. Money is the problem precisely because it leads us to think that wealth behaves like its symbol, money; that because it is possible for a few people to live on interest, it is possible for all to do so; that's because money can be used to buy land and land can yield a permanent revenue, therefore money can yield a permanent revenue.

Because of this fallacy, M. King Hubbert recently had to remind us that **exponential growth**—growth at a constant percentage rate—is a transient phase in human history.[6] The classic example of the power of exponential growth is the story about putting a grain of wheat on the first square of a chessboard, two grains on the second, four on the third, and so on. At the next-to-last, or 63rd, square the board contains 2^{63} grains of wheat, far more than the world's whole wheat crop, and the last, or 64th, square will by itself contain that much again. Hubbert's conclusion was that the world cannot sustain 64 doublings of even a grain of wheat. In our world, many populations are simultaneously doubling—populations

[6]M. King Hubbert, "Exponential Growth as a Transient Phenomenon in Human History," in H. Daly and K. Townsend, eds., *Valuing the Earth*, MIT Press, Cambridge, MA, 1993.

of people, livestock, cars, houses—things much bigger than a grain of wheat. How many times can each of these populations double? How many times can they all double together? A few tens at most, was Hubbert's answer. Our financial conventions, on the other hand, assume that this doubling will go on forever.

This expectation gets played out in reverse when we discount future values to an equivalent present value. We simply run the exponential calculation backward, asking: How much would we have to deposit in the bank today at today's interest (discount) rate in order to have the given future amount at a given future date? This discounting procedure is, as we have seen, at the heart of the financial model of present value maximization, which has displaced the more traditional economic model of profit maximization. The error that bothered Soddy is deeply ingrained in present economic thinking. We have already encountered it in our discussion of why renewable resources are driven to extinction.

It is convenient to dismiss Soddy as a "monetary crank" and to remark what a pity it was that such a brilliant chemist wasted so much of his time on a topic that he was unqualified to think about! This is exactly the treatment that Soddy was given. It was harder to dismiss Irving Fisher and Frank Knight, who also called for 100% reserve requirements, because they were the leading economists of their generation. But their ideas on money were simply classed separately from the rest of their economics, treated as a peccadillo, and were ignored.

Our previous statement—that money does not obey the laws of thermodynamics—needs some qualification. Exchange value is hardly a value if there is nothing for which it can be exchanged. If money is issued without real wealth to back it up, spending that money simply drives up the prices of goods and services, causing inflation, and bringing "real money" back closer into line with real wealth (more on inflation later).

What about virtual wealth? Are there limits to the amount of real wealth people are willing to forego in order to hold money? If not, then the amount of real money in circulation can continue to grow independently of the production of real goods and services. Financial assets are neither money nor real wealth, but they are bought and sold in the market, and people will hold more money to be able to meet their demand for transactions in these assets. In addition, people trade in money itself, using one national currency to buy another, and this similarly increases the demand for money. Both currency speculation and growth in financial assets have increased dramatically in recent years.

The M—C—M* equation previously showed how money has become less a means for facilitating exchange, more an end in itself. In reality, the M—C—M* equation has itself been dwarfed by pure currency speculation and trading in financial paper. John Maynard Keynes warned back in

the 1930s:[7] "Speculators may do no harm as bubbles on a steady stream of enterprise. But the position is serious when enterprise becomes the bubble on a whirlpool of speculation. When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done." While global production of marketed goods and services is roughly on the order of $30 trillion per year,[8] the trade in paper purchasing paper (or, more accurately these days, electrons purchasing electrons) with no intervening commodity is almost $2 trillion *per day*.[9] This means that the buying and selling of paper assets and currencies, M—M*, is more than 20 times greater than exchanges in the real economy! Real enterprise has indeed become a bubble on the whirlpool of speculation. As no productive activity intervenes in these speculative purchases, the sole result seems to be a magical growth in money. But is such growth actually possible indefinitely?

Growth in money is meaningless unless there is a corresponding increase in real wealth, so now we must ask: Does financial speculation lead to growth in real wealth? Some paper-paper purchases are purchases of new stock offerings that do provide financial capital, which can mobilize physical factors of production, but this is only an estimated 4% of stock purchases. Speculation in currency, in which millions of dollars are traded back and forth for very small margins over short time scales, clearly produce nothing. Indeed, such transactions almost certainly contributed to the crises in several South East Asian economies in 1997–1998 as speculators sold off regional currencies, and these crises meant dramatic *decreases* in production from those economies. Yet such speculation would not be undertaken unless some profits were being made somewhere. For example, George Soros, who participated in the financial speculation in South East Asia, is reported to have earned 1 billion pounds speculating on England's currency in 1995.[10] The only possible explanation is that if those who produce nothing are earning, through speculation, more money that entitles them to more real wealth, then those who actually do produce something must be becoming entitled to increasingly less wealth.

In summary, it would appear that the illusion that money can grow without physical limits results from three things. First, as long as the

---

[7]J. M. Keynes, *The General Theory of Employment, Interest and Money*. Orlando, FL: Harcourt Brace, 1991, p. 159.

[8]Official estimates based on purchasing power parity (PPP) are on the order of $40 trillion per year; the numbers for speculation are in nominal dollars, not PPP.

[9]D. Korten, *The Post-Corporate World: Life After Capitalism,* San Francisco: Berrett-Koehler, 1998.

[10]W. Greider, *One World, Ready of Not: The Manic Logic of Global Capitalism*, New York: Simon & Schuster, 1997.

production of real goods and services increases, more money is required to pursue them, so growth in money is justified. But such growth cannot, of course, continue forever on a finite planet. Second, as the number or price of financial assets grows, such as through speculative bubbles, then demand for money grows as well, and supply can increase to meet this demand. The fact is, however, that financial bubbles inevitably burst. Third, holders of financial capital see their capital grow because speculation can serve to transfer resources from those who produce to those who merely speculate. Such transfer of wealth has limits, though the limits are obscured by continued economic growth. Thus, the appearance that money is exempt from the laws of thermodynamics is an illusion that can only be maintained while scale is increasing, or the financial sector is expanding relative to the real sector. It remains impossible for real money to grow without limit.

THINK ABOUT IT!

What do you think would happen if a national government tried the same approach to seigniorage as Ithaca HOURS? For example, the government could impose 100% reserve requirements to prevent banks from creating money, award every new entrant to the economy some lump sum of money (perhaps by providing 18-year-olds sufficient money to pay for a college education or start a business), and lend money into existence at 0% interest for socially desirable projects.

BIG IDEAS to remember

- Money as: medium of exchange, unit of account, store of value
- Barter; simply commodity production; capitalist circulation
- Exchange value vs. use value
- Virtual wealth
- Seigniorage
- Fractional reserve banking
- Money creation
- Money as public good
- Federal Reserve System
- Money and laws of thermodynamics
- Local currencies

CHAPTER

15

Distribution

We have emphasized that ecological economics is concerned with three issues: the allocation of resources, their distribution, and the scale of the economy. We have seen how the ecological sustainability of the Earth is related to the size or scale of the macroeconomy. We have also explored the economist's meaning of efficient allocation in our discussion of microeconomics and the basic market equation. We then looked at the macroeconomic allocation problem in Chapter 16. But the second issue, distribution and the fairness thereof, has remained largely in the background.

Pareto Optimality

In dealing with allocation, we saw that economics defines efficiency as the Pareto optimal allocation of resources by the market. This definition assumes a given distribution of wealth and income. More specifically, an efficient allocation is one that best satisfies individual wants *weighted by the individual's ability to pay*—that is, by her income and wealth. Change the distribution of income and wealth, and we get a different set of efficient prices (since different people want different things), which define a different Pareto optimum. Because different Pareto optima are based on different distributions of income and wealth, economists are reluctant to compare them; one optimum is as good as another. We saw that a major reason for scale expansion—economic growth—has been to avoid the issue of distributive equality. As long as everyone is getting more from aggregate growth, then the distributive issue is less pressing, at least as a cure for poverty. Besides, the efficiency of the allocation of aggregate growth loses its well-defined meaning (Pareto optimality) once we accept the legitimacy of changing distribution in the interest of fairness. Consequently,

economics has tended to address distribution out of logical necessity, but quickly sets it aside in the interests of political convenience.

Box 15-1 Does a Pareto Optimal Allocation Assume a Given Scale as well as a Given Distribution?

If we take the concept of scale literally, as in the scale model of a house, to involve only a proportional change in all linear (scalar) dimensions, then we might say that a scale change is simply an increase or decrease in which all proportions remain constant. All relative prices, measuring unchanged relative scarcities, would also remain constant, defining an unchanging Pareto optimal allocation in terms of proportions. This seems to be what standard economists often have in mind. But is it possible to have everything grow in proportion? No, for two reasons. First, if something is fixed, it obviously cannot grow proportionally to everything else. What is fixed from the ecological economist's perspective is the size of the total ecosystem. As the economic subsystem grows, albeit proportionally in terms of its internal dimensions, the ecosystem itself does not grow. The economy becomes larger as a proportion of the total system—what we have called an increase in its scale, meaning size relative to the ecosystem. Natural capital becomes more scarce relative to manmade capital.

Of course, if the economy were to expand to encompass Earth's entire ecosystem (the model of "economic imperialism" in Chapter 3), the scale issue would disappear. In this sense the neoclassical economist's claim that if only all externalities were perfectly internalized then prices would automatically solve the scale problem (in the process of allocating everything in creation) makes sense. But it's a rather utopian point—like Archimedes' boast that he could move the Earth, if only he had a fulcrum and a long enough lever!

The second difficulty, long noticed by biologists and some economists, is that if you scale up anything (increase all linear dimensions by a fixed factor), you will inevitably change the relative magnitudes of nonlinear dimensions. Doubling length, width, and height will not double area; it will increase area by a factor of four and volume by a factor of eight. Biologists have long noted "the importance of being the right size." If a grasshopper were scaled up to the size of an elephant, it could not jump over a house. It would not even be able to move, because its weight (proportional to volume) would have increased eightfold, while its strength (proportional to a cross-sectional area of muscle and bone) would have increased only fourfold.

Returning to our example of a house, doubling the scale will increase surfaces and materials by fourfold and volumes to be heated or cooled by eightfold. Relative scarcities and relative prices cannot remain the same.

The answer to our question, does the notion of Pareto optimal allocation assume a given scale as well as a given distribution, appears to be yes. Size cannot increase proportionally because (1) there is a fixed factor, namely the size of the ecosystem, and (2) it is mathematically impossible even for all relevant internal dimensions of the subsystem to increase in the same proportion. In sum, it seems quite true that an optimal allocation assumes a given scale, just as it assumes a given distribution.

Economics prides itself on being a "positive science." Allocative efficiency is thought to be a positive, or empirically measurable, issue, even though, as we just saw, it presupposes a given distribution. Whether or not the scale of the economy is sustainable is also considered to be a positive issue involving biophysical constraints, although normative questions of conservation for the future and other species are not far below the surface. Distributive equity, on the other hand, is a normative issue. This is the main question addressed to distribution: Is it fair, not is it efficient, or is it ecologically sustainable? The question "Is it fair?" is directly and unavoidably normative, and for that reason alone given minimal attention by the positivist tradition of economics.

But like other sciences, economics assumes certain cultural values. First, the very criterion of objective efficiency, Pareto optimality, embodies an implicit normative judgment—namely, that malevolence or invidious satisfactions are not acceptable. If everyone but you becomes better off and you remain no worse off, the Pareto criterion tells us that is an objective increase in social welfare. But if everyone else is better off except you, and you are an envious person, then you will be less happy than before, even though your absolute situation is no worse. Economists must either make the (false) positive judgment that people are in fact not invidious and jealous, or the (true) normative judgment that envy at another's good fortune is a moral failing rather than a welfare loss.

There is a second reason that economics is less positive than some think. Redistribution can be efficient in the sense of increasing total social utility, yet economists make the value judgment that this kind of efficiency should not count. For example, redistributing a dollar from the low marginal utility use of the rich to the high marginal utility use of the poor will increase total utility to society and is in that sense efficient. The Pareto criterion forbids such interpersonal comparisons and summations of utility. Some argue that the major function of the Pareto criterion was precisely to sterilize the egalitarian implications of the law of diminishing marginal utility[1]—a law that economics cannot afford to give up, as we saw in our discussion of demand curves (see Chapter 9).

[1]J. Robinson, *Economic Philosophy*, Middlesex, England: Penguin, 1962.

If we admit interpersonal comparisons of utility, then distribution has efficiency implications as well as fairness implications. The extreme individualism of economics insists that people are so qualitatively different in their hermetical isolation one from another that it makes no sense to say that a leg amputation hurts Smith more than a pin prick hurts Jones. If we are all isolated individuals, we can rule out such obviously realistic human characteristics as envy and benevolence. Man as atomistic individual is the *Homo economicus* of neoclassical economics. Ecological economics' concept of the nature of man is "person-in-community," not isolated atom. Community here means community both with other humans and with the rest of the biosphere.

The Distribution of Income and Wealth

Ecological economics distinguishes between the distribution of income and the distribution of wealth, and between the functional and the personal distribution of income.

Wealth is a stock of assets, measured at a point in time, that is, cash in the bank, plus the market value of bonds, corporate shares, land, real estate, and consumer durables as of a given date. Income is the flow of earnings from these assets, plus the earnings of your own labor power (or human capital), between two dates, that is, over a period of time, usually a year. Labor power is not usually counted as capital because one cannot sell it all at once to another person (short of slavery) but can only rent it for certain durations. Income and wealth are thus two different magnitudes, measured in different units, and distributed differently over the population.[2] Wealth is usually more concentrated than income. And financial wealth is even more concentrated than wealth in general. In 1989, the top 1% owned 48% of financial wealth. Virtually all of the growth in wealth between 1983 and 1989 in the U.S. went to the top 20%. The bottom 80% was excluded from this growth, and the bottom 40% saw their wealth decline in absolute terms (Table 15.1).

Economics has a theory that explains income, as discussed next, and one that explains the prices of assets (though not entirely, as the "price" of entrepreneurship is a residual), but no theory at all to explain the distribution of wealth among individuals. It is the historical result of whose ancestors got there first, of marriage, of inheritance, plus individual ability and effort, and just plain luck.

[2]Wealth is measured in dollars (for example) and income in dollars/time. These magnitudes are as different as miles (distance) and miles per hour (speed).

Table 15.1

U.S. Percentage Share of Wealth and Income by Percentile Group

	Percentile Shares			
Year	Top 1%	Next 19%	Bottom 80%	Gini Coefficient
Net Worth (Wealth)				
1983	33.8	47.6	18.7	0.799
1989	37.4	45.3	16.2	0.832
1992	37.2	46.6	16.3	0.823
1995	38.5	45.8	16.1	0.828
1998	38.1	45.3	16.6	0.822
Income				
1983	12.8	39.0	48.1	0.480
1989	16.4	39.0	44.5	0.521
1992	15.7	40.7	43.7	0.528
1995	14.4	40.8	44.9	0.518
1998	16.6	39.6	43.8	0.531

Source: E. N. Wolff, Top Heavy, The Twentieth Century Fund Report, *New York: New Press, 1995, p. 67 (years 1983–1992) and E. N. Wolff,* Recent Trends in Wealth Ownership, 1983–1998, *Working Paper No. 300, Table 2, Jerome Levy Economics Institute, April 2000.*

The Functional and Personal Distribution of Income

Income distributed among people, regardless of its source, is called the personal distribution. Income is also distributed according to how much of total income goes to wages, interest, rent, and profit—the functional distribution. The idea behind the functional distribution is that income is not first created, then distributed. Rather, it is distributed as it is created among the factors combining to create it.

Remember from the circular flow diagram (see Figure 2.4) that supply and demand in the factors market determine the prices of factors—wages, interest, rent, with profit as a residual. Factor prices times the total amount of each factor used yields the functional distribution, usually expressed as percentage of total income going to landowners (rent), laborers (wages), capitalists (interest), and entrepreneurs (profit). Prices of each factor times the amount of the factor owned by each individual yields the personal distribution of income. The amount of each factor owned by each person, including labor power, is the personal distribution of wealth. Therefore, the personal distribution of wealth times the rental price of each type of wealth asset determines the personal distribution of income.

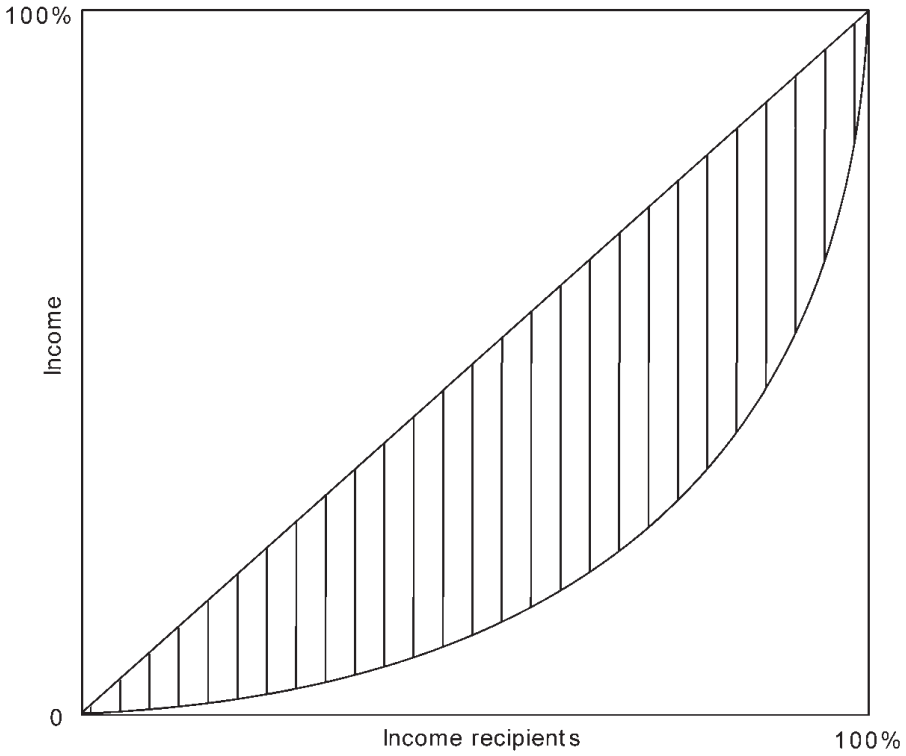

Figure 15.1 • The Lorenz curve. Because the Lorenz curve is in percentages, its shape does not depend on units of measure. It is therefore useful for making comparisons across countries and over time.

Measuring Distribution

Although economists have no good theory by which to explain the distribution of wealth and income, they do have useful ways of measuring and describing it statistically.[3] One useful representation is the **Lorenz curve**, shown in Figure 15.1. The X-axis shows the number of income recipients in terms of cumulative percentages, from lowest to highest income. The Y-axis shows the percentage of total income. The lengths of the axes are equal, so that when closed in, they make a square.

The Lorenz curve plots the percentage of total income going to each percentage of income recipient. We know that 0% of income recipients will get 0% of the income, and that 100% of income recipients will get 100% of the income, so we already know the two extreme points on any Lorenz curve. If each percentage of the population received the same percentage of the income (i.e., the bottom 20% got 20% of total income, the bottom 70% got 70% of income), we would have perfect equality. The Lorenz curve would be the 45-degree line connecting (0, 0) with (100, 100). But suppose the bottom 80% of recipients get 44% of the income.

[3]For a clear and insightful exposition, see J. Pen, *Income Distribution*, New York: Praeger, 1971.

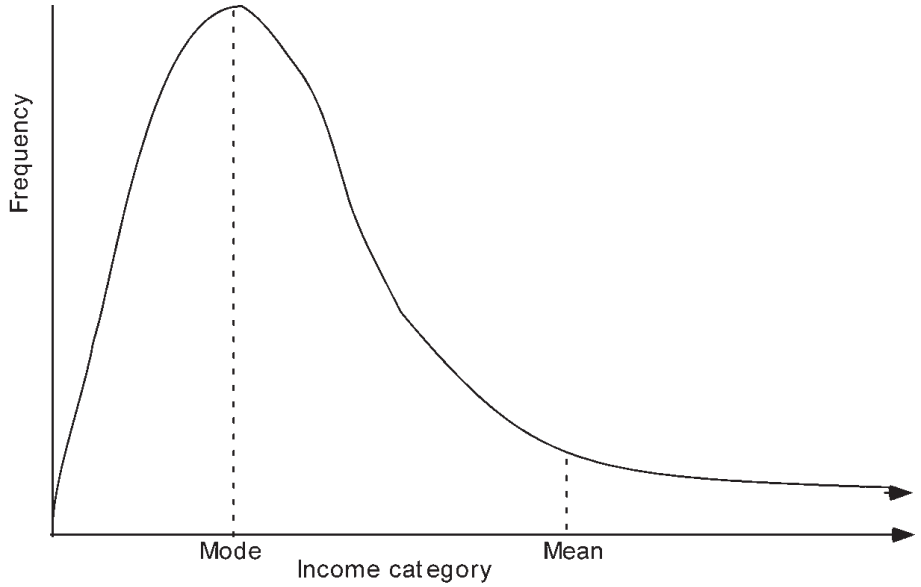

Figure 15.2 • The frequency distribution of income.

That gives us another point, one that lies well below the 45-degree line. If we fill in many points between the extremes, we get a curve shaped like the one in Figure 15.1. The closer the curve to the 45-degree line, the more equal the distribution; the farther away, the less equal. The shaded area defined by the curve and the 45-degree line measures inequality. In the limit, if one person got 100% of income and everyone else got 0%, the Lorenz curve would coincide with the axes and look like a backwards L.

The ratio of the shaded area (between the curve and the 45-degree line) to the total triangular area under the 45-degree line is called the **Gini coefficient**. For perfect equality the shaded area is zero, and consequently the Gini coefficient is 0; for perfect inequality the shaded area takes up the whole area under the 45-degree line, and consequently the Gini coefficient is 1. Values of the Gini coefficient for U.S. wealth and income distribution are given in Table 15.1.

The *Gini coefficient* is used to measure the inequality of the distribution of wealth or income across a population. A Gini coefficient of 1 implies perfect inequality (one person owns everything) and a coefficient of zero indicates a perfectly equal distribution.

A more familiar statistical description is the common frequency distribution, shown in Figure 15.2. The X-axis shows income category and the Y-axis shows number of members in each income category (frequency). Income distribution does not follow a normal distribution, as does height or many other personal characteristics. Rather, it is highly skewed, with the mode well below the mean, and a very, very long tail to the right needed to reach the top income.

If we wanted to show the maximum income on Figure 15.2, we would need a fold-out extending the horizontal axis by the length of a football field. Graphical representations generally do not capture the extreme

inequality at the upper income range. Income categories are frequently truncated at a maximum category of "$100,000 and over," where "over" means four orders of magnitude over.

Moreover, these data are just for the United States. The distribution of wealth and income *between* countries is far greater than that found *within* countries.

What is the proper range of inequality in the distribution of income? Surely it is impossible to have one person owning everything, and everyone else owning nothing. Maybe we could have everyone else getting a subsistence wage, and the fortunate one person enjoying the entire social surplus above subsistence. But most people would not consider that fair, even though possible. At the other extreme, few people think a perfectly equal distribution—a Gini coefficient equal to zero—would be fair either. After all, some people work harder than others, and some jobs are more difficult than others. Fairness in a larger sense would require some income differences. There is a legitimate case to be made that differences in distribution provide a socially useful incentive for industriousness and innovation.

Is there a legitimate range of inequality, beyond which further inequality becomes either unfair or dysfunctional? What might such a range be? Plato thought that the richest citizen should be four times wealthier than the poorest. Ben Cohen and Jerry Greenfield, of Ben and Jerry's ice cream fame, at one time reportedly pledged that the highest paid executive would receive no more than five times the salary of the lowest paid employee. Maybe Plato, Ben, and Jerry were wrong, though, and maybe a factor of ten would be better. Or 20 or 50. Currently the acceptable ratio is not defined, and in 1999 in the United States, the *typical* CEO earned 475 times the *typical* worker.[4] (Of course, comparing the *top* CEOs to the *bottom* workers would yield even a higher ratio.) Ecological economics does not accept the current notion that real total output can grow forever. If the total is limited, then the maximum for one person is implicitly limited. The issue of a proper range of inequality in distribution is therefore critical for ecological economics, even though it has not yet received due attention. The standard economist's effort to keep distribution at bay forever by eternal growth is not a satisfactory solution.

Finally a word on the functional distribution of income. For industrial countries, the division varies around the following: wages = 70%, profits = 20%, interest = 8%, and land rent = 2%. For ecological economics, what is striking is that essentially none of the value of the total product is at-

[4]J. Reingold and F. Jesperson, Executive Pay: It Continues to Explode—and Options Alone Are Creating Paper Billionaires, *Business Week*, April 17, 2000. Online: http://www.businessweek.com/careers/content/jan1990/b3677014.htm.

tributed to natural resources or services. Even land rent is mainly a locational premium, not a payment for resources *in situ* or natural services—one more piece of evidence that the flow of low entropy from nature is treated as a free good. If we think of two social classes struggling to divide the pie, we have laborers getting 70% and capitalists, business owners, and landowners together getting about 30%. This division represents a kind of balance of power in the social struggle. Neither side wants to include nature as a participant in production, which would require that nature's services be paid according to their scarcity and productivity.

Even if one wanted to pay for nature's contribution, who would collect on nature's behalf? There is no social class analogous to labor or capital that has an interest in seeing to it that nature's services are properly accounted and paid for. Historically the landlord class may to some extent have played the role of defender of nature's services, but that class hardly exists anymore, and few lament its demise. The government is the biggest landholder in the U.S., and it has followed a policy of cheap resources in order to benefit and ease the tensions between labor and capital. The existing classes, labor and capital, see it in their mutual interest not to share with a third party. Since in reality there is no third party, all that would be necessary is to pay into a fund a scarcity rent for natural resources and services, and then redistribute the fund back to labor and capital, perhaps on the same 70-30 division. This would get the cost accounting and prices right, and improve the efficiency of allocation, without necessarily affecting the distribution. Alternatively, since many of the goods and services provided by natural capital are nonmarket goods, the scarcity rent could go toward supplying other nonmarket goods. The government could either do this directly or could subsidize the private production of such goods. The rent could also be redistributed progressively by financing the abolition of regressive taxes.

Consequences of Distribution for Community and Health

The existing distribution of wealth is not only a precondition for efficient allocation; it is also a fundamental dimension of justice in society. As such, it affects us more directly than we might at first think. Evidence indicates that inequality of income distribution (independently of absolute poverty) has a substantial effect on rates of morbidity and mortality.[5] The relatively poor have higher incidences of death and sickness than the relatively rich, regardless of the absolute level of income of the relatively poor. The main reason investigators suggest is the extra stress associated with being

[5]See. G. Wilkinson, *Mind the Gap*, New Haven, CT: Yale University Press, 2001.

relatively poor, being at the bottom of a dominance hierarchy. This extra stress is caused by less control over the circumstances of one's life, greater risks of job loss, a lower level of social standing and respect, more frequent experiences of disrespect and shame, with consequent anger and violence. Life at the bottom is more threatening, and the threat often comes from stressful relations with people higher up, including bosses, landlords, and government officials. Stress, of course, has well-known negative direct physiological effects on health.

In addition to these direct effects, inequality has indirect social effects on health. It is more difficult to form friendships across wider income gaps, as well as more difficult to form civic associations when wealth levels and economic interests are very disparate. Lack of friends and civic cohesion is also correlated with ill health. Treating people as atomistic, isolated individuals, unaffected by social relationships, literally makes them sick. As seen from our discussion of Max-Neef's human needs matrix (see Table 13.1), we are persons-in-community, related to each other internally—that is, our personal identity is largely constituted by our relation to others in the community. We are not independently defined entities held together only by external relations of the cash nexus. When these identity-constituting social connections become strained and corrupted by excessive inequality, we get sick more often and we die younger. We are also less happy.

Intertemporal Distribution of Wealth

Every bit as important as the distribution of wealth and income within a generation is the distribution of resources between generations. However, while people have pondered the distribution of resources within a generation for millennia, the concern for distribution between generations is more recent. For the vast majority of human history, natural resources appeared limitless and technological advance was slow. People had approximately the same resource endowment as their great grandparents had enjoyed, and they expected their great-grandchildren to inherit the same endowment as well. As the pace of technological change accelerated with the Industrial Revolution, change became noticeable from one generation to the next, and people began to expect a better life for their children than they themselves had enjoyed. The "Protestant work ethic" asked people to work hard and invest for their children. At least up through the 1960s, the question most economists asked was: How much consumption should this generation sacrifice for the ever-growing well-being of the next?[6]

[6]E.g., J. Robinson, *Essays in the Theory of Economic Growth*, London: Macmillan, 1968; E. Phelps, Second Essay on the Golden Rule of Accumulation, *American Economic Review*: 793–814 (1965).

However, the onset of the atomic age made it apparent that technological advance had the capacity to bring harm as well as good. Growth in population and per-capita consumption raised the specter of resource depletion. Worsening pollution caused alarm, and ecologists began to worry that many systems were nearing irreversible, catastrophic thresholds. The relevant question was no longer: How much should we sacrifice to make the future even better off? Now it was: How much should we sacrifice to keep the future from being worse off than the present? Curiously, at least in the United States, a culture change was occurring at about the same time. The work ethic was no longer "work hard, live frugally and invest in the future" but rather "work hard and consume as much as possible now." As a result, savings rates in the U.S. are currently at historic lows and rapidly approaching zero. For much of the population, savings rates are negative.

Should people strive to make the future better off than the present? Do we have at least an obligation to make sure it is not worse off than the present? There are no easy answers to the "appropriate" distribution of wealth between generations. Even a brief survey of philosophies is beyond the scope of this text. We will, however, quickly examine two alternative approaches: the ecological economics approach, based on ethical judgments concerning obligations to future generations (intergenerational justice), and the more mainstream approach in economics that argues for an "objective" decision-making rule (intergenerational allocation).

The Normative Approach of Ecological Economics

Ecological economists generally take the position that intergenerational resource distribution is an ethical issue. The generation into which someone is born is entirely based on chance. There is therefore no moral justification for claiming that one generation has any more right to natural resources, the building blocks of the economy, than any other. At the very least, future generations have an inalienable right to sufficient resources to provide a satisfactory quality of life. The current generation thus has a corresponding duty to preserve an adequate amount of resources. What is adequate depends on both technological and ecological change, both of which are characterized by pure uncertainty (ignorance). How we choose to deal with uncertainty is also an ethical decision.

What does this mean in practical terms?

Renewable and nonrenewable resources are fundamentally different and must be treated separately. An equal distribution of finite nonrenewable resources among a virtually infinite number of future generations would imply no resource use by any single generation. But there is no point in leaving resources in the ground forever, never to do anyone any good, so an upper limit to exhaustible resources for any one generation

might be determined by the waste absorption capacity of the environment. As long as the use of the resource generates waste no faster than the ecosystem can absorb it, the use of exhaustible resources by one generation will not reduce renewable natural capital. Keeping fossil fuel use within such limits would automatically limit our ability to extract other mineral resources.

Even with a limited ability to extract nonrenewable but recyclable resources, each generation would have a further obligation to either efficiently recycle such resources, or at least minimize the generation and dispersion of "garbo-junk" as much as possible to make such resources as intergenerationally nonrival as possible. If existing technologies makes our well-being dependent on nonrenewable resources—as is currently the case—then we are simply obliged to develop substitutes for these resources. One option would be to capture marginal user costs, the unearned income from nonrenewable resources, and invest them toward developing such substitutes.[7]

Renewable resources as fund-services provide essential life-support functions, and these functions clearly must be maintained. Renewable resources as stock-flows must also be harvested at sustainable levels. No one created renewable resources, and therefore no single generation has the right to reduce the amount of the resource a future generation can sustainably consume, suggesting resource stocks must be at least as large as that which provides the maximum sustainable yield. As we saw in Chapter 12, sustainable management of renewable resources in a manner that "optimizes" both stock-flow and fund-service benefits will in general maintain these resources far from any catastrophic ecological thresholds. It is worth bearing in mind that as nonrenewable resources are finite, the exhaustion of these resources is a finite loss to future generations. Renewable resources, as both stock-flows and fund-services, produce a finite flow over an immeasurable number of future generations, and their irreversible loss therefore imposes a perpetual cost to the future.

The "Positive" Approach of Neoclassical Economics

Conventional economists, in contrast, favor an objective decision rule to determine the intergenerational allocation of resources. The problem thus becomes simply a technical one of comparing future benefits and costs with those that occur in the present. Fortunately, the market can tell us the value of things in the future relative to things today, and therefore the market can solve the problem of intergenerational allocation.

[7]For practical guidelines on investing scarcity rents, see S. El Serafy, "The Proper Calculation of Income from Depletable Natural Resources." In Y. J. Ahmad, S. El Serafy, and E. Lutz, eds., Washington, DC: *Environmental Accounting for Sustainable Development*, World Bank, 1989.

Intertemporal Discounting. How does the market reveal future values? Where adequate financial markets exist, people can borrow money today at interest, which requires them to pay back more money in the future. The fact that people engage in this activity reveals that people prefer things now rather than in the future, and economics must respect people's preferences.

There are three basic reasons why people might prefer things now to things in the future. First, people may simply be impatient. Anyone who goes into interest-bearing debt to purchase something is willing to sacrifice a greater quantity in the future for a smaller quantity now. Some of this impatience may come from uncertainty—no one knows for sure if he or she will be alive tomorrow, so why not eat, drink, and be merry today? This rationale for discounting is known as the **pure time rate of preference (PTRP)**.

Second, for things that reproduce, it makes sense that a given quantity in the future would be worth less than a given quantity now. For example, a handful of seed corn now can become a bushel of marketable corn a few months from now, so if growing corn was risk-free and required no resources or effort, then a handful now would be at least as valuable as a bushel in a few months. Of course, growing corn is risky and requires land, labor, and resources. However, market goods (in this case, seed corn) can be sold for money. Investing the money earned from the sale of the seed corn in an insured bank is not very risky, and for the individual investor basically requires no further resources or labor. As we explained in Chapter 10, this rationale for discounting is known as **opportunity cost**, the lost opportunity to invest. If money is a substitute for any other resource, then we should give more weight to any resource today over the same resource tomorrow.

Third, the economy has grown fairly steadily for hundreds of years. People therefore expect that they will be richer in the future than they are today. Just as an extra $1000 provides less utility to Bill Gates than to a pauper, the law of diminishing marginal utility means that money in the future will be worth less than the same amount of money today. This is sometimes referred to as the "richer future" argument.

In general, this process of valuing the future less than the present is known as **intertemporal discounting**, introduced in Chapter 10. Business people explicitly discount the future when making investment decisions, and mainstream economists argue that people automatically apply this concept to all of their purchase decisions. As a result, they conclude, the market efficiently allocates goods between the present and future.

What's more, if intertemporal discounting leads to allocative efficiency in the market, then it should also be applied to nonmarket investments. For example, one of the biggest nonmarket decisions we face today is how

to deal with global climate change. Virtually all economic analyses of climate change place a lower weight on future costs and benefits than on present ones. These analyses look at different policy scenarios and for each sum up the present costs and benefits with discounted future costs and benefits to arrive at a **net present value (NPV)**. NPV basically tells us what present and future costs and benefits are worth to us *today* (not to the future *tomorrow*), which implies that future generations have no particular right to any resources, and we have no obligation to preserve any. Under this type of benefit-cost analysis, the higher the NPV relative to required investments, the better the project.

Such analyses can carry a great deal of weight as society decides how to address some of the most pressing problems we now confront. The central importance of the discount rate in determining the outcome of such analyses means the topic deserves our attention.

Box 15-2 Intertemporal Discounting and Global Climate Change

Policy makers seeking an objective decision-making tool for resolving the problems of global climate change have turned to economists. Economists typically respond to the problem by creating complex models of future costs and benefits of climate change, and compare these to the costs of mitigation measures in a cost-benefit analysis designed to calculate net present value. Not surprisingly, analyses using a fairly high discount rate find that future damages from global warming do not justify efforts today to reduce greenhouse gases. The 6% discount rate used in one study would have us believe that we should not invest $300 million today to prevent $30 trillion (a rough estimate of today's global GNP) in damages in 200 years.[a] A similar study using a 2% discount rate in contrast finds that we should make substantial investments now to reduce the impacts of global warming in the future. Similarly, what we decide to do with an old-growth forest that supplies a small but steady flow of benefits forever if left intact or a large, one-time return if it is clear-cut will depend on the discount rate we choose.

A frequently asked question is: Does a higher or lower discount rate favor the environment? For a given fishery or mine, as we have seen, higher discount rates increase the intensity and rate of exploitation and are thus bad for the environment. But a higher interest rate (discount rate) slows down aggregate growth in GNP and throughput, thus easing pressure on the environment. In terms of evaluating a given project, a high discount rate favors projects whose costs are mainly in the future and whose benefits are in the present, and penalizes those who costs are in the present with benefits in the future. Most issues in economics are not simple, and that certainly holds for discounting.

In many such models, the choice of a discount rate may be the single most important factor, yet respected economists addressing the same problem use dramatically different rates and arrive at dramatically different results. Are such models actually objective decision-making tools?

[a]30 trillion is a number that's hard to wrap your mind around. Putting it into perspective: 30 trillion seconds is slightly less than a million years (951,294 years, to be precise)!

Discounting Reconsidered

We have already explained why intertemporal discounting can make sense for the individual and for market goods. We must now examine whether it also makes sense for society and for nonmarket goods.

We saw why individuals might have a pure time rate of preference: People are impatient; they don't live forever; possessions can be lost, destroyed, or stolen, and opportunities disappear. A reasonable individual may discount the future for any one of these reasons—why should I pay money now to reduce damages from global warming that will only occur after I am dead?—but the same logic does not apply to society. Relative to the individuals of which they are composed, societies are immortal, and uncertainties are averaged out. For this reason, there is, in fact, fairly wide consensus within the economics profession that social discount rates should indeed be lower than individual discount rates. The **social discount rate** is a rate of conversion of future value to present value that reflects society's collective ethical judgment, as opposed to an individualistic judgment, such as the market rate of interest.

When it comes to the opportunity cost of capital, however, the consensus changes. Financial capital does function as a productive asset, and if we have it now instead of in the future, we have the opportunity to invest it in productive activities that will increase the quantity of market goods in the future. There are a number of important issues we must bear in mind, however.

First, the real value of money can only grow if the production of goods and services that money can acquire also grows, and we know that the production of goods and services cannot grow forever on a finite planet. While there may always be some areas that are growing, justifying a discount rate for the individual, the economy as a whole cannot grow indefinitely, in which case a social discount rate into the indefinite future may be inappropriate.

Second, we must recognize that many investments are "profitable" because we ignore many of the costs of production. We know that all human productive activities use up natural resources and return waste to the environment, and these costs of production are often ignored. Many of these

costs, such as contributions to global warming, have greater impacts on future generations. Thus, ignoring costs to future generations allows us to earn higher returns on investments. We then use these higher returns to justify the fact that we ignore costs imposed on future generations! Even in the short run, then, it seems that market-determined interest rates are not suitable discount rates.

Related to the opportunity cost of capital is the argument that the future will be richer than the present because of investments we make now. Of course, if the economy does not continue to grow, the future will not be richer, and if we deplete our natural resource stock, there is every chance the future will be poorer. In fact, measures such as the ISEW suggest that society is already growing poorer, not richer, if we take into account external costs. Also, if we believe that natural capital must be treated separately from manmade capital (because they are complements rather than substitutes and natural capital has become the limiting factor), then the decline in natural capital, coupled with the law of diminishing marginal utility, suggests we should apply a negative discount rate to natural capital. At the very least, we might consider applying a positive discount rate only to those goods and services that are actually highly fungible with money—that is, that can be converted into money and back again with little effort. Basically, this would mean that we should only discount market goods and services.

Third, there are only finite opportunities for productive investment in an economy; investments, like other things, show diminishing marginal returns. For example, someone borrows money to explore for oil, and someone else borrows money to build a car factory. The next person to borrow money to explore for oil will have fewer places to explore, and therefore will expect lower returns. The next person to borrow to build a car factory will face a more saturated market, and therefore expect to sell fewer cars. As more and more people borrow to invest and opportunities are used up, the returns on investments can be expected to decrease, ultimately falling to zero. More likely, if interest rates are determined in the market by the supply and demand of money for borrowing and returns on investments, a balance will be reached in which investors cannot afford to pay high enough interest rates for consumers to be interested in deferring consumption. Theoretically then, in a perfect market situation, the opportunity cost of capital at the margin will just equal the PTRP of the existing generation. (Obviously, future generations cannot take part in financial markets any more than they can in any other market.) However, a high PTRP means consumption will be high and investment and growth low,[8] and a low PTRP implies the op-

[8]Returns on investments will be high, but the total amount invested will be low, and hence growth will be low.

posite. Thus, if we allow the market interest rate to determine the discount rate, there would theoretically be an inverse correlation between discount rate and economic growth, the exact opposite of what would justify the "richer future" rationale for discounting.

Box 15-3 Discounting, Psychology, and Economics

Economists argue that economics is the science of human preferences, so human preferences must be respected. If people value the present more than the future, we must respect that. The question is: Do people *exponentially* discount the future? While it is true you may prefer to have something now rather than the same thing in 5 years, how do you value something that happens 100 years in the future compared to 105 years? If you heard that global warming was going to result in the deaths of 50 million Bangladeshis in 125 years, would that make you feel only half as bad as finding out it would actually kill those 50 million Bangladeshis in 100 years? If you are like most people, you would feel equally bad in either case, yet influential economic models of the impacts of global warming really do assume we would care only half as much about those deaths if they occurred 25 years later.

Empirical studies show that people do discount the future, but do not do so exponentially. We might give more weight to what happens now than to what happens in the near future, but we are nearly indifferent between the same outcome occurring at different times in the more distant future. One approach to modeling this type of behavior mathematically is known as **hyperbolic discounting**. While this precise formulation of intertemporal discounting may not be perfect, considerable evidence suggests it is far more representative of human preferences than exponential discounting. While the approach was first introduced over 30 years ago and has gained increasing attention in the last few years, it is still fairly rare to see it in use.

An increasing number of studies in the area of Psychology and Economics are finding that the traditional economic assumptions of human behavior are often seriously flawed. If economics is serious about becoming the science of human preferences, it would do well to pay more attention to how humans really behave.[a]

[a] *For a good introduction to the field of Psychology and Economics, see M. Rabin, Psychology and Economics,* Journal of Economic Literature *36(1): 11–46 (March 1998).*

Finally, many economists argue that technology is the driving force for economic growth. Not only does technology ensure we won't run out of resources and the economy won't stop growing, but it offers yet another reason to discount the value of resources in the future. Technology is

likely to develop substitutes for natural resources. When these substitutes become available, the resources they replace will lose value. Therefore, they will be worth less in the future than they are today. After all, hasn't oil largely replaced coal, and haven't fiber optics replaced copper in many uses? However, technology ultimately complements resources and can never completely replace them. Some 150 years ago, oil had little value. Today, it is an integral part of an overwhelming number of industrial processes and products. As we saw in Chapter 5, we are actually developing new uses for oil and other raw materials faster than we are developing replacements, again suggesting that the value of raw materials will increase in the future, not shrink.

What in the end can we say about discount rates? They do make sense for individuals in the short run. For some small-scale, short-term social projects, they may also make sense. However, justifications for discounting the future on a large scale and over long-time horizons are questionable at best. **Intertemporal allocation** is the apportionment of resources across different stages in the lifetimes of basically the same set of people (the same generation). Discounting can make sense for someone efficiently allocating resources intertemporally. But as we lengthen the time period we are more and more talking about different people (different generations) and less and less about the same people at different stages in their lives. **Intertemporal distribution** is the apportioning of resources across different generations, across different people. Distribution is fundamentally different from allocation, and, consequently, justice replaces efficiency as the relevant criterion for policy when time periods become intergenerational.

BIG IDEAS to remember

- Pareto optimality
- Role of scale and distribution in defining Pareto optimal allocation
- Income distribution vs. wealth distribution
- Functional vs. personal distribution
- Social limits to range of inequality
- Lorenz curve
- Gini coefficient
- Inequality and health
- Intertemporal distribution vs. intertemporal allocation
- Discounting and net present value
- Pure time rate of preference (PTRP)
- Individual vs. social discount rates

CHAPTER

16

The IS-LM Model

We have now explored three of the major issues with which macroeconomics is concerned: gross national product (GNP), money, and distribution. We questioned the appropriateness of GNP as the desirable end for economic policy, and emphasized the importance of a just distribution as a desirable end, but said little about policies for attaining these ends. In this chapter, we examine the policy tools at the macroeconomist's disposable that can help us attain an economy with sustainable scale, just distribution, and efficient allocation.

Of course, to know how policies work, we have to know how the macroeconomy works. One way of doing this might be to build on microeconomic principles to construct a model in which supply and demand of all goods and services balances simultaneously. This approach would extend the basic market equation presented in Chapter 8—MU*xn**MPP*ax* = MU*yn**MPP*ay*—into a general equilibrium model encompassing all goods (*x, y, z* . . .), all commodities (*a, b, c* . . .), and all consumers (*n, m, o* . . .). Such a model can easily become overwhelming. A thousand simultaneous equations with a thousand unknowns is hard to come into mental contact with. It does show that everything depends on everything else, which is interesting and usefully humbling, but it is also crippling from a policy perspective to have to face the implication that in order to predict anything, you have first to know everything. But a smaller system of two or three or five especially important aggregate sectors interacting through two or three or five simultaneous equations that reflect key behavior can aid the understanding and give basic policy insights. This is the kind of model that most macroeconomists have sought. They still look at the whole economy, but they divide it into fewer but larger aggregate

sectors than does the general equilibrium model of microeconomics.[1] A model of this type, first offered in 1937 by Sir John Hicks[2] and now called the IS-LM model, has proven to be a good "two-digit" compromise between completeness and simplicity. It has become the "workhorse" model in macroeconomics. Below we will explain this model, and then discuss its applications to ecological economics.

The model divides the economy into two sectors: the real sector (dealing with national income, savings, investment, rates of productivity of capital, government spending, taxation, etc.) and the monetary sector (money supply, interest rates, demand for liquid cash balances). The real sector reflects the theories and insights of classical economics, and the monetary sector reflects the insights of John Maynard Keynes, which in 1937 were still quite new. The model seeks to explain how the interdependent behavior of consumers and savers, lenders and borrowers, and monetary authorities interact to determine the level of national income and the rate of interest.

Box 16-1 The Quantity of Money Theory of Income

Another way of relating the real and monetary sectors in an aggregate way is through the "identity of exchange," MV = PQ, where Q is quantity of final commodities sold to households, P is average price of exchange, M is stock of money, and V is velocity of circulation of money (number of times an average dollar is spent per year on final goods and services). Since by definition V = PQ/M, the equation of exchange is an identity or truism. To the extent that V is a constant or slow to change, reflecting stable payment habits and settlement periods, the identity becomes the "quantity of money theory of income," stating that changes in PQ are proportional to changes in M. If the economy is at full employment, it will be very hard to increase Q in the short run, and the change in PQ will be mainly a change in P—i.e., inflation. Historically M and P have often moved in direct proportion, yielding a quantity of money theory of the price level.

[1]Our measures of the two most basic magnitudes of macroeconomics, GNP and money, are too dialectical and uncertain to be able to support exact calculations implicit in complicated models. As Oskar Morganstern remarked in his classic *On the Accuracy of Economic Observations*, "economics is a two-digit science."

[2]J. Hicks, Mr. Keynes and the "Classics," *Econometrica* 5(2) (April 1937).

IS: The Real Sector

Let's begin with the real or classical sector. The real sector is in equilibrium when the supply of goods by firms is just equal to the demand for goods by households (the lower half of the circular economy in Figure 2.4). Of course, the demand for goods by households is determined by their income—the money firms pay households for their factors of production (e.g., labor), and the supply of goods is determined by the firms' employment of those factors of production (the upper half of the circular economy in Figure 2.4). In equilibrium, income (Y) equals output (GNP). Remember from the circular flow diagram in Figure 2.5 that the equilibrium condition for the continued flow of national income at a given level is that leakages equal injections. In the simplest case, the leakage is savings (S) by households, the new injection is investment (I) by firms. Therefore, the equilibrium condition for the real sector is S = I.

But how do S and I get determined? Let r be the interest rate and Y be national income (GNP). In equilibrium, income paid to the factors of production will just equal the output of goods provided by those factors of production, and the income will be used to purchase the output. Savers (i.e., households) will save more if their income Y is higher than if it is lower. Also savers will save more with a higher interest rate r than with a lower one. Investors (i.e., firms) will borrow and invest more if the interest rate is lower, and if income is higher. In other words, savings is some function of the interest rate and national income, say S = S(r, Y). Likewise, investment is some different function (representing the behavior of firms instead of households) of the same two variables, say I = I(r, Y).

In equilibrium,

$$S = I$$

or

$$S(r, Y) = I(r, Y)$$

The above equation is satisfied for all combinations of r and Y such that S = I, that is, such that savers and investors are both satisfied.

There are many such combinations of r and Y—we have only one equation with two unknowns. Plotting all the combinations of r and Y that result in S = I gives us Hicks' so-called IS curve, short for I = S (Figure 16.1). To reiterate, this is the combination of r and Y that leads to equilibrium in the real sector: leakages (savings) equal injections (investment), and the demand for goods is just equal to the supply.

Why is the IS curve drawn with a negative slope? Businesses will only borrow money to invest if they can make sufficient returns from the investment to pay off the loan plus interest and still have money left over for profit. A businessperson would not borrow money at 6% interest to invest

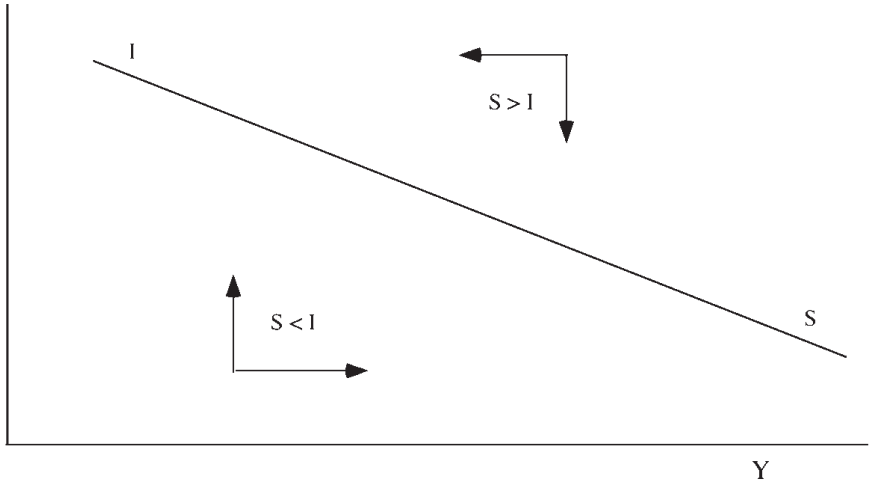

Figure 16.1 • The IS curve: At low (high) levels of income Y, there is a correspondingly low level of savings. At high (low) rates of interest r, there is a low (high) demand for investment. Therefore, at low (high) levels of Y, savings and investment will only be in equilibrium when r is high (low). If interest rates are too high for a given level of income, savings (leakages) will be greater than investment (injections). Firms producing more goods than people consume reduce production and the economy shrinks. Firms with excess capacity borrow less, so the price of borrowing (the interest rate) falls to clear the market. The converse is true when investment is greater than savings.

in a project expected to return 5% annually on the investment, but would borrow at a rate of 4%. As interest rates go down further, more and more investments become profitable, and therefore more investments are made. More investment leads to higher Y. Therefore, high interest rates lead to low rates of investment and low income, while low interest rates lead to high rates of investment and high income. Savings, in contrast, are probably determined more by income than by interest rates.[3] When income is low, all money has to be spent simply to meet basic consumption needs, and none is available to save. As income increases, basic consumption needs require a smaller percentage of income, and more is left over to save, so in general higher incomes lead to greater savings.

Combining these two tendencies, we would expect that at high levels of income when lots of money is being saved, investors will only borrow all that money to invest if interest rates are low. At low levels of income, savings are low, and unless interest rates are high, businesses will demand more money than is being saved. For some readers, a diagrammatic explanation for the negative slope of the IS curve will be easier to follow (Figure 16.2).

[3]Savings rates should also increase as interest rates increase, as under these circumstances savings yield higher returns, and consumption has a higher opportunity cost. But empirical evidence does not support this. One reason may be that if savers are motivated by attainment of a target future amount, a higher interest rate would mean less saving is needed to reach the future target.

Box 16-2 A Graphic Derivation of the IS curve

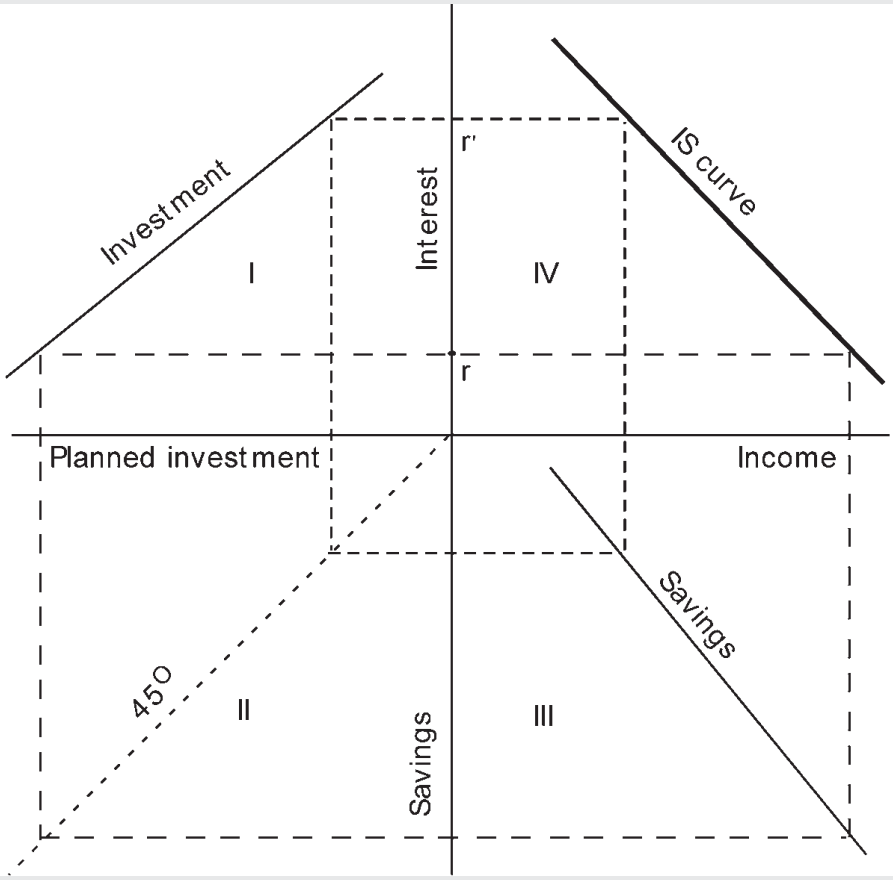

Figure 16.2 • A graphic depiction of the derivation of the IS curve.

Figure 16.2 illustrates one way to derive the IS curve. Quadrant I shows the basic relationship between interest rates and investment—high interest rates lead to low levels of investment, and low interest rates lead to high levels of investment. There is a negative correlation between interest rates and investment, as depicted on the graph. Quadrant III shows the relationship between savings and income. Poor people must spend all their income to meet their basic needs and cannot afford to save anything. As income increases, people begin to save, so there is a positive correlation between income and savings, as depicted on the graph. We know that in equilibrium (which is what the IS curve depicts), investment equals savings. Quadrant II contains a 45-degree line that allows us to translate a given rate of investment from quadrant I to an identical rate of savings in quadrant III. Quadrant IV shows the relationship between income and interest in a real sector equilibrium.

If we start with interest rate r, we can see from quadrant I that this

will correspond to level of investment I. Dropping a line down from I in quadrant I to the 45-degree line then across to quadrant III lets us determine the equilibrium level of savings S in quadrant III. We can see from quadrant III that this level of savings corresponds to income Y. The point in quadrant IV where income Y meets interest rate r gives us one point on the IS curve. If we do the same for interest rate r′, we have two points on the IS curve. We see that a low level of interest leads to equilibrium only when income is high, and a high level of interest leads to equilibrium when income is low. Perhaps the simplest way to remember this relationship is that at low interest rates, investment will be high, and high investment leads to high income.

Macroeconomics does not assume that the economy is always in equilibrium, but it does assume that it is at least moving in that direction. For example, we know that if r rises, then savers will try to save more, and investors will be less willing to borrow and invest, leading to a condition in which planned S > I. In other words, savers want to save more than investors want to invest at the new higher r. This will have two impacts. First, leakages will be greater than new injections, causing income to fall. Second, savers earn interest on their savings because investors are willing to pay that interest to borrow the money. Interest is the price of money. When the supply of savings is greater than the demand for savings, the interest rate must fall. The mechanism is the same as for any other good, as explained in Figure 9.2. At a lower Y savers save less, and at a lower r investors borrow more, and both r and Y continue to fall until I again equals S at a lower income (Y) and higher interest rate (r) than before. If the interest rate falls, then investment will become greater than savings, and adjustment will occur in the opposite manner. These dynamics are indicated by the arrows in Figure 16.1.

LM: The Monetary Sector

We turn now to the monetary sector and the LM curve, which shows the levels of income (Y) and interest (r) at which the demand for money balances (money held by people) equals the supply of money. We must first ask why individuals want to hold money balances when they could easily exchange them for real assets. From our earlier discussion the answer is clear—people hold cash balances to avoid the inconvenience of barter. Keynes referred to this as the **transaction demand for money**. He also spoke of a related **liquidity preference**, meaning that, other things being equal, people prefer liquid assets to "frozen" assets because they are so easily convertible into anything else, therefore fungible. Money is the most liquid of all assets. But of course other things are seldom equal, and the cost

of holding wealth in the form of fungible money is to forego the interest that could be had by lending the money, or the utility from spending it on a real asset or commodity. Yet if too much of your wealth is tied up in nonliquid forms, you will have difficulty making necessary transactions in a timely manner and meeting unexpected contingencies. The higher the national income, the more need for transactions and consequently the more money everyone will need (a higher transactions demand for money), and the higher the interest rate will have to be to induce owners of those transactions balances to sacrifice liquidity by lending them.

The demand for money balances (DM) thus depends on r and Y, by means of a relation of liquidity preference (L). Thus:

$$DM = L(r, Y)$$

The equilibrium condition is that the demand for money equals the supply of money (SM):

$$DM = SM$$

What determines the supply of money? In earlier times it was the geology and technology of gold or silver mining (a part of the real sector!), but today we have not real commodity money, but fiat or token money, controlled by the government through the private banking sector, as discussed in Chapter 14. For simplicity, the model usually takes SM as given by the government, equal to M. Thus:

$$L(r, Y) = M$$

is the equilibrium condition for the monetary sector, and the LM curve consists of all those combinations of r and Y such that the aggregate demand for cash balances is equal to the given money supply (Figure 16.3).

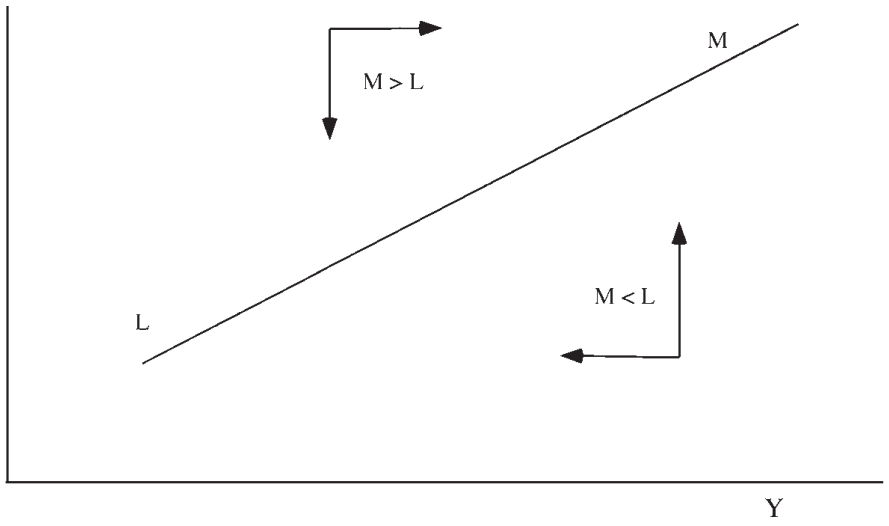

Figure 16.3 • The LM curve.

Since we have one equation with two unknowns, we cannot get unique values of the unknowns, but we can determine all those combinations of r and Y that satisfy our one equation.

Box 16-3 The Federal Reserve Bank

In the U.S., money is *not* controlled by the democratically elected government but rather by the Federal Reserve Bank (the Fed), a nonelected "branch" of government. Decisions concerning monetary policy are decided upon by a seven-member board of governors with lesser influence by the directors of the 12 regional Federal Reserve Banks. Members are appointed by the President (with Senate approval) for 14-year staggered terms, and the chair and vice chair are appointed for 4-year terms. Despite the importance of monetary policy in the functioning of our economy, the system is specifically designed to insulate the Fed from pressure by democratically elected politicians! The Fed is not expected to respond to voters. This does not mean that the Fed does not have a constituency to which it feels responsible, as we shall discuss later.

Why is the LM (short for L = M) curve drawn with a positive slope? Let's ask ourselves what are the consequences on the interest rate (r) of an increase in income (Y). A larger Y means a larger volume of transactions and will cause a greater demand for transactions balances. This will lead to a higher r to compensate for the loss in liquidity from lending those balances. Thus, a higher Y will require a higher r for money holders to again be satisfied (for L to equal M). Hence the positive slope of the LM curve. This relationship seems to be sufficiently clear that a more detailed graphic explanation is unnecessary.

When the monetary sector is out of equilibrium, what specific mechanisms drive it toward equilibrium? Say the monetary authority increases the money supply, so there is more money available than people actually desire to hold at the existing interest rate—that is, M > L. Excess money is used to buy bonds and other nonliquid interest-bearing assets (which we will refer to jointly as "bonds" for convenience). More money chasing the same number of bonds will drive up their price.

There are many types of bonds, but in the simplest case, when someone buys a bond, they are paying something now to receive a fixed amount when the bond matures. For example, if I pay $50 today for a $100 bond that matures in 10 years, my rate of return is about 7.2%. An increase in the money supply might drive the price of the bond up to $60, which provides a rate of return of only 5.24%. The higher the price for a bond, the

lower the interest rate on that bond. Hence, an increase in the supply of money increases the demand for bonds and drives down the interest rate.

At lower interest rates, there is less opportunity cost to holding money, and hence a higher demand for money. Lower interest rates also stimulate investment, leading to economic growth, which further stimulates the demand for money. The result is a new equilibrium at lower interest rates and higher income. A decrease in the money supply of course leads to the opposite result. These forces are illustrated by the arrows in Figure 16.3.

Combining IS and LM

Putting the IS and LM curves together lets us determine a unique combination of r and Y (namely r*, Y*) that satisfies both the S = I condition of the real sector and the L = M condition of the monetary sector (Figure 16.4). The point of intersection is the only point common to both curves, the only point that gives equilibrium in both real and monetary sectors. Basically we now have two simultaneous equations determining two unknowns, r and Y.

The **IS-LM model** is used in a comparative statics[4] way to analyze the effect on r and Y of changes in the underlying determinants—namely,

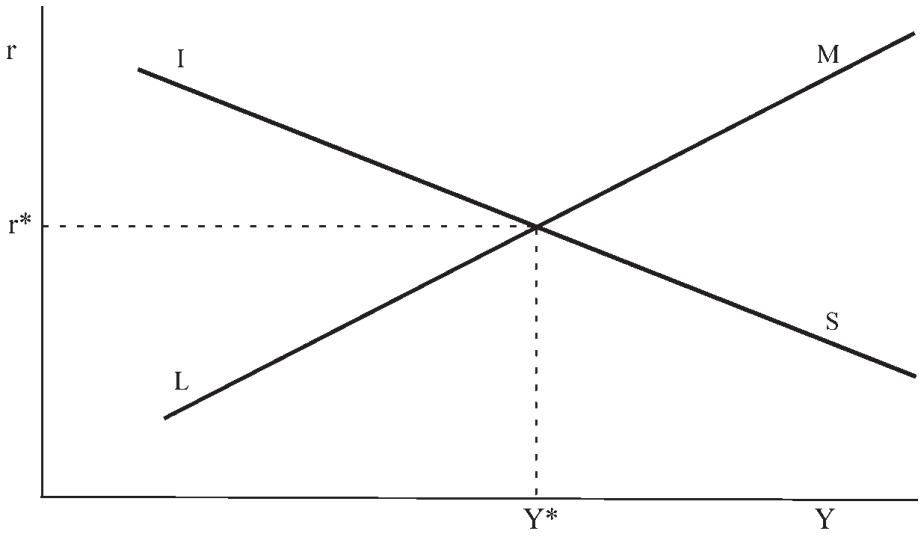

Figure 16.4 • The IS-LM model.

[4]Comparative statics is the analysis of what happens to endogenous variables in a model (in this case, r and Y) as a result of change in exogenous parameters (in this case, propensity to save, efficiency of capital investment, and liquidity preference). It compares the new equilibrium variables with the old ones, without explaining the precise dynamic path leading from the old to the new equilibrium.

propensity to save, the efficiency (productivity) of capital investment, and liquidity preference. Of particular interest to policy makers is the impact of policy variables on r and Y—namely, government expenditure, taxation, and the money supply. Each of these changes results in a shift in one of the curves, and consequently in a move along the other curve to a new intersection point. What we are really interested in, then, is how the economy moves toward equilibrium after policies or outside (exogenous) changes push it away.

Exogenous Changes in IS and LM

First let's look at some exogenous changes, those that are basically independent of fiscal and monetary policy, and therefore outside the IS-LM model. Consider an increase in the marginal propensity to save. Such a change in savings rate might result from fears of an economic downturn that would lead to lower wages and greater unemployment. (We might hope that people might one day simply decide to consume less in order to protect the environment!) In either case, people decide to save more and spend less of their extra income. This means that now $S > I$ for all the combinations of r and Y on the IS curve. We need a new IS curve for which $S = I$ again. If people save more at every r, this means $S > I$, or leakage greater than injection, so the flow of income will fall to the level at which $S = I$ again.

Even though people are saving a larger fraction of their income, they will end up having a smaller income out of which to save, with the result that S will be the same, only now the product of a higher rate of savings per dollar times a lower number of dollars in income. The very act of saving more and spending less will have caused a fall in aggregate income to the extent that the lower income times the higher fraction saved gives the same total savings, equal to the unchanged level of investment. So every r will be paired with a smaller Y in the new IS than with the old one. The IS curve will have shifted to the left. The new intersection with LM will occur at a lower r^* and lower Y^* than before. An increase in the marginal propensity to save will therefore result in a fall in national income and a fall in the interest rate.

Of course if the marginal propensity to save increases, then the marginal propensity to consume must decrease. As people consume less, businesses will be unable to sell their goods, leading to unplanned accumulations of inventory. This in turn will lead businesses to reduce planned investment and production—perhaps laying people off. Unemployment resulting from layoffs further decreases consumption, requiring another round of adjustment, lowering Y still more. The final result, when $S = I$ again, may well be that S will be lower than the level at which

we started. Thus, the effort of everyone to save more in the aggregate could result in everyone actually saving less—the so-called **paradox of thrift**. In such a case, a higher savings rate induced by fear of recession could itself cause a recession—a self-fulfilling prophecy.

Now suppose there's an increase in the efficiency of investment (an increase in the marginal productivity of capital), thanks to a new invention. For example, many people claim that this is exactly what has happened in today's "new economy," in which information technology is said to have increased productivity. This would increase I, so that $I > S$ now along the old IS curve. With $I > S$, injections are greater than leakages out of the circular flow, so the flow of income will grow until $S = I$ again. The new IS curve will have a higher Y for each r. The curve shifts rightward. The new equilibrium occurs with a higher Y^* and a higher r^*. An improvement in the marginal efficiency of capital raises both income and the interest rate.

Finally, turning to the LM curve, suppose there was an increase in liquidity preference, so that $L > M$. Such a change could result from increasing uncertainty over future economic conditions, and a desire by people to be prepared for the unforeseen with cash on hand. Alternatively, the deregulation of banking in the United States during the mid-1970s allowed certain checking accounts to pay interest. This reduced the opportunity cost of holding money, and therefore probably increased the liquidity preference as well. In either case, for any income and associated level of needed transactions balances, there is a greater willingness to hold those balances, to hold more than strictly needed. It takes a higher r to induce holders of money to lend. Consequently each level of Y will be associated with a higher r on the new LM than on the old one. The new LM will shift upward. The new equilibrium will occur at a higher r^* and lower

Box 16-4 Junk Bonds and Timber Companies

Seemingly abstract things like interest rates on bonds and Wall Street transactions can affect real economic production and the provision of environmental services. For example, during the 1980s, hostile takeovers and the introduction of junk bonds on Wall Street led to deforestation on the West Coast. How did this happen? Mergers, when two companies join together, and acquisitions, when one company purchases another, are a normal part of corporate activity in the U.S. Sometimes, however, one company does not wish to be taken over by another. For example, mergers and acquisitions (M&A) focused primarily on short-term profits

can weaken or destroy the company being acquired, leading to massive layoffs.

Reasonably enough, managers not eager to be laid off will be opposed to a merger, and under such circumstances, takeover attempts are "hostile." One company acquires another through the purchase of a controlling share of stocks. As soon as someone starts buying enough stock to control a company, the stock price rises. A company threatened by a hostile takeover can attempt to defend itself by repurchasing its own stock, driving up the price of stock even further. To get enough money for a hostile takeover, the company attempting the takeover can offer high-yield, high-risk bonds known as "junk bonds" in Wall Street jargon.[a]

The best target for takeover is a company that has lots of assets in a nonliquid form that can be liquidated after takeover to pay off the junk bonds, but that cannot be sold quickly to defend against the takeover. Timber companies have valuable assets in the form of forests that can be liquidated after takeover, but cannot be sold quickly to buy back stock and prevent a hostile takeover. This made them popular takeover targets during the 1980s.

A classic example is Charles Hurwitz's acquisition of Pacific Lumber in the mid-1980s. Pacific Lumber fought the takeover, but using a combination of junk bonds and short-term loans, Hurwitz won out, acquiring with the company's 196,000 acres of forest, including the largest unprotected stands of virgin redwood in the world. Hurwitz was saddled with an enormous debt and crushing interest payments. To repay the debt, Hurwitz liquidated much of the forest stock, including many old-growth redwood groves. Some illegal cuts were conducted on weekends and holidays to avoid state regulators. Wall Street innovations during the 1980s accelerated the decimation of the nation's last remaining virgin forests, and of the environmental services those forests once provided.[b]

---

[a] *Different companies (and cities and countries) have different credit ratings based on their financial soundness. Bonds from financially sound companies are themselves very sound, bonds from less sound companies are not, and the risk of default is higher.*

[b] *N. Daly, "Ravaging the Redwood: Charles Hurwitz, Michael Milken and the Costs of Greed,"* Multinational Monitor, *16(9) (1994). Online: http://multinationalmonitor.org/hyper/issues/1994/09/mm0994_07.html.*

Y*. An increase in liquidity preference raises the interest rate and lowers national income.

The above analysis of changes in propensity to save, efficiency of investment, liquidity preference, and money supply is summarized in Figure 16.5.

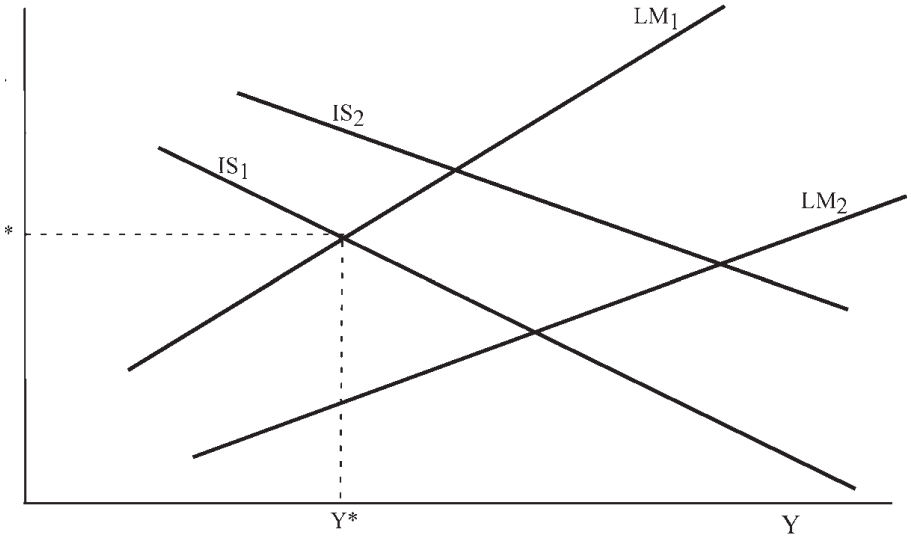

Figure 16.5 • Shifting of the IS and LM curves. The shift from IS_1 to IS_2 could be the result of either a decrease in the marginal propensity to save or an increase in the marginal efficiency of investment. The shift from LM_1 to LM_2 is caused by either a reduction in liquidity preference or an increase in the money supply. Can you work out the changes in r and Y resulting from an increase or decrease in each of the four parameters, others remaining constant? What about changes in two or more parameters at the same time?

IS-LM and Monetary and Fiscal Policy

Changes in the propensity to save, the efficiency of investment, and liquidity preferences are not brought about directly by policy interventions; they are affected by psychology and technology and as a result are difficult, if not impossible, to predict. However, policy makers do have two sets of economic levers by which they can influence these variables: monetary policy and fiscal policy.

What does the IS-LM model tell us about different monetary and fiscal policy levers? The analysis of monetary and fiscal policy in macroeconomics can be worked out by tracing the effects on IS or LM of changes to the money supply and of government taxing and spending.

Monetary policy basically affects the money supply. When the monetary authority (the Federal Reserve, in the U.S.) increases money supply, the LM curve shifts downward and M > L, which drives interest rates down, as explained earlier. Lower interest rates stimulate the economy, and income grows. If the money supply is increased by too much, there can be too much money chasing too few goods, and inflation threatens. Reducing the money supply drives interest rates up, shrinks the economy, and can help control inflation.

Fiscal policy is basically government expenditure and taxation. When the government spends money, industry has to produce more goods and

services to meet the increased demand. This drives up income and also increases the demand for investments, driving up interest rates. The IS curve shifts to the right. Decreasing government spending has the opposite effect.

There are three ways the government can finance expenditures. First, it can impose taxes. When the government increases taxes, people have less to spend, decreasing demand, and leading to less investment. The economy shrinks, and interest rates go down. However, if the government spends the entire tax increase, the stimulus of increased expenditure outweighs the contraction caused by taxes, since some of the tax money now being spent would have been saved. Second, the government can use debt financing and borrow money. The government borrows money by selling bonds. An increase in the supply of bonds drives down the price and drives up the interest rate. Third, the government can use its right to seigniorage to simply print and spend money, which increases the money supply. As we noted previously, the increased money supply will further stimulate the economy but will have a countervailing impact on interest rates.

Seigniorage-financed fiscal policy seems the logical choice for stimulating the economy, but it carries the threat of inflation. Governments could dramatically increase their ability to use seigniorage if they increased reserve ratios to 100%, as suggested by Frederick Soddy so long ago. The government would then be able to print and spend money when the price index started to fall, and tax and destroy money if inflation threatened to become a problem. The government would also be able to target monetary policy much more effectively, using it to address issues of scale, distribution, and allocation.

The impact of fiscal and monetary policy depends on how much excess capacity exists in the economy. Consider a bowling alley in a small, isolated town where the government is undertaking a large project to stimulate economic growth. When unemployment is high, wages may be fairly low, and few people have disposable income to spend on bowling. As a result, the bowling alley is virtually empty. If the government funds a large project in town, some people are directly employed by the project, and they spend much of their money in town, inducing other local businesses to hire to meet the increased demand. People use their extra income to go bowling, and the bowling alley's income grows.

Now imagine that the government implements the same project in a town with very low unemployment. Bowling is popular, and the alley is full every night. The government needs employees for the project, but increasing demand when supply is low drives up wages, the price of employees. Disposable income increases, but every new bowler at the alley simply "crowds out" another bowler, who would have to leave the alley. The alley might like to expand, but the government is borrowing money

to finance its project, driving up interest rates, making it too expensive for the alley owner to expand. The alley can raise its prices with the increased demand, but it must also pay higher prices for its labor force, and therefore can only break even. When an economy is at full employment, the bowling alley owners might be much better off with an expansionary monetary policy that lowered interest rates so they could expand. In contrast, if the government lowered interest rates when the alley had considerable excess capacity, expansion would do the owners no good at all.

The failure of lowering interest rates to stimulate economies with low demand is known as a **liquidity trap**. In general, the economy is somewhere between the extremes of depression and operation at full capacity (i.e., most bowling alleys are full sometimes, but very few are always full). While increased government expenditure leads to some degree of crowding out and increased interest rates, it also increases income.

Table 16.1 summarizes the impacts of fiscal and monetary policy on interest and income. In each case, the impact is obviously the opposite for the opposite policy.

Inflation and Disinflation

If we looked only at the IS-LM model, and if our goal was continued economic growth, the superior policy option would be clear: keep increasing the money supply to lower interest rates and stimulate investment, and use fiscal policy when necessary to stimulate demand. However, when we first presented the LM model, we saw that the real money supply is equal to the nominal money supply divided by prices—that is, the real money supply equals M/P. There is, therefore, another path toward equilibrium between supply and demand for money in response to an increase in money supply—price inflation. A larger nominal money supply divided by higher prices can lead to no change in real money supply. The closer the system is to full output, the less output is likely to increase in response to lower interest rates, and the more likely that monetary expansion will result in inflation. **Inflation** is an increasing general level of prices (not a state of high prices).

Why are governments and monetary authorities so worried about inflation? Are their concerns justified? How does inflation affect the real economy? The first point to make is that people appear not to like inflation, which alone is some justification for trying to avoid it. Many economists argue that inflation is regressive, but empirical support for this argument is difficult to find.[5] Empirical evidence does show, however,

[5]A. Bulir and A.-M. Gulde, Inflation and Income Distribution: Further Evidence on Empirical Links, IMF Working Papers, no. 95/86. Washington, DC: International Monetary Fund, 1995.

Table 16.1

Expected Impacts of Basic Monetary and Fiscal Policies on Interest Rate and Income

Policy	Interest Rate	Income
Monetary expansion can be accomplished by: • Reduced reserve requirements • Selling bonds on the open market • Lowering the discount (interest) rate	(–) When economy is weak (high unemployment, low investment), monetary policy may have little to no impact on interest rates.	(+) When economy is weak, no impact on income. Known as the liquidity trap.
Tax increase	(–) Taxes (especially progressive income taxes) help stabilize the economy.	(–) Taxes collect more money when income grows and less when it shrinks.
Increased government expenditure	(+) Can be spent on market or nonmarket goods.	(+)
• Financed by deficit spending	(+) Impact on interest rate may be small when economy is weak, large when economy is operating at full capacity.	(Probably +) Income will increase when economy is weak, but may not increase when it is already operating at full capacity; latter condition is known as crowding out.
• Financed by taxes	(+) Increase in interest rate is less than occurs with deficit spending.	(+) Growth rate is less than occurs with deficit spending.
• Financed by seignorage	o	+ Likely to cause inflation under crowding out conditions, with no real growth in income.

that real wages can fall substantially during prolonged episodes of high inflation.[6] In addition, during episodes of high inflation, it is likely that the wealthy and educated are better able to take advantage of investments and contracts that protect their money than the poor. Thus, with continuous high inflation, the poor may well lose ground to the rich.

Hyperinflation, often defined as inflation greater than 50% per month, can also destabilize the economy. In hyperinflation, money fails not only as a store of value, but also as a medium of exchange. Impacts of moderate inflation depend to a large extent on whether it is expected or unexpected.

[6]B. Braumann, High Inflation and Real Wages, Western Hemisphere Department Series: Working Paper WP/01/50, May 1, 2001.

If everyone expects a certain rate of inflation, and their expectation comes to pass, then inflation is incorporated into contracts and causes very few problems. The only groups one would expect to lose from an expected inflation are holders of money (which pays no interest) and people on fixed incomes. However, with expected inflation, most people will hold less money, and incomes are likely to be inflation adjusted. **Disinflation** is a decrease in the rate of inflation. Deflation is a decline in the overall price level. Unexpected inflation, disinflation, and deflation have entirely different outcomes than expected inflation. The most useful way to assess the impacts of these unexpected changes is to look at debtors versus creditors.

Unexpected inflation has entirely different outcomes. The most useful way to assess the impacts of unexpected inflation and disinflation is to look at debtors versus creditors. When there is unexpected inflation, any loans with nominal interest rates (i.e., interest rates that are not pegged to inflation) will be worth less and less every year. Debtors benefit and creditors suffer. For example, people in the 1960s got 30-year house mortgages at around 6%. When inflation in the 1970s climbed over 12%, some home owners ended up paying back less than they originally borrowed. In general, unexpected inflation systematically redistributes wealth from creditors (generally the rich) to debtors (generally the poor). The government is a net debtor, and therefore benefits, as do the future generations that are expected to pay off the government's debts. However, a country cannot have unexpected inflation forever—eventually it becomes expected, or else becomes hyperinflation, with its accompanying problems.

What happens when the government tries to cause disinflation or deflation? Obviously, just as unexpected inflation benefits debtors, unexpected disinflation must benefit creditors. In 1980, a 30-year mortgage at 14% didn't look so bad when inflation was 13% annually, and people expected their incomes to rise by at least that rate. By 1986, however, inflation (and wage increases) had fallen to less than 2%, and creditors were collecting a 12% annual real return on their loans. Thus, existing debtors suffer and existing creditors benefit from disinflation.

Other impacts of disinflation depend on whether it is brought about by fiscal or monetary policy. Theoretically inflation can be reduced by decreasing aggregate demand or increasing aggregate supply, but policy usually acts on demand. Fiscal policy can only decrease aggregate demand through greater taxation or reduced expenditure, both of which should lower the real interest rate to the benefit of new debtors. Other distributional impacts depend on the specific policy used. For example, demand could be reduced by reducing subsidies for big business or by reducing transfer payments to the poor.

THINK ABOUT IT!

Under President Reagan there was a big emphasis on supply-side economics, increasing income by providing incentives for production (i.e., supply). Policy measures for achieving this include investment subsidies, reduced capital gains taxes, and reduced taxes for the rich. Can you explain why these policies would theoretically increase supply and reduce inflation?

The monetary authority, on the other hand, can only act to reduce demand by reducing the money supply, which increases real interest rates, to the detriment of debtors. Interest-sensitive sectors of the economy, such as farming and construction, also lose out. If losers are forced into liquidation or bankruptcy, they may be forced to sell their assets at bargain prices, and it is the well-to-do who maintain the liquidity necessary to purchase those assets. Thus, recessions may generate corporate mergers and increased concentration of the means of production.

The claim made for disinflationary policies is that in the short term the economy suffers, but in the long term stable money allows for steady growth and higher real wages. The problem is that short-term suffering can be severe, especially when monetary policy is used to decrease demand. While the jury is still out on the distributional impacts of moderate inflation, the distributional impacts of unemployment caused by disinflationary policies, as we will see below, are clear.

Unemployment

In the world of microeconomics, involuntary unemployment should not exist. Prices are set by supply and demand, and when the demand for labor is low, the price falls. At a lower wage, fewer people are willing to work, and supply falls accordingly, returning the system to equilibrium. Clearly, however, unemployment is a persistent problem in modern economies. We particularly want to examine two issues: the link between unemployment and inflation, and the implications of unemployment for distribution.

Some unemployment is inevitable. People are constantly entering and leaving the labor market, changing jobs, and moving from place to place. Businesses go bankrupt, or suffer downturns and lay people off. It always takes time to find a new job. This is known as "frictional" or "natural" unemployment. According to theory, if policy makers tried to reduce unemployment below this level, the result would be greater demand for a fixed number of workers. Workers would have more bargaining power and would demand higher wages, thereby causing inflation.[7] Thus, a wide-

[7]This theory was originally introduced by Milton Friedman in his 1967 American Economics Association presidential address.

spread euphemism for "natural" unemployment is NAIRU, the non-accelerating inflation rate of unemployment.[8] There is considerable disagreement over what NAIRU actually is. James K. Galbraith argues that economists are really quite practical—their estimates of NAIRU simply reflect actual unemployment.[9]

But the link between low unemployment and inflation is not clear empirically. Why not? We offer two explanations. First, in the era of globalization, large corporations are free to move their capital and production to other countries. Even when unemployment is low, corporations can counter demands for higher wages by a local workforce with the threat of moving to a lower-wage country. This explains how the low unemployment of the 1990s in the U.S. was accompanied by stagnant wages and a diminished share of national income going to wage earners.[10] Second, we must point out that income from production is divided between wages, profit, and rent. Increased bargaining power by wage earners need not lead to "wage-push" inflation—it could instead simply increase the share of income going to wage earners and decrease the share going to rent or profit. Does increased bargaining power by owners lead to "profit-push" inflation?

In summary, then, low unemployment increases the bargaining power of wage earners, which translates into higher wages (though this effect is diminished by globalization). Higher wages can cause inflation, which then erodes the higher wages, or it can change distribution patterns between wages and profit. High unemployment, in contrast, increases the bargaining power of corporations and leads to redistribution toward the owners of capital. Whatever the validity of the theory behind NAIRU, it is quite clear that monetary authorities pay close attention to unemployment as an indicator of inflationary pressures. For example, when unemployment falls too low, the Fed tends to raise interest rates to reduce investment, employment, and demand. Distributional impacts of inflation are uncertain, but unemployment caused by disinflationary policies has clearly negative impacts on some of the poorest sectors in society.

Finally, it is worth noting that increasing unemployment can set up a vicious cycle. As people lose jobs, they lose money to purchase goods and services. With less demand, businesses respond by reducing supply, perhaps laying off more workers to do so, and further reducing demand.

---

[8]In the 1960s, economists found an inverse empirical relationship between unemployment and inflation, which was dubbed the Phillips curve. But during the 1970s, a number of economies experienced increasing unemployment and increasing inflation simultaneously.

[9]J. K. Galbraith, Well, Excuuuuse Me! *The International Economy*, December 1995.

[10]R. J. Gordon, The Time-Varying NAIRU and Its Implications for Economic Policy, NBER Working Paper No. W5735, May 1997.

Many fiscal policies such as welfare payments, unemployment insurance, and other transfer payments are designed to diminish this impact, adding stability to the economy. Economic stability is a public good, and an important policy objective.

The Impact of Policies on Scale, Distribution, and Allocation

Now that we understand the basic elements of fiscal and monetary policy, we can turn to their particular applications. How we apply these policies, of course, depends on what we wish to achieve. Mainstream macroeconomists primarily pursue continuous economic growth, with a lesser emphasis on distribution. Allocation is left to microeconomic forces. Ecological economists are primarily concerned with the impact of macroeconomic policies on scale (i.e., growth) but with a different goal than mainstream economists—to make sure that the costs of additional growth in material throughput are not greater than the benefits. Ecological economists assume that eventually the costs will exceed the benefits if they haven't already. They also place much more importance on distribution than mainstream economists. In short, ecological economics strives to create an economy in which there is no growth in physical throughput, while avoiding the suffering caused by recession or depression. The allocation of resources between market and nonmarket goods and services can play an extremely important role in this regard.

Macro-allocation

As we have discussed earlier, free markets work very well at allocating resources among market goods but very poorly at allocating nonmarket goods, typically failing to provide them in satisfactory quantities. Many policy makers already recognize this point, as can be clearly seen in government budgets, the bulk of which are spent on public goods such as defense, health care, education, road systems, bridges, streetlights, national parks, and so on.[11] In fact, few institutions besides government allocate resources toward nonmarket goods, and only the government is able to use policy to reduce demand and hence expenditure for market goods and shift it toward nonmarket goods.

Macro-allocation is the allocation of resources among market and nonmarket goods and services.

For simplicity, we refer to the allocation between market and nonmarket goods as **macro-allocation**, and allocation among market goods as **micro-allocation**.

[11]National defense is generally considered a public good, though arms races, nuclear weapons, and excessive defense expenditures may do more to undermine national security than to ensure it. To the extent that disease is communicable and individuals are made uncomfortable by the suffering of others, health care is also a public good.

In the private sector, monetary policy directly affects only the *market* economy, by stimulating or discouraging investment in the production and consumption of *market goods* for profit. Why is this so? Monetary policy acts primarily through its impact on interest rates and hence on borrowing and lending. The private sector invests little in nonmarket goods, since such goods generate no profit that can be used to pay back loans. Therefore, lower interest rates will not affect the production of nonmarket goods by the private sector. Not only will monetary expansion do nothing to provide public goods and open access resources, it can actually increase the degradation of these resources if the production of the market goods is accompanied by negative externalities affecting the environment. Return to our example of the bowling alley. If lower interest rates induce it to expand, it will not expand into a void and may expand into some ecosystem—a wetland, for example—that currently provides valuable nonmarket services to the local community. As we discussed earlier, such negative externalities are an inevitable outcome of market production.

Therefore, if our policy objective is sustainable scale, monetary expansion is very problematic. Even if the economic scale is well within the constraints imposed by the ecosystem, monetary expansion acts on only one type of good, market goods. Market goods do not always offer the highest marginal contribution to human well-being, however. The microeconomic law of the equimarginal principle of maximization thus applies not only to the scale of the economic system relative to the ecosystem that sustains it, but also to the division of market and nonmarket goods produced by an economy. In ecological economics, macro-allocation is every bit as important as micro-allocation.

Theoretically, federal money in a democratic society will be directed toward the goods and services that provide the greatest marginal utility for society as a whole. As we have discussed, an important role of government expenditure is to provide nonmarket goods.

It is important to distinguish between two classes of nonmarket goods, which have different effects on scale. Manmade nonmarket goods affect scale to the same degree as market goods. If the government project in the bowling alley town is a big government building, it may also encroach upon some valuable ecosystem and destroy the services it provides. In contrast, what if the government project in the bowling alley town created a park that provides recreation and restores wetlands that serve as filters for the town's water supply and as a buffer against catastrophic floods? Protecting and restoring the ecosystems that provide nonmarket environmental services can effectively decrease scale, or at least help ensure that we do not surpass optimal scale. As the world becomes more full, the marginal benefits from protecting and restoring ecosystem funds, and hence the nonmarket services they provide, will increase relative to those

from market goods and manmade public goods. As this happens, and if politicians come to understand the benefits and public good nature of ecosystem services better, more and more federal money should be allocated toward providing such services.

It is important to recognize, however, that government expenditure on ecosystem funds can still increase scale. How is this so? We must remember that once the initial expenditure enters the economy, the multiplier effect takes hold. Money spent to restore ecosystem funds will in turn be spent by its recipients on market goods—workers restoring the wetland may spend their money on bowling, pressuring the bowling alley to expand. The larger the multiplier, the larger the impact on the market sector of the economy, and the less control the government has over composition. As most macroeconomics textbooks explain, tax increases decrease the multiplier and also reduce income. A smaller multiplier increases the ability of the government to affect macro-allocation, and reduced income reduces scale. Taxes can also be used to discourage undesired behaviors, such as pollution, and transfers can be used to encourage desired ones, such as environmental preservation. The full impact of taxation on scale and macro-allocation depends on how the taxes are spent, but taxes can certainly play an extremely important role in achieving an optimal scale—a point we will examine at greater length in Chapter 21.

Another important point must be made here. Under traditional analysis of the IS-LM curve, fiscal policy when the economy is operating at full capacity results in crowding out (remember the full bowling alley) and should be avoided. However, in terms of macro-allocation and scale, full output conditions can increase the effectiveness of fiscal policy. With full employment, if the government spends money to create a park and restore wetlands, interest rates and labor costs go up, and it is more difficult for the bowling alley to expand. (Fortunately, the park offers a recreational alternative to bowling that does not displace ecosystem services.) Government expenditure on restoring ecosystems under such conditions will therefore have an unambiguous impact on reducing scale.

What are the distributional impacts of fiscal and monetary policy? Fiscal policy in the form of taxation and government transfers can be easily and effectively distributed as desired. Government transfers such as welfare, unemployment insurance, Medicare, Medicaid, and Social Security all play an important role in distribution. Corporate welfare programs (which outweigh transfer payments to the poor[12]) affect distribution in the opposite direction. Public goods are available to all, and their provi-

[12]C. M. Sennott, "The $150 Billion 'Welfare' Recipients: U.S. Corporations," *Boston Globe*, July 7, 1996.

sion improves distribution. In terms of income, progressive taxation can help reduce gross inequalities in income distribution, a necessary condition morally and practically if we are to achieve a sustainable scale. Monetary policy can also play an important but narrower role in distribution. High interest rates caused by tight monetary policy can lead to unemployment, and they favor creditors over debtors, as discussed earlier in the section on inflation and disinflation. Low interest rates have the opposite effect.

In summary, in terms of ecological economic goals, monetary policy is a blunt instrument directed only toward the production and consumption of market goods, with limited flexibility in terms of distribution and macro-allocation. Expansive monetary policy increases scale. Fiscal policy has far greater flexibility in terms of scale, distribution, and macro-allocation.

IS-LM in the Real World

While the IS-LM model is very useful, it has important limitations.[13] The model is deceivingly simple and does an inadequate job of conveying the real-life complexity of monetary and fiscal policies. While the model shows the general impacts of such policies, it fails to incorporate the issues of uncertainty, time lags, and structural changes, as well as the difficulty of choosing the appropriate policy variables to manipulate.

Economists typically have a poor understanding of what is happening in the economy at any given moment. Is unemployment too high? Is the economy growing too fast, threatening inflation? Are we headed for recession? Weathermen may be inaccurate at predicting the future, but at least they can look outside right now and tell you what the weather is. Economics is less advanced. Viewing the same data concerning the economy, economists frequently disagree on how to interpret them. For example, in the United States, the Federal Reserve Bank raised interest rates in May 2000, fearing that the economy was growing too fast, and continued to voice fears of inflation through November 2000. Less than 2 months later, the Fed initiated the first of several interest rate cuts designed to ward off recession. Part of the problem is that the economic system is evolving rapidly in response to technological, environmental, cultural, and structural changes.

[13]See J. R. Hicks, IS-LM: An Explanation, *Journal of Post-Keynesian Economics* III (2): 139–154 (Winter 1980–1981). Hicks, the originator of the model, expresses reservations about how well it fits the real world once expectations and dynamics are recognized. He considers the model useful for understanding the past, but less so for understanding the future. We agree, but feel that many of Hicks' caveats apply to all equilibrium models, and that, rare though it is among scholars, Hicks was too hard on himself!

Compounding the difficulty of an inadequate understanding of the economy are the time lags involved in policy. There are two types of lags: lags in decision making (the inside lag) and lags between the time the decision is made and the policy takes effect (the outside lag). In fiscal policy, decisions such as tax cuts and expenditure increases are typically debated at length. Both legislative and executive branches must agree, and appropriate legislation must be passed. The decision lag can therefore be substantial. Once the decision to increase or decrease expenditure has been made and carried out, the outside lag may be relatively short, as such policies have an immediate effect on aggregate demand (though the full effect of the multiplier will take some time). Tax cuts or increases, on the other hand, have much slower results and are often not even felt until the next tax year.

The Fed, in contrast, generally has a much shorter decision-making lag. The Federal Open Market Committee (FOMC), responsible for Fed Policy, meets about eight times a year.[14] Policy is generally decided at the meeting, and open market transactions can take place almost immediately. However, the most relevant impact of these policies is on interest rates and their effect on investment and consumption decisions. Investment decisions are rarely spur of the moment; they generally have a long gestation period. Thus, the Fed has a short decision-making lag, and a long lag before the policy takes effect.

These lags are very important to consider when deciding on a policy. It is quite possible that by the time a decision is made and the resulting policy takes effect, the problem the policy was designed to address will have disappeared, and a policy with the opposite effect may even be required.

Another problem is disagreement over what type of policy should be pursued, and what the impact will be, especially for monetary policy. The Fed usually tries to manipulate one of two targets: the money supply or the interest rate. Not only is there considerable debate over which course the Fed should pursue, there are serious obstacles to achieving either goal. For example, as Fed chairman Alan Greenspan admitted in congressional testimony: ". . . we have a problem trying to define exactly what money is . . . the current definition of money is not sufficient to give us a good means for controlling the Money Supply. . . ."[15]

Psychology can also make it difficult to manipulate interest rates. As we discussed earlier, interest rates are ultimately determined by the bond markets. Bonds, of course, mature in the future, and the amount someone

[14]The Federal Reserve Act mandates that the FOMC meet at least four times a year, and since 1981 it has met eight times a year.

[15]Congressional testimony, February 17, 2000.

is willing to pay for a bond depends on their *expectation* of future inflation. The Fed might implement an expansionary monetary policy to bring down interest rates, but if bond marketers believe this expansion will instead induce inflation or force monetary contraction in the near future, it could paradoxically serve to drive interest rates *up*.

A final problem with policy in countries with independent monetary authorities is the difficulty in coordinating between monetary and fiscal policy. This problem can become acute when the monetary authority and the government have different policy objectives. The elected government is mainly concerned with growth and employment, two issues that affect voters, and hence their elected representatives. In contrast, the Fed is mainly concerned with "sound" money (i.e., no inflation) and has frequently pursued this policy in the past, even when it has caused significant hardship in the form of unemployment and worse.

Box 16-5 Why Is the Fed So Anti-Inflation?

From our discussion of inflation, it would seem that inflation is less harmful than unemployment induced by anti-inflationary policies. Why then is the Fed so anti-inflation? In answering this question, it is worth bearing in mind who the natural constituency of the Fed is. Most members of the FOMC are bankers or Wall Street professionals, and the Fed seems to listen closely to the concerns of these groups. These two groups form the bulk of the wealthy creditors who benefit from low inflation and disinflationary episodes, and who are unable to increase their share of national income as readily during inflationary periods.[a]

[a] *W. Greider,* Secrets of the Temple: How the Federal Reserve Runs the Country, *New York: Simon & Schuster, 1987.*

Despite its shortcomings, the IS-LM model is a vast improvement over prior models. It is a two-sector general equilibrium model, the sectors being the real sector and the monetary sector. Before Hicks' model, economists often tried to explain the interest rate as a purely monetary phenomenon (liquidity preference and money supply) or a purely real phenomenon (savings and investment). There was an investment rate of interest and a money rate of interest, and confusion about which set of factors "really" determined the interest rate. Hicks showed that the real and monetary sectors simultaneously interact to determine both the interest rate and national income. But Hicks said nothing about the ecosystem and biological rates of growth. In 1937 the world was still considered

"empty." Thus, the IS-LM model treats all economic growth as identical—it does not distinguish between government expenditures on market goods, manmade public goods, or investments in ecological restoration, nor does it address distribution.

Adapting IS-LM to Ecological Economics

How might the IS-LM model be adapted to ecological economics? Remembering our basic vision of the macroeconomy as a subsystem of the finite and nongrowing ecosystem, the most obvious suggestion would be to impose an external constraint on the model representing the biophysical limits of the ecosystem. For example, we could assume a fixed throughput intensity per dollar of Y (i.e., GNP), so that a given Y in money terms implied a given physical throughput. Then we could estimate the maximum ecologically sustainable throughput, convert that into the equivalent Y, and impose that as an exogenous constraint on the model. Based on Figure 16.4, it would be represented by a vertical line at the Y corresponding to maximum sustainable throughput. It would not be a function of the interest rate at all.[16] Let's call the vertical line EC for "ecological capacity." It reflects a *biophysical* equilibrium, not an *economic* equilibrium. It is ignored by the actors whose behavior is captured in the IS and LM curves.[17]

The most obvious approach is not always the best, but it is usually a good place to start. Also, this approach closely parallels the macroeconomist's representation of full employment of labor as a perpendicular at the level of Y corresponding to full employment at an assumed labor intensity of GNP. Our EC line represents "full employment" of the environment at an assumed throughput intensity of GNP. Later we will discuss further the assumption of fixed throughput intensity.

Let's consider the three possible positions of the biophysical equilibrium relative to the economic equilibrium, shown in Figure 16.6. The first case represents the "empty world" scenario. The biophysical limit is not binding. The distance Y*C may be thought of as excess carrying capacity.

[16]A. Heyes, A Proposal for the Greening of Textbook Macro: "IS-LM-EE," *Ecological Economics* 32 (1) (2000); P. Lawn, *Toward Sustainable Development*: An Ecological Economics Approach, Boca Raton, FL: Lewis Publishers, 2001. Heyes and Lawn have proposed an EE curve corresponding to the ecological limits discussed here that would be a function of the interest rate. Several technologies produce income, some of which require or degrade more natural capital than others. Less natural capital-intensive technologies require investments and are thus more likely at lower rates of interests. One problem is that the investments themselves would require natural capital.

[17]However, it need not be ignored by government, which affects the IS curve. The government is perfectly capable of investing in environmental services produced by natural capital and other nonmarket goods. But it is completely ignored by monetary policy (the LM curve), which acts on the economy through its effect on interest rates and hence market goods.

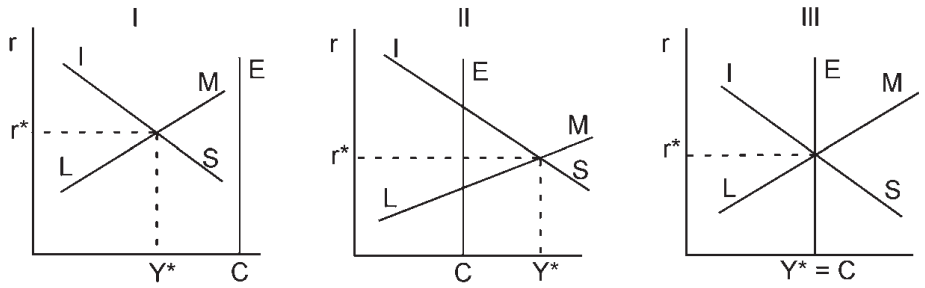

Figure 16.6 • The biophysical equilibrium relative to the economic equilibrium.

Most macroeconomists who use the IS-LM model would have this case in mind, if indeed they thought at all about EC. If the distance Y*C is large then there is for practical purposes of short run policy no point in conceiving or drawing the EC line.

The second case is the "full world" (or overfull) scenario. The economic equilibrium has overshot the biophysical equilibrium. The distance CY*, the overshoot, can be given two interpretations. In neither case, however, would it represent real income. In the first case we might think of it as a purely monetary phenomenon, inflation. After hitting EC the real sector is effectively at the end of its tether. Even though the monetary sector keeps on churning, it only generates price increases for the same real income C. The other interpretation assumes that the real sector keeps producing real output, but by unsustainable drawdown of natural capital. Thus, CY* would represent capital consumption counted as income. As natural capital is consumed, the EC line eventually has to shift even farther to the left, increasing the overshoot. Most ecological economists believe the second case to be a rather accurate description of the present state of affairs. We seem to be avoiding short-term inflation by long-term capital drawdown. Most conventional economists do not worry about long-term capital drawdown and shifting the EC curve farther to the left because they believe that knowledge is shifting EC to the right and thereby restoring the empty world situation.

The third case represents a big coincidence under our assumptions. For the economic equilibrium to coincide with the biophysical equilibrium would require either extraordinary good luck, or purposeful coordination and planning. There is nothing in the model to make it happen, just as, currently, there seems to be nothing in our institutions or behavior that would make it happen. In Chapters 20–23, we will discuss policy changes that could theoretically lead to this outcome.

Recall that we previously discussed the concept of full employment, which we might represent by an FE limit for labor similar to the EC limit for natural capital. Ideally, a FE labor line should coincide with the ISLM

equilibrium point—make IS = LM at full employment. If FE is beyond the intersection of IS and LM, then policy makers might pursue FE through growth in Y. But what if FE is beyond EC? The problem is no longer to pursue FE by growth in Y, but instead through structural change, such as shifting factor intensity away from fossil fuels and manmade capital (both of which rapidly draw down natural capital) and toward labor. We have already explained that when IS = LM beyond EC, we are likely to draw down natural capital, and it is implicit in the acronym NAIRU (the non-accelerating inflation rate of unemployment) that going beyond FE results in inflation. Why the difference? Why doesn't moving beyond EC also simply cause inflation? The answer is that natural resources are either free or cheap to begin with; they are not appropriately priced by the market mechanism, and excessive use therefore does not affect the price signal.

It remains true, however, that the assumption of constant throughput intensity of Y is troublesome. We know that throughput intensity of Y changes with new technology, and with shifts in the mix of goods that make up Y, even if probably not with factor substitution of capital funds for throughput flows. Differing assumptions about throughput intensity of Y can at least be represented by a shift in the EC perpendicular. (Curve shifting is, as you have no doubt realized by now, not an uncommon device in economic analysis!) However, in terms of practical policy recommendations, perhaps the best approach would be simply to impose the ecological constraint as a limit on throughput. For any given technology, a fixed limit on throughput will also limit Y, but over time, new technologies and a different mix of goods and services can allow Y to increase without increasing throughput, thereby threatening the life-support functions of the ecosystem.

BIG IDEAS to remember

- Macroeconomic model vs. general equilibrium
- Real vs. monetary sectors
- IS = LM
- MV = PQ
- Transactions demand for money
- Liquidity preference
- Relation of bond prices to interest rate
- IS-LM analysis of monetary and fiscal policy
- Comparative statics
- Crowding out
- Inflation and disinflation
- Unemployment
- Macro-allocation
- IS-LM adapted to show economic and biophysical equilibria

Conclusions to Part IV

Chapters 13–16 offered a whirlwind tour of the macroeconomy. The real economy consists of the physical transformation of low-entropy matter-energy from nature into forms that enhance human welfare. GNP is an inadequate measure of the real economy, because it lumps together both goods and bads. The monetary sector of the economy functions as a lubricant that helps us allocate what is produced. The law of diminishing marginal utility tells us that distribution, both within and between generations, can be an important tool for increasing human welfare. Macroeconomic policy levers include government expenditure, taxes, money supply, and interest rates. These policy levers must be used to attain the goals of sustainable scale and just distribution—goals essentially ignored by the market microeconomy. We next turn our attention to international economics, to see how it affects the policy levers we just discussed.