ECONOMICS of the PUBLIC SECTOR

THIRD EDITION

JOSEPH E. STIGLITZ

W.W. NORTON & COMPANY
NEW YORK / LONDON

20 Optimal Taxation

FOCUS QUESTIONS

1 What are the trade-offs involved in designing a progressive income tax system?

2 What should be the role of the taxation of commodities (such as luxuries) and savings in achieving greater equity in taxation?

3 If the government imposes taxes on different commodities, how should the tax rates be set so as to minimize the total deadweight loss?

In the previous chapter we observed that there may be a significant welfare loss (the *deadweight loss*) associated with any tax other than a lump-sum tax. Two questions immediately arise: Why, if this is the case, do we not just impose a lump-sum tax? And if we are to impose distortionary taxes, is there some way that they can be designed to minimize the deadweight loss? These questions have been at the center of theoretical research in taxation. The research has produced some remarkably simple and insightful answers, answers that may help to design better tax systems in the future.

The chapter is divided into four sections. The first section disposes of two fallacies that have long confused discussions of tax design. Next the basic principles of optimal taxation are described, and then applied to analyze the design of income tax structures. The final two sections analyze commodity taxation. The third section focuses on the effectiveness of taxing consumers' purchases of different commodities at different rates in achieving redistributive goals, and the fourth on the role of taxation of producers.

TWO FALLACIES OF OPTIMAL TAXATION

Before turning to the details of the analysis, we need to dispose of two fallacies which have misled discussions of tax design—one suggesting an overly simplified approach, the other that the world is so complex that nothing can be said.

THE FALLACY OF COUNTING DISTORTIONS

The first fallacy says we should simply have a tax on wage income. Additional taxes—taxes on commodities, such as cigarettes or alcohol, or taxes on savings—just add to the number of distortions and thus to economic inefficiency. One distortion is better than several distortions.

A tax on wage income would be optimal if there were no distortions associated with that tax; for then that tax would be equivalent to a lump-sum tax. But we showed in the previous chapter that an income tax distorts individuals' decisions to work, and it is not necessarily the case that one large distortion is better than several smaller distortions. Chapter 19 showed that the deadweight loss from a tax was proportional to the square of the tax rate. This suggests that it may be better to have a number of small taxes than a single large tax.

MISINTERPRETATIONS OF THE THEORY OF THE SECOND BEST

In earlier chapters we characterized Pareto efficient resource allocations. All of the required conditions are seldom satisfied. The **theory of the second best** is concerned with the design of government policy in situations where the economy is characterized by some important distortions that cannot be removed.[1] This is in contrast to "first-best" economies, where all the conditions for Pareto efficiency can be satisfied. Second-best considerations say that it may not be desirable to remove distortions in those sectors where they can be removed. The theory of the second best is often interpreted fallaciously as saying that as long as there are some distortions, economic theory has nothing to say. This is incorrect, as we shall shortly show. Economic theory can tell us under what circumstances two small distortions are preferable to one large one; when it is better to have inefficiencies in both consumption and production; and when it is better not to have inefficiencies in production. Second-best theory tells us that we cannot blindly apply the lessons of first-best economics. Finding out what we should do when some distortions exist is often a difficult task, but it is not impossible.

[1] Early formulations of the theory of the second best include those of James Meade, *Trade and Welfare: Mathematical Supplement* (Oxford: Oxford University Press, 1955), and R. G. Lipsey and K. Lancaster, "The General Theory of Second Best," *Review of Economic Studies* 24 (1956–1957): 11–32.

OPTIMAL AND PARETO EFFICIENT TAXATION

Chapter 3 introduced the concept of Pareto efficiency. Recall that a resource allocation was Pareto efficient if no one could be made better off without someone else's being made worse off. Similarly, in judging tax structures we again use the concept of Pareto efficiency: a **Pareto efficient tax structure** is one such that there exists no alternative tax structure which can make some individuals better off without making other individuals worse off.[2] If such an alternative tax system exists, then the current tax system is clearly inefficient.

There are many Pareto efficient tax structures, just as there are many Pareto efficient resource allocations without taxes. In each, no one can be made better off without someone else's being made worse off. They differ in distribution. In the two-person economy of Robinson Crusoe and Friday, Crusoe is better off in some Pareto efficient allocations, Friday is better off in others.

In Chapter 5, we learned how one can choose among Pareto efficient resource allocations using a social welfare function. So too in choosing among Pareto efficient tax structures: the **optimal tax system** is the set of taxes which maximizes social welfare. Clearly, different social welfare functions will generate different optimal tax structures. At a practical level, for instance, a social welfare function which reflects a greater concern for equality (such as a Rawlsian social welfare function) may imply that the optimal tax structure is more progressive, with the rich bearing a larger fraction of the burden for paying for public goods. One of the objectives of optimal tax theory is to determine whether there are some general properties of all Pareto efficient tax systems, that is, properties which hold regardless of the social welfare function.

LUMP-SUM TAXES

If all individuals were identical and were treated for tax purposes identically, a lump-sum tax would be the only efficient tax: any other tax would introduce distortions, so that the government could raise the same amount of revenue *and* make each individual better off. And if everyone were identical, there would be no reason to redistribute income. Both equity and efficiency would thus require that any revenue that the government needed be raised by imposing a uniform lump-sum tax on all individuals.

In the real world, things are more complicated. Individuals differ, governments wish to redistribute income, and in any case, there is a strong belief that individuals who can more easily pay taxes should pay more taxes than those who cannot easily pay. Even if the government wishes to make different people pay different taxes, it does not follow that it would have to impose distortionary taxes, like income or excise taxes.

[2] For a more detailed description of Pareto efficient tax structures, see J. E. Stiglitz, "Self-Selection and Pareto-Efficient Taxation," *Journal of Public Economics* 17 (1982): 213–40, and J. E. Stiglitz, "Pareto Efficient and Optimal Taxation and the New New Welfare Economics," in *Handbook of Public Economics*, ed. Alan J. Auerbach and Martin Feldstein (Amsterdam and New York: North Holland; distributed in Canada and U.S. by Elsevier Science Publishers, 1987), pp. 991–1042.

A Pareto efficient tax structure is one such that there exists no other tax structure which can make some individuals better off without making others worse off.

The optimal tax structure, given a particular social welfare function, is the Pareto efficient tax structure which maximizes that social welfare function.

WHY IMPOSE DISTORTIONARY TAXES?

Indeed, it can be argued that if the government had perfect information about the characteristics of each individual in our society, it would not impose distortionary taxes. If the government could ascertain who had greater abilities, and who therefore was in a better position to pay taxes, it would simply impose higher lump-sum taxes on those individuals.

But how can abilities be measured? Consider a family. Parents often believe that they have good information concerning the abilities of their children. A parent who has two children, one of whom has a great deal of ability but chooses to become a beachcomber, and the other of whom has limited ability that he uses to the fullest, is more likely to provide financial assistance to the latter than to the former; the assistance is not made on the basis of income—the beachcomber may in fact have a lower income than a hardworking but low-ability brother.

The government, however, is not in the position of the parent who can observe the ability and drive of his children. The government can base its tax only on *observable variables*, such as income and expenditure (and even these, as we shall see, are not easily observable). The choice facing the government is to have either a *uniform lump-sum tax*, one that individuals pay regardless of what they do or what their abilities are; or a tax that depends on easily measured variables, such as expenditures or income—and such a tax is inevitably distortionary. An income tax does not always succeed in taxing those whom we might think ought to be taxed—it treats equally the individual who has low ability but works extremely hard and the individual who is of high ability and takes it easy, provided the two have the same income. Still, most people believe that those who have a higher income ought to pay a higher share of government costs because those with a higher income are, *on average*, more able or have had better than average luck. Moreover, society may reasonably value the loss of income by the rich (implying, say, one less yacht) less than it values the loss of income to lower-income individuals.

The use of distortionary taxes is thus an inevitable consequence of our desire to redistribute income, in a world in which the government can observe the characteristics of individuals only imperfectly. Still, some tax systems are less distortionary than others.

DESIGNING AN INCOME TAX SYSTEM

Pareto efficient tax structures minimize distortions. For instance, one might ask, is it better to redistribute income just through a progressive income tax, or to supplement a progressive income tax with a tax on luxuries consumed

by the rich? Before addressing that question, however, we ask a simpler one: Assuming there are no savings, so the only source of income is wages, and the only tax is an income tax, how progressive should the tax system be? That is, how much larger a portion of their income should rich people pay?

As always, economists focus on trade-offs. Here, the more progressive the tax, the larger the deadweight loss, the inefficiencies from the tax, but the less the degree of inequality. We can view much of the political debate concerning how progressive the tax structure should be as one involving differences in values, in how much deadweight loss one is willing to accept for a given decrease in inequality.

There may be disagreements not only about values but also about the empirical question of what the trade-offs are. Those who advocate more progressive taxes tend also to argue that the cost, in terms of the deadweight loss, of reducing inequality is relatively small. In Chapter 19 we showed that the magnitude of the deadweight loss from a tax was related to the substitution effect. If leisure and consumption goods are very substitutable, then the compensated labor supply schedule will be very elastic, and there will be a large deadweight loss from a tax on consumption or labor income. If consumption this period and consumption next period are very substitutable, then the savings schedule will be very elastic, and the deadweight loss associated with an interest income tax will be large. Those who believe that the deadweight losses are small are often referred to as *elasticity optimists*; they believe, for instance, that the (compensated) labor supply and savings elasticities are low, so that the distortions associated with high tax rates are low; while those who believe that the distortions are large are often referred to as *elasticity pessimists*, because they believe that the labor supply and savings elasticities are large.

WHY DOES MORE PROGRESSIVITY IMPLY MORE DEADWEIGHT LOSS?

The preceding section argued that as we use our tax system to attain greater equality, the deadweight loss increases. Panels A and B of Figure 20.1 illustrate this general proposition by contrasting two tax schedules. The first (in color) is a proportional income tax, in which the tax liability is the same percentage of income for all individuals, no matter how large or small their income. The second is a simple progressive income tax that imposes a tax at a flat rate on the difference between the individual's income and some critical level of income, $\hat{Y}$. Individuals whose income falls below the critical level receive a grant from the government equal to the tax rate times the shortfall between their income and the critical level. Notice from panel B that the marginal tax rate, the extra tax an individual pays or receives on an extra dollar of income, is constant for both tax systems. Therefore, both are called **flat-rate taxes.** But with the progressive tax, the *average* tax rate, the ratio of the total tax payments to the individual's income, increases with income. This is why we call the tax progressive.[3]

[3] Usage is not standardized. Some prefer to reserve the term *progressive* for tax structures where the *marginal* tax rate increases. Nothing important hinges on these semantic points. Notice that a flat-rate tax combined with a lump-sum tax is regressive, in the sense that the average tax rate decreases with income. For a more general discussion of the definition of progressive and regressive tax structures, see A. B. Atkinson and J. E. Stiglitz, *Lectures on Public Economics* (New York: McGraw-Hill, 1980), Chapter 2.

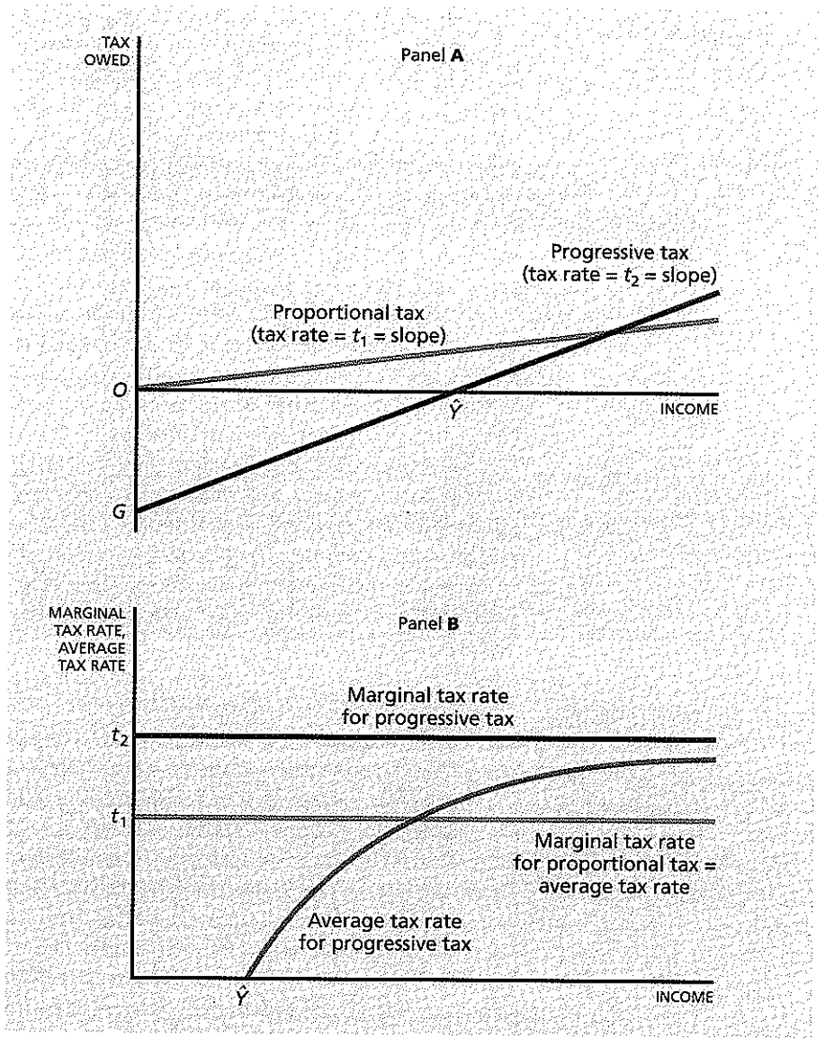

FIGURE 20.1 **Flat-Rate Income Tax Schedules** Panel A compares the tax schedule of a proportional flat-rate income tax with that of a progressive flat-rate income tax. Panel B compares average and marginal tax rates for these two taxes.

Because, as we have depicted it, the progressive flat tax provides for a payment to individuals whose income falls short of the critical level, we sometimes refer to that portion of the tax schedule below $\hat{Y}$ as a **negative income tax.**[4]

[4] In some tax systems, those with income above $\hat{Y}$ are taxed on the difference between their income and this exemption level, but those below the critical level neither pay taxes nor receive a rebate.

The progressive flat tax can be thought of as a combination of a uniform lump-sum grant to all individuals and a proportional income tax. Thus, in Figure 20.1A, a proportional tax at the rate t_2, combined with a grant of OG, is identical to an income tax on incomes in excess of $\hat{Y}$ ($\hat{Y}$ is the exemption level) at a rate of t_2, provided that those with incomes less than $\hat{Y}$ receive a rebate equal to t_2 times the difference between $\hat{Y}$ and their income. If the government is both to finance its public goods and other public expenditures *and* pay everyone a uniform lump-sum grant, the revenue raised must be higher than if it just financed the public goods, so the marginal tax rate must be higher than with just a proportional tax.

In the last chapter, we learned that the deadweight loss increases with the *marginal* tax rate: the magnitude of the deadweight loss is related to the substitution effect; and the magnitude of the substitution effect is related to the marginal tax rate. More progressive taxes have higher marginal tax rates, and thus greater deadweight loss.

Moreover, the more progressive the tax, the greater the *likelihood* of a smaller labor supply and national output necessitating on that account a still higher tax rate. All lower-income individuals are better off, so both substitution and income effects lead to smaller labor supply. For higher-income individuals, income and substitution effects are offsetting. Unless they have very backward-bending labor supply schedules, overall labor supply will be reduced.

A DIAGRAMMATIC ANALYSIS OF THE DEADWEIGHT LOSS OF PROGRESSIVE TAXATION

The fact that more progressive tax results in greater deadweight loss can be seen by looking at any individual, and comparing the revenues the government can obtain with two taxes which leave the individual just as well off. The more progressive tax has a higher marginal tax rate.

Figure 20.2 shows a budget constraint with a proportional tax and a budget constraint with a progressive tax, one which gives the individual a fixed income, even if she does not work. The marginal rate with the progressive tax is higher, but is set so that the individual is on the same indifference curve. We compare the total revenue. It is reflected in the distance between the before- and after-tax budget constraints. Since income is measured along the vertical axis (and hours along the horizontal axis), the tax revenue in dollar terms is measured as the vertical distance between the two budget constraints; for the proportional tax, by the distance $E'A'$, and for the progressive tax, by the distance EA. It is apparent that $E'A'$ is much larger than EA: for any given effect on utility, the progressive tax yields lower revenue.[5] It is less efficient than the proportional tax.

[5] The vertical distance between the indifference curve and the before-tax budget constraint is maximized at the point where the slope of the indifference curve is the same as the slope of the before-tax budget constraint. That is why a lump-sum tax, which does not alter the slope, maximizes revenue for any given impact on utility. Since at E' and E the slope of the indifference curve is flatter than the slope of the budget constraint, the vertical distance is larger the further "up" the indifference curve we move.

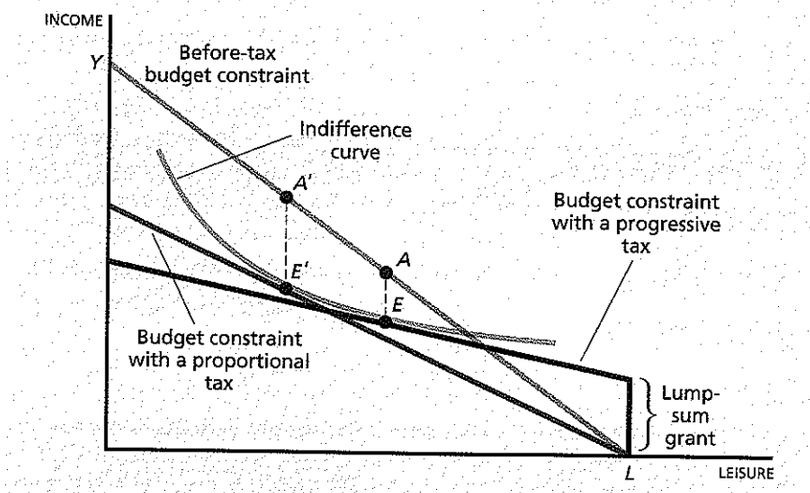

FIGURE 20.2 **Comparing a Progressive and a Proportional Tax Which Have the Same Effect on Utility** Tax revenue is higher with the proportional tax.

CHOOSING AMONG FLAT-RATE TAX SCHEDULES

The analysis has clarified the trade-offs faced as we increase the degree of progressivity. Poorer individuals gain, richer individuals lose. Like a leaky bucket, the dollar gains of the poor are less than the "dollar equivalent" losses of the rich, because of the deadweight losses associated with the tax. But the *social value* of the gains of the poor may well exceed the social value of the losses of the rich. Whether this is so depends, of course, on the social welfare function.

In Chapter 5 we introduced the concept of a Rawlsian social welfare function, where society is concerned about the welfare of the worst-off individuals. The worst-off individuals are those at the bottom of the income distribution, and their welfare is typically related directly to the size of the lump-sum grant. For a Rawlsian, the optimal tax structure is simply that which maximizes the lump-sum grant, that is, which maximizes the revenue that can be extracted from taxpayers. Such a tax rate may be quite high—one estimate[6] put the number at 80 percent, though others have estimated

[6] Nicholas H. Stern, "On the Specification of Models of Optimum Income Taxation," *Journal of Public Economics* 6 (1976): 123–62. He assumed that expenditures on public goods amount to 20 percent of national income. The results of the calculations are very sensitive to all the assumptions made, and in particular to assumptions concerning the compensated and uncompensated elasticities of labor supply. As we noted in Chapter 19, there is considerable controversy concerning their magnitude. Those who believe that the uncompensated elasticity is quite high believe that revenues that can be obtained from taxing the rich, to pay the lump-sum grant to the poor, peak out at much lower rates of taxation. Those who believe that the compensated elasticity is quite high believe that the deadweight loss from the progressive taxation is very high.

THE 1993 TAX INCREASE ON UPPER-INCOME INDIVIDUALS: A PARETO INEFFICIENT TAX?

In 1993, Congress raised taxes on upper-income individuals. The tax increase clearly made them worse off. Critics claimed that it was a Pareto inefficient tax change (though the popular press did not use that vocabulary). The claim was made that these individuals would reduce their work effort—the substitution effect was larger than any income effect—so that tax revenues would be reduced. Thus, funds available to redistribute to the poor would actually be lowered. This did not happen. Instead, tax revenues for the rich increased faster than for others, and far faster than national income. To be sure, upper-income individuals are worse off than they would have been with lower taxes, but the taxes they paid increased. (Of course, it is possible that without the tax rate increase, incomes of the rich would have increased even more, enough so that tax revenues would have increased. But this would have required an implausibly large growth in their income, not commensurate with historical experience.) Thus, whether the tax change was desirable depends on the social welfare function, but it does not appear to have been Pareto inefficient.

lower rates. At 80 percent, the deadweight loss incurred by higher-income individuals is quite high. Other social welfare functions, which put more weight on middle- and upper-income individuals, accordingly suggest a lower optimal tax rate. One estimate[7] put the optimal tax with a utilitarian social welfare function—where all individuals are weighed equally—at 19 percent.

GENERAL EQUILIBRIUM EFFECTS

So far we have assumed that the income tax has no effect on before-tax incomes; that there is, in other words, no *shifting* of the income tax. Some economists, however, believe that there may be considerable shifting. In particular, it has been argued that the income tax system has increased the degree of before-tax inequality.

There are some who believe, first, that the wages and fees of managers and professionals adjust to the taxes, leaving their after-tax income relatively unchanged. Moreover, if as a result of the income tax, skilled workers supply less labor and investment is discouraged, unskilled laborers' productivity and, hence, their wage, will decline. At the present time, unfortunately, we do not know the quantitative significance of these effects. If they are important, it suggests that the benefits of progressivity are less than they seem when these effects are ignored.[8]

[7] Stern, "On the Specification of Models of Optimum Income Taxation."

[8] The importance of these general equilibrium effects in the design of optimal taxes was noted by Martin Feldstein, "On the Optimal Progressivity of the Income Tax," *Journal of Public Economics* 2 (1973): 357–76, using a simulation model. His results were corroborated and extended in subsequent theoretical work by N. Stern, "Optimum Taxation with Errors in Administration," *Journal of Public Economics* 17 (1982): 181–211; F. Allen, "Optimal Linear Income Taxation with General Equilibrium Effects on Wages," *Journal of Public Economics* 17 (1982): 135–43; J. E. Stiglitz, "Self-Selection and Pareto-Efficient Taxation." See also Laurence J. Kotlikoff and Lawrence H. Summers, "Tax Incidence," Chapter 16 in *Handbook of Public Economics*, vol.2, pp. 1043–92.

RAISING BENEFITS FOR THE POOR The analysis also makes clear why it is so difficult to provide increased benefits for the poor. It is not "just" that financing those benefits requires raising taxes. There is a real problem in designing the "phaseout," that is, the rules stipulating how benefits get reduced as income increases. A rapid phaseout implies a high marginal tax rate (since benefits are reduced greatly for each extra dollar earned) over the phaseout income range, thus weakening work incentives. A slow phaseout reduces the magnitude of the disincentive effect, but, if the poorest are to receive the same benefit, raises the benefit levels of others, including lower-middle-income individuals, necessitating further tax increases. The objectives of *targeting* and *good incentives* are inevitably in conflict.

Consider the earned income tax credit (EITC), meant to supplement the wage income of poor families with dependent children. The idea behind the EITC was simple: reward the poor for working, thus encouraging them to work more and acquire more skills. In 1993, the EITC was greatly expanded and indexed, so that in 1997, the maximum benefit, for a family with two or more eligible children, of $3656 phased out over the range of $11,950 to $29,290, implying a marginal tax rate of 22 percent from the EITC. When the Clinton administration took office, it had hoped to expand the EITC so that all those working full-time would be out of poverty. But the overriding desire to reduce the deficit led to a lower maximum benefit than

BASIC TRADE-OFF IN TAX DESIGN

More progressive tax systems entail greater deadweight loss; more "equalitarian" social welfare functions (placing more weight on equality) will choose more progressive tax systems.

FLAT-RATE TAXES ARRIVE ON THE POLITICAL SCENE

The simplicity of the flat-rate tax system has long attracted academic economists. In the early 1980s, Robert Hall and Alvin Rabuschka of Stanford University wrote a widely read book advocating the flat tax. In 1996, Malcolm Forbes ran a presidential primary campaign centered around the flat-rate tax. He proposed a high exemption level, and a low rate. Like the supply-siders of the 1980s, he believed the supply response to the lower tax rate would be a huge increase in national income. But most economists thought that the supply response would be far smaller, leaving a huge deficit—estimated in the hundreds of billions per year. Raising the flat rate to eliminate the deficit made the proposal sound less attractive; but even then it would have represented a huge change in who bears the burden of taxation, with the rich facing markedly lowered burdens and the middle class facing higher burdens. As people examined the idea more closely, their enthusiasm for it languished, and so did Forbes's campaign. Still, the idea of a flat-rate tax is likely to be an active one on the political scene for years to come.

was required to achieve this goal. Fiscal constraints forced a shorter phaseout range, thus leading to greater marginal disincentives.

NONLINEAR TAX STRUCTURES

The discussion so far has focused on the optimal flat-rate tax. In fact, the United States has had, for a long time, a highly nonlinear schedule, with marginal rates varying from zero to almost 40 percent. Nonlinear tax structures increase complexity and, for reasons that are explained more fully in Chapter 24, increase incentives and opportunities for tax avoidance. At the same time, they can reduce the total deadweight loss associated with attaining any set of revenue and distributive goals.

Earlier, we saw that the deadweight loss is related to the marginal tax rate and the elasticity of the (compensated) labor supply. The basic principles of efficient progressive income taxation are derived from that insight:

1 Impose high average tax rates with low marginal tax rates.
2 Make as few people as possible face high marginal tax rates.
3 Impose high marginal tax rates on those for whom the tax is least distorting.

Figure 20.3 compares two tax structures, a progressive flat-rate tax and a tax structure with high marginal tax rates at low incomes and very low marginal tax rates at very high incomes. *OB* is the lump-sum grant given to ssomeone who does not work and has no other source of income. This grant gets phased out as income rises. The high marginal tax rates (high phaseout rate) over the interval *BC* mean that at incomes beyond *C*, average rates can be higher while marginal rates are lower. For these middle- and upper-income individuals, this means the government is collecting more taxes with less distortion. The price is greater distortions for those with income in the range *BC*. The total deadweight loss will be low, however, if there are relatively few people in this interval, or if those in it have a relatively low labor supply elasticity. Even if they reduce their labor supply significantly, however, the total social loss may be relatively unproductive.

LOWERING TAX RATES FOR THE RICH Figure 20.4 shows why lowering the marginal tax rate for the highest income groups may be desirable. The figure depicts the budget constraint facing the highest income group, with individuals in the group choosing point *E*. The revenue raised is the amount *EA*. If, for those who have income above Y_A, we now lower the marginal tax rate to zero, the government still collects the same revenue; but with a lower marginal tax

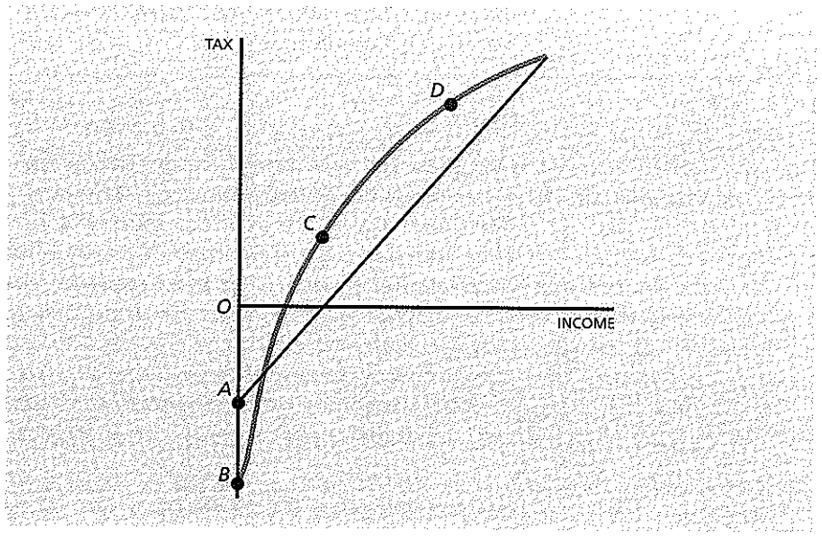

FIGURE 20.3 **Linear versus Nonlinear Tax Structures** Nonlinear tax structures may be able to increase the amount of redistribution without increasing the deadweight loss associated with the tax. The nonlinear schedule *ABCD* has a higher marginal tax rate among the very poor and low marginal tax rates at upper income ranges. On the other hand, higher earners face a higher average tax rate.

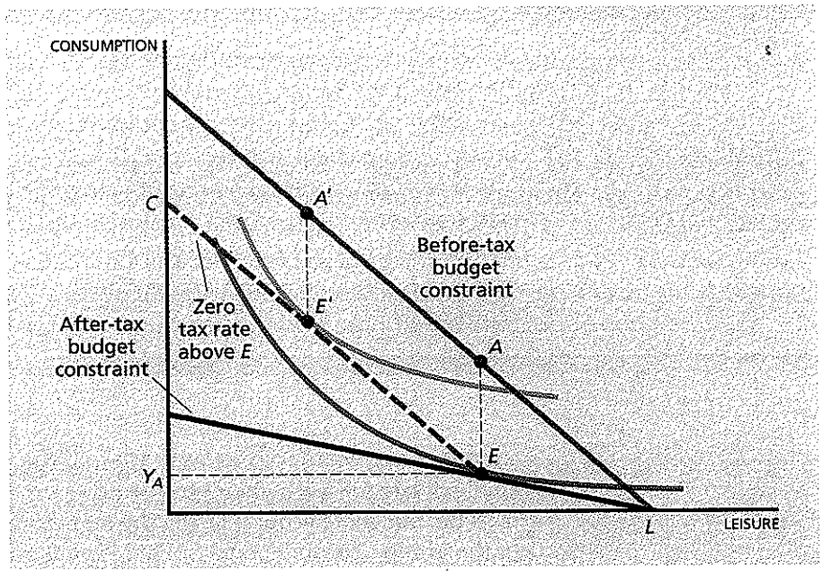

FIGURE 20.4 **Impact of Lowering Marginal Tax Rates for Upper-Income Individuals** Lowering marginal tax rates for upper-income individuals may improve the welfare of this group without reducing government revenue. Here, lowering the marginal tax rate beyond E to zero makes the individual better off, but has no effect on revenue.

rate, these individuals work harder and are still better off. (The new budget constraint is the dotted line $EE'C$, and they choose point E', on a higher indifference curve.) Thus, the tax reform is a Pareto improvement: the rich are better off, and no one is worse off. Now, if instead of imposing a zero marginal tax rate beyond Y_A we had imposed a low marginal rate, the higher-income individuals would still be better off, but there would be additional tax revenues collected, which could be used to reduce taxes on the middle class and/or increase subsidies for the poor. *All individuals could be made better off.*

Such reasoning provided part of the rationale for the reduction in the tax rates at the upper end of the income distribution enacted in 1986.

DIFFERENTIAL TAXATION

The government imposes a huge array of taxes on various commodities, from airline tickets to tires to gasoline to perfume. Taxes that are imposed at different rates on different commodities are called **differential taxes.** Some of these taxes, such as the airline ticket tax, are designed as *benefit taxes*—that is, to make those who benefit from airline travel pay for the costs of the air traffic controller system and airports. Others, such as the taxes on gasoline, tobacco, and alcohol, are partially designed as "*corrective taxes,*" to

ameliorate some of the externalities they generate, such as traffic congestion and air pollution from automobiles. Finally, some, such as the tax on perfume, are *luxury taxes,* intended to increase the redistributive nature of our tax system.

In this section we address two key questions. First, if the government cannot impose an income tax to redistribute income—as is the case in many less-developed countries—what rates should it impose on different commodities? And second, if the government *can* impose an income tax to redistribute income, should it *also* impose taxes on different commodities at different rates? The two questions turn out to have markedly different answers.

RAMSEY TAXES

We begin with an even simpler question posed by the great Cambridge economist Frank Ramsey. Ramsey was not concerned with redistribution, only with efficiency. But he assumed that the government could not impose a lump-sum tax.[9] Hence, it had to raise revenues through commodity taxation. The question he asked was, what is the least distortionary pattern of taxes? Should, for instance, every commodity be taxed at the same rate—in which case the tax is just a tax on income? (Recall the discussion of equivalent taxes in Chapter 18.) That was the answer suggested by those who simply wanted to count distortions, for such a tax would have only one distortion. Ramsey showed not only that that was wrong, but that there was a simple formula for the optimal tax rate.

The commodity taxes that minimize the deadweight loss are called **Ramsey taxes.** Under certain simplifying conditions, Ramsey taxes are proportional to the sum of the reciprocals of the elasticities of demand and supply:

$$\frac{t}{p} = k(1/\eta_u^d + 1/\eta^s),$$

where k is a proportionality factor that depends on the total amount of revenue the government is attempting to raise, t is the per unit tax, p is the (after-tax) price, η_u^d is the compensated elasticity of demand, and η^s is the elasticity of supply. If the elasticity of supply is infinite (a horizontal supply schedule), the tax should simply be inversely proportional to the compensated elasticity of demand. Ramsey's result should not come as a surprise. In Chapter 19 we showed that the deadweight loss from a tax increased with the compensated elasticity of demand and with the elasticity of supply. (Recall Figure 18.8.)[10]

[9] F. Ramsey, "A Contribution to the Theory of Taxation," *Economic Journal* 37 (1927): 47–61. The question had been posed to him by his teacher, A. C. Pigou. See A. C. Pigou, *A Study in Public Finance,* 3rd ed. (London: Macmillan, 1947).

[10] If there is a tax rate t on corporate profits, the Ramsey formula is modified to:

$$\frac{t}{p} = k\left(1/\eta_u^d + \frac{(1-t)}{\eta^s}\right).$$

Hence, if corporate profits are taxed at 100 percent, the tax rate is simply inversely proportional to the elasticity of (compensated) demand.

Figure 20.5 shows the solution to the optimal commodity tax problem. Panel A depicts the deadweight loss as a function of the tax rate imposed on commodity *i*. Panel B shows the revenue raised as a function of the tax rate imposed on commodity *i*. From these two diagrams we can calculate, at each tax rate, the ratio of the increase in deadweight loss to the increase in

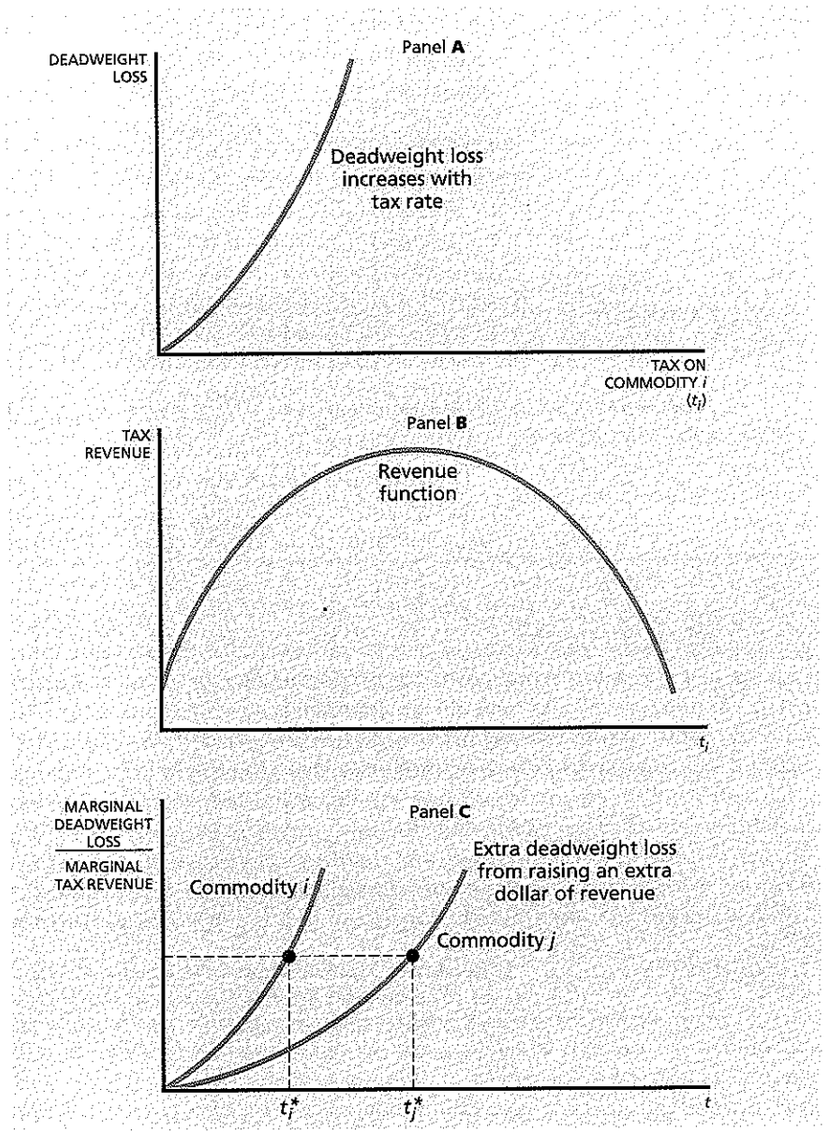

FIGURE 20.5 **Optimal Commodity Taxation** The marginal excess burden (deadweight loss) per marginal dollar raised must be the same for all commodities.

tax revenues from raising the tax a little bit—that is, the marginal deadweight loss from raising an extra dollar of revenue from a tax on commodity *i*. Notice that we have drawn the curve not only so that excess burden increases as the revenue raised increases, but also so that the extra deadweight from raising an extra dollar of revenue increases with the tax rate (and thus with the revenue raised). This follows from the fact that the deadweight loss increases with the square of the tax rate.

A similar curve can be derived for commodity *j*, as shown in panel C. The tax rates should be set so that the increase in deadweight loss per extra dollar raised is the same for each commodity. If the increase in excess burden per extra dollar raised were greater for one commodity than for another, by adjusting tax rates so that one less dollar was raised on the first commodity and one more dollar was raised on the second commodity, total deadweight loss would be reduced.

In panel C, the marginal deadweight loss per marginal dollar of revenue raised is higher for commodity *i* than for commodity *j* at any given tax rate. To equate the marginal deadweight loss per marginal dollar of revenue raised we must impose a lower tax rate on *i* than on *j*. Ramsey's basic insight was to observe that *commodities with low elasticity of demand (or low elasticity of supply) have a lower marginal deadweight loss per marginal dollar of revenue raised, and thus should face higher marginal tax rates.*

OPTIMAL COMMODITY TAXATION WITH INTERDEPENDENT DEMANDS* The result we have just given requires that the compensated demand curves of each commodity are independent; that is, the demand for one commodity does not depend on the price of another. Another interpretation of Ramsey's result holds when supply curves are infinitely elastic, whether or not demand curves are interdependent: *The optimal tax structure is such that the percentage reduction in the compensated demand for each commodity is the same.*[11]

ALTERNATIVE INTERPRETATION: OPTIMAL COMMODITY TAX STRUCTURE WITH INTERDEPENDENT DEMANDS An income tax is distortionary because it induces individuals to make "incorrect" decisions concerning the amount of labor they wish to supply. Commodity taxation may help correct that distortion. If we tax commodities that are complements for leisure and subsidize commodities that are complements for work, we encourage individuals to work, and thus reduce the distortion caused by a uniform commodity tax (which is equivalent to just a wage tax). For instance, by taxing ski equipment and subsidizing commuter costs we induce individuals to work more and consume less leisure.[12]

* This subsection and the remaining subsections of this part of the chapter deal with more advanced topics and can be omitted.

[11] Note that if $\eta^s = \infty$, $t/p = k/\eta_u^d$; with horizontal supply curves, the percentage tax is inversely proportional to the compensated demand elasticity. The percentage change in output is equal to the percentage increase in price × percent change in demand from a percent change in price = $(k/\eta_u^d) \times \eta_u^d = k$—i.e., it is the same for all commodities.

[12] This interpretation was noted in W. J. Corlett and D. C. Hague, "Complementarity and the Excess Burden of Taxation," *Review of Economic Studies* 21 (1953): 21–30.

RAMSEY TAXES

In the absence of any income or profit taxes, and with all individuals identical, raising revenues so as to minimize deadweight loss requires imposing taxes in inverse relationship to the elasticity of demand and supply.

REDISTRIBUTION AND RAMSEY TAXES: COMMODITY TAXATION IN LDCs There is one very disturbing feature of Ramsey's analysis. The major reason why governments use distortionary rather than uniform lump-sum taxes is that they have certain redistributive goals that they cannot achieve otherwise. However, the early discussions of optimal taxation assumed that all individuals were identical (in which case the natural assumption would be that the government would employ uniform lump-sum taxation).

This was particularly vexing since the results described earlier suggest that high tax rates should be imposed on commodities with low price elasticities, such as food. These commodities often have low income elasticities, so that if a high tax is imposed on them the poor will bear a larger burden than the rich. But the original reason for employing commodity taxation was to shift more of the burden onto the rich than they would face, say, with a uniform lump-sum tax. Ramsey's analysis thus seemed to provide little guidance for any serious policy analysis and was, accordingly, largely dismissed.

Subsequent research has extended Ramsey's original analysis to include redistributive goals.[13] Not surprisingly, whether one wishes to tax income-elastic and price-elastic commodities, such as perfume, at a higher or lower rate than income-inelastic and price-inelastic commodities like food depends, in part, on the strength of one's concern for income redistribution.

Less developed countries typically place little reliance on income taxes, since they have difficulty monitoring income. Indeed, in many cases, they cannot even impose a tax on all commodities, but only on commodities that are imported or exported (since they have some control over what passes over their borders), and on commodities manufactured in the urban sector. Most LDCs have sufficient concern for redistribution that they tax luxuries at higher rates than basic necessities.

[13] See, in particular, Peter Diamond and James Mirrlees, "Optimal Taxation and Public Production, I: Production Efficiency and II: Tax Rules," *American Economic Review* 61 (1971): 8–27 and 261–78; P. Diamond, "A Many-Person Ramsey Tax Rule," *Journal of Public Economics* 4 (1975): 335–42; A. B. Atkinson and J. E. Stiglitz, "The Structure of Indirect Taxation and Economic Efficiency," *Journal of Public Economics* 1 (1972): 97–119; and A. B. Atkinson and J. E. Stiglitz, "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics* 6 (1976): 55–75; reprinted in A. B. Atkinson (ed.), *Modern Public Finance,* vol 2. International Library of Critical Writings in Economics, no. 15 (Aldershot, U.K., and Brookfield, Vt.: Elgar, 1991), pp. 82–102.

DIFFERENTIAL COMMODITY TAXES IN ADVANCED COUNTRIES WITH PROGRESSIVE INCOME TAXES

All advanced industrialized countries, however, do have a progressive income tax. For them, the issue is markedly different from that posed by Ramsey. They ask: *If* there is an optimally designed income tax, does the marginal benefit of the extra redistribution which, say, a tax on luxuries provides exceed the marginal cost, in terms of the excess deadweight loss? The naïve answer to this question was rejected in the introduction to this chapter. Earlier discussions had suggested that introducing more distortions was bad, and that therefore differential commodity taxation was undesirable, but this fallacy was dismissed: one simply cannot count the number of distortions. Yet, remarkably enough, the conclusion of these earlier discussions was correct: If an income tax is well designed, adding differential commodity taxation is likely to increase the ability to redistribute income little, if at all. The objective of taxation is to redistribute income, or to impose the burden of taxation on those most able to afford it, and it turns out that the best way to do this, after all, is to focus taxation on what we are really interested in, namely income.[14]

INTEREST INCOME TAXATION AND COMMODITY TAXATION

In our earlier discussion we showed how a tax on interest income discourages future consumption. It changes the slope of the budget constraint in the same way that a tax on future consumption only would.

Thus an income tax that taxes interest can be viewed as a differential commodity tax in which future consumption is taxed more heavily than current consumption. The question whether it is desirable to tax interest income is then equivalent to the question whether it is desirable to tax future consumption at higher rates than current consumption.

Just as little may be gained by adding differential commodity taxation with a well-designed income tax, so little is gained from taxing consumption at different dates at different rates. This means, in effect, that interest income should be exempt from taxation. An income tax that exempts interest income is, of course, equivalent to a wage tax, and we showed in Chapter 18 that a wage tax was equivalent to a consumption tax (in the absence of bequests). This suggests that it may be optimal to have a consumption tax. We discuss this further in Chapter 25.

TAXES ON PRODUCERS

So far this chapter has focused on taxes on households, on their wage and interest income and their consumption. Many people believe that it is only fair that firms pay taxes too. Such reasoning is misguided: firms never bear the incidence of a tax, as we have seen, but individuals do, as shareholders, workers, or consumers. Figuring out the incidence of taxes on corporations is a complicated matter.

[14] Indeed, under standard assumptions, Pareto efficient taxation requires that there be no differential taxation of commodities. See Atkinson and Stiglitz, "The Design of Tax Structure."

But we can ask a more general question: Does Pareto efficient taxation imply that taxes should be imposed on production processes? The taxes described thus far interfere with one of the three conditions for Pareto efficiency discussed in Chapter 3, product mix efficiency: the marginal rate of transformation differs from the marginal rate of substitution.[15, 16] Do we want to maintain production efficiency, even if we cannot maintain product mix efficiency?

Many of our taxes also affect the production efficiency of the economy, which is to say that they result in the economy's not being on its production possibilities schedule. Production efficiency requires that the marginal rate of technical substitution between any two inputs be the same in all firms, and that the marginal rate of transformation between any two outputs (or between an input and an output) be the same in all firms. Productive efficiency is attained when all firms face the same prices for inputs and outputs. Thus, any tax on an input that is not uniform across all firms, or any tax on an output that is not uniform across all firms, results in the economy's not being productively efficient. For instance, the corporation income tax is widely viewed as a tax on capital inputs used in incorporated firms, because it raises the after-tax cost of capital in corporations above that in unincorporated businesses. Also, while gasoline that is used for most business purposes is taxed, gasoline used for farming is not. But these are only the most obvious examples.

Many production activities are performed in both the market and nonmarket sectors. Only activities performed in the market sector are taxed. Thus an individual driving himself to work is performing the same service that a taxicab driver who drives the individual to work performs. Yet there is a tax on the latter and not on the former. A person who bakes a loaf of bread at home is performing a service similar to that of a baker but is not taxed in the same way that the baker is taxed. There is thus a distortion between the marketed and nonmarketed sectors, and the economy is not productively efficient.

Any tax on *intermediate goods*—goods used to produce other goods—is distortionary. To see this most clearly, consider a firm that produces and uses computers in its own production plants; the cost of the computer is simply the cost of the factors of production (including the return to capital employed in the production). In a competitive economy this firm would be forced to sell the computers at its costs of production, so that the cost of any

[15] With a tax on wage income, the marginal rate of transformation, the wage, exceeds the marginal rate of substitution (the after-tax wage); with differential commodity taxes, relative producer prices (which equal the marginal rate of transformation) differ from relative consumer prices.

[16] We can also ask, if it is possible to charge individuals with different incomes different taxes on consumption, whether it is desirable to do so. In other words, is it desirable to maintain *exchange efficiency*? Under the conditions under which no differential commodity taxation is desirable, of course, there is exchange efficiency in the consumption of all goods; but when differential taxation is desirable, it is also in general desirable to have relative tax rates dependent on income.

other firm using a computer would be the same as the cost of the manufacturing firm in using it. But now, when a sales tax is imposed, the cost to the firm manufacturing the computer and using it is less than the cost to another firm using the computer in its production processes. There is thus an important distortion, and the economy is no longer productively efficient.

Should the government impose such distortionary taxes if it wishes to minimize the deadweight loss of the tax system? One naïve answer to this question is to say, of course not; the government should not introduce any additional distortions that it does not need to. This kind of argument is similar to the arguments we discussed earlier concerning differential commodity taxes. It makes no sense simply to count the number of distortions. Yet it turns out that under some circumstances, the conclusion of the naïve argument is correct.

If the government is able to tax away all profits in the private sector, and if there are no other restrictions on the ability of the government to impose taxes (other than the ability to impose lump-sum taxes), it is possible to show that productive efficiency is desirable. Hence, the government should impose no distortionary taxes on businesses. *Whatever the government could do with a distortionary tax on producers, it could do better with a direct tax on consumers that maintained the economy on the production possibilities schedule.*[17]

This analysis has some very strong implications. It suggests, in particular, the undesirability of import duties and of taxes on corporations that differ from taxes on unincorporated businesses.

There are many instances, however, when governments face difficulties in imposing taxes. For instance, governments cannot distinguish between final consumer use of a commodity and the use of the commodity by a business; thus, if a government is to impose a tax on consumers it must also impose a tax on business use. Whenever the government is not able to identify and tax away *all* pure profits in the private sector, and whenever there are other restrictions on the ability of the government to impose taxes, it may be desirable to impose distortionary taxes on producers.[18]

But the basic insight, suggesting that one look unfavorably on taxes that interfere with productive efficiency, is still a valuable one. Taxes on imports, for example, introduce an important inefficiency in the economy; at least in more developed countries, governments can impose a tax on the consumption of these goods rather than on just imports; and, *in general*, such consumption-based taxes are preferable.[19]

[17] This result was originally established in the important paper by Diamond and Mirrlees, "Optimal Taxation and Public Production, I: Production Efficiency." See also Alan J. Auerbach, "The Theory of Excess Burden and Optimal Taxation," Chapter 2 in *Handbook of Public Economics,* vol. 1, pp. 100–101.

[18] This result was established in J. E. Stiglitz and P. Dasgupta, "Differential Taxation, Public Goods and Economic Efficiency," *Review of Economic Studies* 39 (1971): 151–74.

[19] For a more extended discussion of the relationship between trade taxes and commodity taxes, see P. Dasgupta and J. E. Stiglitz, "Benefit-Cost Analysis and Trade Policies," *Journal of Political Economy* 82 (January–February 1974): 1–33.

THE DEPENDENCE OF OPTIMAL TAX STRUCTURE ON THE SET OF AVAILABLE TAXES

Throughout this chapter, we have noted the dependence of the optimal tax results on the assumptions made concerning the set of available taxes. This was particularly true for commodity taxation. Whether there should be differential commodity taxation, and, if so, how the difference in rates should be chosen, depends on whether there is an income tax and if there is, on its structure. Ramsey showed that *in the absence of any income tax* (and assuming no redistributional objectives), different commodities should be taxed at different rates depending only on the elasticities of demand and supply. When there is an optimally chosen income tax, it may not be desirable to impose differential commodity taxes. When it is desirable to impose differential commodity taxes, they do not depend simply on the elasticities of demand.[20]

It should be emphasized, however, that the set of taxes that is feasible should itself be a subject for analysis: it depends, in particular, on what variables are easily observable and verifiable. In developing countries in which there are many barter transactions (trade not for cash) and in which the level of record keeping is low, it is difficult to enforce an income tax, and commodity taxes must be relied on to redistribute income and to ensure that the burden of taxation is equitably borne. But in the United States, the case for the use of redistributive commodity taxation is weak.

[20] The central question is whether the additional redistribution that might be obtained from differential commodity taxation is worth the extra deadweight loss.

When there is a flat-rate income tax, with the tax rate chosen optimally, the optimal tax rate on a commodity is simply inversely proportional to the elasticity of demand and proportional to a parameter that measures the extent to which the good is consumed relatively more by the rich (so that a tax on that good is progressive). In some simple cases, that distributional parameter itself is proportional to the price elasticity of demand; goods with low elasticities of demand (like food) have low deadweight losses but a tax on them is regressive. The two effects (efficiency, or deadweight loss, and distribution) are offsetting, and there should either be no differential taxation on different commodities, or it should depend on parameters other than the elasticity of demand.

In the more general case where an optimal income tax can be imposed that is not necessarily flat (that is, marginal rates can vary with income), a critical determinant of the commodity tax structure is how the marginal rate of substitution between two commodities depends on leisure; in the case where marginal rates of substitution among commodities do not depend at all on leisure, there should be no differential commodity taxation.

REVIEW AND PRACTICE

SUMMARY

1 Pareto efficient tax structures are such that there is no alternative that can make any individual better off without making some other individual(s) worse off. The nature of the Pareto efficient tax structure, in turn, depends on the information available to the government.

2 There are important trade-offs between distributional goals and efficiency in the design of tax structures. The optimal tax structure balances the gains from additional redistribution with the costs in terms of loss in efficiency.

3 The deadweight loss associated with the magnitude of the substitution effect suggests that it is desirable to have low marginal tax rates in those parts of the income distribution where there are a large number of individuals, which is to say in the middle income ranges. On the other hand, high marginal rates in such ranges enable the government to collect the same or greater revenue with a lower marginal tax rate from upper-income individuals. This reduces the deadweight loss per dollar of revenue raised from upper-income individuals.

4 Ramsey taxes minimize the deadweight loss associated with raising a given revenue through commodity taxes alone. In the simple case of independent demand and supply curves, the higher a good's supply and compensated demand elasticities, the lower the tax rate on a good.

5 Whether different commodities should be taxed at different rates depends on the taxes that are available to the government. If the government has imposed an optimal income tax, there may be little if any gain from the imposition of differential commodity taxes.

6 If there are no pure profits in the private sector (the economy is perfectly competitive, or the government can impose a 100 percent profits tax) and if there are no other restrictions on the ability of the government to impose taxes, then the government should not impose any taxes that interfere with the productive efficiency of the economy. When these stringent assumptions are removed, it may be desirable to introduce taxes that interfere with productive efficiency.

KEY CONCEPTS

Theory of the second best
Pareto efficient taxe structure
Optimal tax system
Flat-rate taxes
Negative income tax
Differential taxes
Ramsey tax

QUESTIONS AND PROBLEMS

1 "If there are groups in the population who differ in their labor supply elasticity, they should be taxed at different rates." Justify this in terms of the theory of optimal taxation, and discuss its implications for the taxation of working spouses.

2 Earlier, we noted that consumption at different dates could be interpreted just like consumption of different commodities at the same date. What do the results on optimal taxation imply about the desirability of taxing interest income? (Hint: Recall that the price of consumption tomorrow relative to the price of consumption today is just $1/1 + r$, where r is the rate of interest.)

3 Explain why it might be desirable to have a regressive tax structure, even if the social welfare function is utilitarian, when general equilibrium effects of taxes are taken into account. Would it ever be desirable to impose a negative marginal tax rate on very high income individuals?

4 If you believed that those who were more productive in earning income also had a higher marginal utility of income (they were more efficient in consumption), what would that imply for the design of tax structures? Discuss the reasonableness of alternative assumptions.

5 Under what circumstances will an increase in the progressivity of the tax schedule increase the degree of before-tax inequality?

6 To what extent do you think that differences in views concerning how progressive our tax structure should be reflect differences in values, and to what extent do they reflect differences in judgments concerning the economic consequences of progressivity (deadweight loss, shifting)?

7 One argument sometimes made in favor of the use of commodity taxation rather than income taxation is that people do not accurately perceive the amount they pay in commodity taxes. They will object less to a 20 percent income tax supplemented by a 10 percent sales tax than to a 30 percent income tax. Do you think this is true? If it is, what do you think it implies about the design of tax policy?

8 Explain why the EITC may actually lower total work effort of the poor even if it increases labor force participation. (Hint: Focus separately on those below and above the maximum benefit level.)

APPENDIX A DERIVING RAMSEY TAXES ON COMMODITIES

The formula for Ramsey taxation, given horizontal supply schedules, may be derived using calculus and certain standard results from microeconomic theory. We represent the individual's utility by her *indirect utility function*, giving her level of utility as a function of consumer prices (p_1, p_2, p_3, . . .)

and of income (I): $V = V(p_1, p_2, p_3, \ldots, I)$. A standard result[21] is that the change in utility from a change in price is just equal to the (negative of the) quantity consumed times the marginal utility of income $\frac{\partial V}{\partial I}$:

$$\frac{\partial V}{\partial p_i} = -Q_i \frac{\partial V}{\partial I}.$$

Let us now increase the per unit tax on, say, the first commodity (t_1) and reduce the per unit tax on the second commodity (t_2) in such a way as to leave utility unchanged. Since with horizontal supply curves producer prices are fixed, the change in the consumer price is just equal to the change in the tax: Then $dp_1 = dt_1 > 0$, $dp_2 = dt_2 < 0$. Clearly, to keep utility unchanged, the required change in the tax on the second commodity must satisfy $dV = \frac{\partial V}{\partial p_i} dt_1 + \frac{\partial V}{\partial p_2} dt_2 = 0$. We can substitute in the values of $\partial V/\partial p_1$ to obtain

$$\frac{dt_2}{dt_1} = -\frac{Q_1}{Q_2}.$$

Thus, if the quantity consumed of the first commodity is large (so the loss in welfare from the tax increase is large), the reduction in taxes on the second commodity must be large.

If the demand for each commodity depends only on its own price, then the change in revenue induced by an increase in the tax on the first commodity is just

$$\frac{\partial(t_1 Q_1)}{\partial t_1} = Q_1 + t_1 dQ/dp_1 = Q_1\left(1 + \frac{t_1}{p_1}\frac{dQ_1}{dp_1}\frac{p_1}{Q_1}\right) = Q_1\left(1 - \frac{t_1}{p_1}\eta_u^1\right),$$

where η_u^1 is the compensated demand elasticity for good 1. The term $t_1 \frac{dQ_1}{dp_1}$ represents the *loss* in revenue resulting from reduced sales in response to the changed price. The reason why it is the compensated demand elasticities that are relevant is that we are considering variations in two tax rates that, together, *leave the individual at the same level of welfare.*

Similarly, for each change in the tax on the second commodity, the change in revenue is given by

$$Q_2\left(1 - \frac{t_2}{p_2}\eta_u^2\right).$$

[21] This result is known as Roy's Identity. For a proof, see H. Varian, *Microeconomic Analysis* Third Edition (New York: Norton, 1982), pp. 106–7, or A. Deaton and J. Muellbauer, *Economics and Consumer Behavior* (London: Cambridge University Press, 1980), pp. 37–41.

The total change in revenue is thus

$$\frac{dR}{dt_1} = Q_1\left(1 - \frac{t_1}{p_1}\eta_u^1\right) + \frac{dt_2}{dt_1}Q_2\left(1 - \frac{t_2}{p_2}\eta_u^2\right)$$

$$= Q_1\left[\left(1 - \frac{t_1}{p_1}\eta_u^1\right) - \left(1 - \frac{t_2}{p_2}\eta_u^2\right)\right] = Q_1\left[\frac{t_2}{p_2}\eta_u^2 - \frac{t_1}{p_1}\eta_u^1\right].$$

With an optimal tax structure, this must be zero, i.e., given that we are keeping the level of utility of the individual constant, revenues must be maximized. But this requires that

$$\frac{t_2}{p_2}\eta_u^2 - \frac{t_1}{p_1}\eta_u^1 = 0.$$

Generalizing this condition to all commodity taxes, $t_1, t_2, \ldots t_i, \ldots$, we know that $\frac{t_i}{p_i}\eta_u^i$ must be the same for all, that is, for all commodities. Let k be that value, so that

$$\frac{t_i}{p_i} = \frac{k}{\eta_u^i}.$$

This means that tax rates must be inversely proportional to compensated demand elasticities. This is the Ramsey rule.

APPENDIX B DERIVATION OF RAMSEY FORMULA FOR LINEAR DEMAND SCHEDULE

Figure 20.6 illustrates a linear compensated demand schedule, $Q = a - b(p + t)$, with a fixed producer price (infinite elasticity supply schedule) and a tax t. The slope of the demand schedule is b. The deadweight loss

$$\text{DWL} = \tfrac{1}{2}\, bt^2,$$

so the marginal deadweight loss from increasing the tax is

$$\text{MDWL} = bt.$$

The revenue raised by the government is

$$R = tQ = at - b(pt + t^2),$$

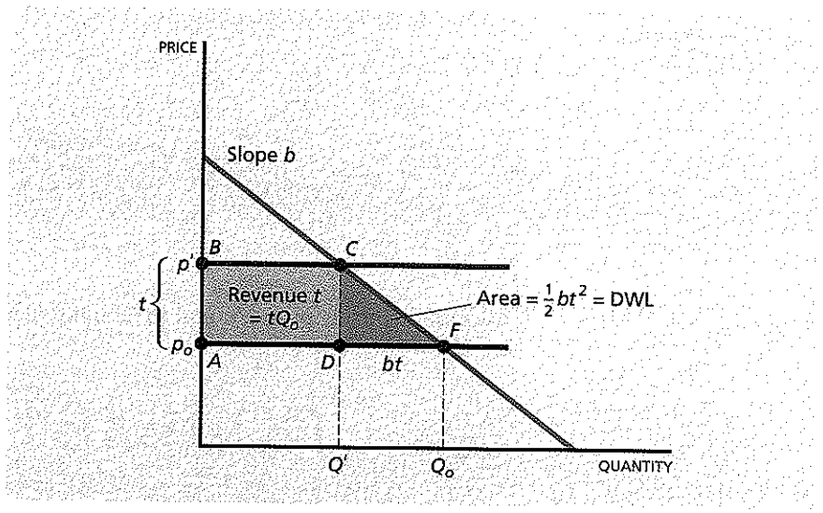

FIGURE 20.6 **Ramsey Pricing Calculation with Linear Demand Curves** With a linear demand schedule, the revenue raised by a tax at the rate t is the shaded square $ABCD$ (equals tQ_o', where t is the tax rate and tQ_o' is the output after the tax. The deadweight loss is the triangle DCF, where DC equals the tax, t, and DF equals the change in output, which is just bt, where b is the slope of the demand curve. The total deadweight loss is just $\frac{1}{2}\, bt^2$. Ramsey looked at the extra deadweight loss associated with raising an extra dollar of revenues.

so the marginal revenue from increasing the tax is

$$\text{MR} = a - b\,(p + 2t).$$

The ratio of marginal revenue to marginal deadweight loss is

$$\frac{\text{MR}}{\text{MDWL}} = \frac{a - b(p + 2t)}{bt}$$

$$= \frac{Q}{bt} - 1$$

$$= k', \text{ the same for all commodities}$$

or

$$\frac{Q}{bt} = 1 + k' \quad \equiv 1/k$$

or

$$t = k\,Q/b.$$

But the elasticity of demand is just

$$\eta_u^d = -\frac{\Delta Q/Q}{\Delta p/p} = \frac{bp}{Q}$$

so

$$\frac{t}{p} = \frac{kQ}{bp} = \frac{k}{\eta_u^d},$$

taxes are inversely proportional to demand elasticities.