Optimal Progressivity

To this point, we have assumed that all individuals are the same. To consider the distributional impact of the tax system, we will have to alter that assumption. We have seen that differential lump sum transfers can move the economy along the Pareto frontier. However, the fact that a point like A (Figure 11) is Pareto efficient does not necessarily imply that it is better than point B. As Samuelson said, "Efficiency has no normative implications." The question of how far the government should go in moving toward total equality, represented by the 45° line in Figure 11, cannot be answered by an analysis based purely on efficiency. In fact, there are several reasons that the government might not choose to redistribute income at all. The issue is divisive because the government must take income from some in order to give it to others. The economic cost might even be so high that redistribution is not indicated in practical terms, even if it is preferred in theory.

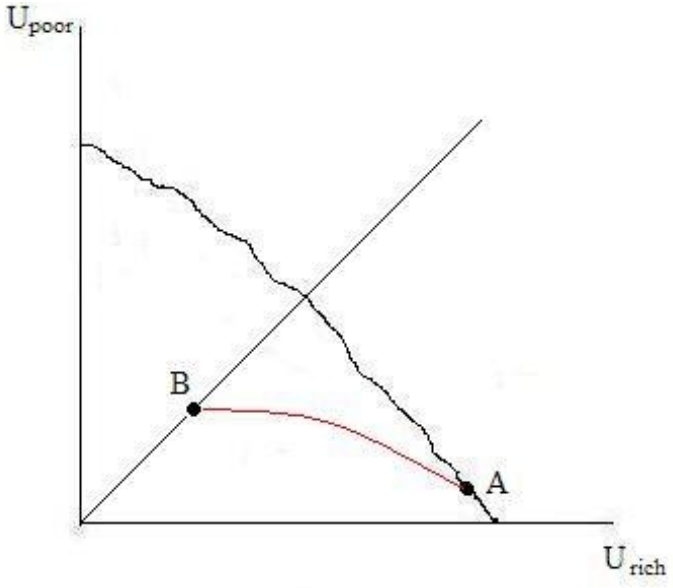


Figure 11

Redistribution is generally accomplished by the use of progressive taxation. Progressive taxation means that the average tax rate rises with income, in contrast to regressive taxation, in which the average tax rate falls with income. A proportional tax system has an average tax rate that is constant over income.

The analysis of redistribution in public finance has changed over time. In the past, it was based primarily on two principles: benefit and ability to pay. The benefit principle considers taxes a charge for services provided by the government; according to this view, the burden of provision by the government

should be borne by those who benefit from the service provided. This principle is easy to apply to services like roads, where fees for use are simply assessed. However, the burden of a public good like national defense is more difficult to apportion. Because those who are well-off derive more benefit from the government's protections, this principle implies that progressivity is desirable, but it does not give much guidance on the form of that progressivity. Finally, the benefit principle precludes redistribution; any benefit distributed must be paid for. The ability to pay, or equal sacrifice, principle calls for increased well-being to be associated with increased taxes. In this framework, the sacrifice of a dollar in taxes is smaller the more one has, so more income implies more taxes. This criterion, like the benefit principle, suggests progressivity but does have concrete conclusions regarding exactly how much more the well-off should pay in taxes.

Over time, the use of these principles has been replaced with a different framework. Modern public finance examines the Pareto frontier and asks, what are the consequences of moving toward full equality? Because redistribution must be accomplished with the available, second-best instruments, it is costly. Economists now concentrate on discovering the minimum efficiency cost of redistribution, tracing out the most efficient second-best outcomes for a given amount of redistribution. This is represented by the red curve in Figure 11. The difference between that curve and the Pareto frontier is the cost of redistribution, which rises as the amount of redistribution rises. However, while economists can give guidance regarding the location of second-best outcomes, choosing the correct balance between redistribution and efficiency requires value judgements about social welfare.

The social welfare functions we will consider have two common characteristics. First, they are Paretian; social welfare increases when any individual's well-being increases and no others' decrease. Second, they are anonymous; social welfare depends only on the amount of income held by each individual, not on who specifically has that income. This rules out cases in which the same level of income produces different levels of well-being for different people.

A utilitarian social welfare function values social welfare as simply the sum of individual welfare, $W = U_1 + U_2$. In this framework, the first-best outcome would be complete equality. The second-best outcome is where a 45 degree line is tangent to the second-best curve (see Figure 12).

A Rawlsian social welfare function, also known as a maximin function, has social welfare equal to the minimum individual welfare, $W = \min(U_1, U_2)$. The Rawlsian function places no premium on equality as long as the well-being of the worst-off improves. For example, a transfer that makes the rich much better off, the near poor much worse off, and gives one cent to the poorest person is considered an improvement by Rawlsian standards. In this case, social indifference curves are L-shaped, as shown in Figure 13.

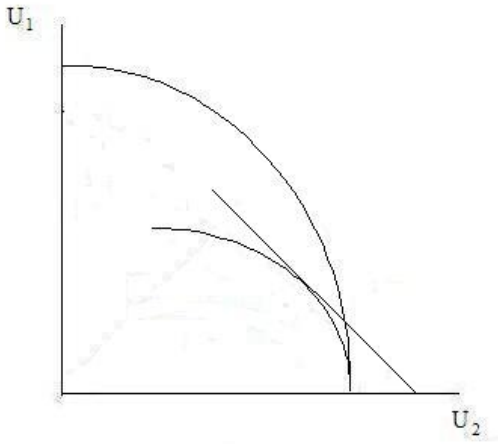


Figure 12

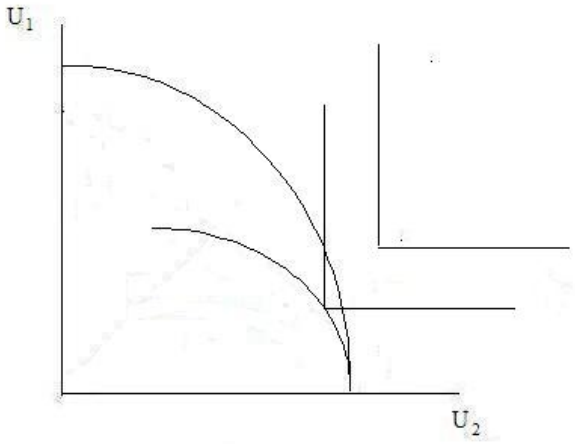


Figure 13

Generalized utilitarian social welfare functions cover a wide variety of social preferences between these two extremes. An example is $W = \frac{1}{\alpha}U_1^{\alpha} + \frac{1}{\alpha}U_2^{\alpha}$. The parameter α is generally less than one (α equal to one is the strict utilitarian

function discussed above, while the function moves toward Rawlsian as $\alpha \to \infty$). As α falls, the curvature of the social indifference curves increases and society places an increasing value on equality. The changes in societal attitudes toward equality symbolized by changes in α will have implications for the optimal degree of progressivity in the tax system. Risk aversion may be important for choosing the correct alpha if the appropriate decisions are made from behind the Rawlsian veil of ignorance. The social welfare functions considered here do not rule out altruism; the income of others can enter into the individual's utility function. The level of inequality can also enter directly into the social welfare function. However, such social preferences can choose allocations in which everyone has low but relatively equal well-being over those in which the level of inequality and every individual's welfare are higher.

Once the preferred level of redistribution is selected, the proper instruments must be chosen. If income taxes are infeasible, a consumption tax could be used for redistribution. We have seen that purely efficient commodity taxation will under certain assumptions follow the inverse elasticity rule, $t_i \eta_i = k$. Thus, the efficient ratio of taxes on two commodities, which is independent of the revenue requirement, is

$$\left(\frac{t_1}{t_2}\right)^* = \frac{\eta_2}{\eta_1}$$

If the distributional characteristics of each good are taken into account, the ratio for the optimal commodity tax becomes

$$\left(\frac{t_1}{t_2}\right)^* = \frac{\eta_2}{\eta_1}\left(\frac{\lambda - R_1}{\lambda - R_2}\right)$$

where R_1 and R_2 capture the distributional characteristics of the respective good, or the weighted marginal social utility of those consuming the good. If, for example, good one is food, while good two is yachts, R_1 will be larger than R_2. The λ parameter is the loss in social utility for every extra dollar raised in taxes. However, in most cases, commodity taxes will be the optimal instrument for redistribution only when the income tax is unavailable. This is because consumption taxes are poorly targeted if purchase baskets are similar across income groups; low taxes on food redistribute to food purchasers, which is not necessarily exactly the target group. Of course, in some cases, the targeted group will be easily identified by purchases; targeted redistribution toward those in poor health will be simply accomplished by a subsidy on health expenditures.

Optimal Linear and Nonlinear Income Taxation

To accomplish redistribution, additional revenue must be raised, which will entail higher marginal tax rates. Higher marginal tax rates on income induce additional distortion in labor supply. To further explore the relationship between distortion and the average marginal tax rate, a graphical illustration of the relationship between systems of taxation and average/marginal tax rates is helpful:

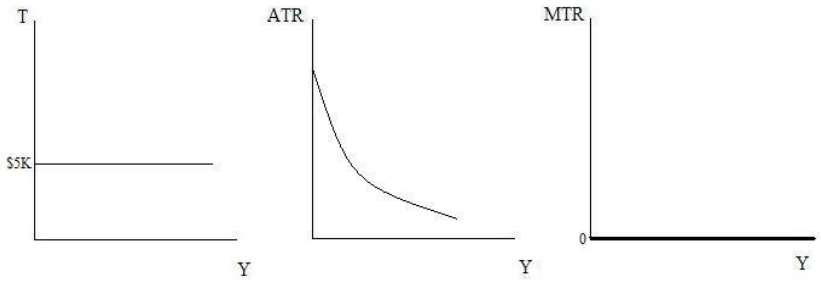


Figure 14

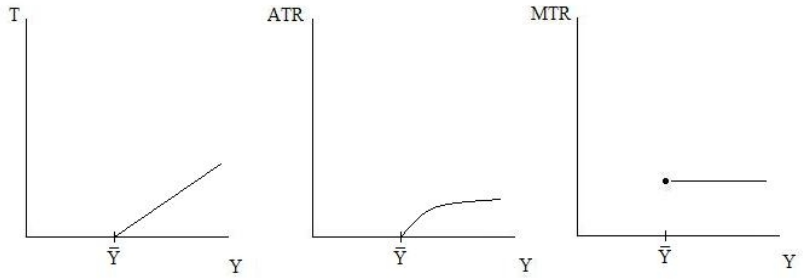


Figure 15

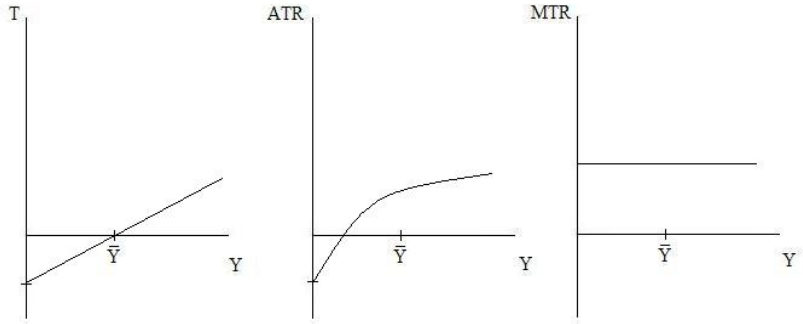


Figure 16

With a uniform lump sum tax, as in Figure 14, the average tax rate falls with income; the marginal tax rate is constant at zero. A flat tax with exemption, or a proportional tax, has an average tax rate that rises with income, while the marginal tax rate is constant after the exemption point (Figure 15). A

negative income tax has a rising average tax rate, with marginal tax rates that are constant from zero income (Figure 16). Note that the negative income tax implies a higher marginal tax rate on the poor even though they are better off. Higher marginal tax rates on low incomes are also applied to those above them, which raises additional revenue that can be used for redistribution. This additional revenue is raised from those with high incomes without the increased distortion of high marginal rates. Keeping marginal rates high only at the low end means that the labor supply distortion is relatively less for high-income people. Therefore, marginal tax rates themselves are not a good measure of what is happening to the poor.

We generally assume that the elasticity of labor supply is the same across incomes, but if high-income individuals have more elastic labor supply, there is an additional argument for marginal tax rates that fall with income. Basic models also assume that movement between labor supply and leisure is the only margin of response, but the form of income may be important for high-income individuals. If high-income individuals do a lot of income switching, their overall elasticity of taxable income may be high. On the other hand, labor force participation may be a very elastic margin for those with low incomes. Variations in assumptions about the relevance and size of these elasticities will have implications for optimal progressivity.

Mirrlees (1971) Model of Optimal Income Tax Progressivity In considering the optimal level of progressivity in the income tax, we will use five basic assumptions:

1. People differ only in earning ability, which is equivalent to the wage rate.
2. There is no capital income.
3. People value both consumption and leisure and have identical preferences.
4. The revenue requirement is exogenous. Note that the revenue requirement could be zero, if there is only redistribution and no financing of public goods.
5. An ability tax is not possible.

Using these assumptions, Mirrlees developed five conclusions about any general income tax and transfer system:

1. The system is very difficult analytically.
2. There are few general results. One of these is that the marginal tax rate on the very top income earner should be zero.
3. Numerical simulations show that the optimal structure is approximately linear, with tax rates always below 40% and usually in the 20-30% range. Optimal marginal tax rates decline with income.

4. Even with a Rawlsian social welfare function, the highest marginal tax rate is below 50%.

5. The labor supply elasticity used is critical to the results.

The lack of many general conclusions stems from the general nature of the problem. The system Mirrlees analyzed had no limits on the sign or structure of taxes. However, there are some results that hold regardless of assumptions about the social welfare function or the distribution of abilities in the population. The marginal tax rate should always (trivially) be between zero and one hundred percent, except on the top income earner, whose optimal marginal rate is zero. This does not imply that the top income earner should have an average tax rate or total tax burden equal to zero. To see why the optimal top rate is zero, imagine the following: If the top earner's marginal rate is more than zero, what happens if it is reduced to zero? The top person works more, since the return to working has increased and any income effect would be small at such a high income. Deadweight loss has decreased, and the top earner is better off. The extra labor supply is taxed, and the proceeds can be distributed to other individuals. This is a Pareto improvement.

The key intuition here is that marginal tax rates on the highest earner distort his or her labor supply, but since there is no one for whom the rate is inframarginal, it raises no additional revenue. Increasing the tax on the 10,001st dollar distorts the behavior of those at $10,000, but it also picks up additional revenue from every individual who earns more than that. This additional revenue can be redistributed to everyone with income below $10,000, making everyone better off. The only individual that this argument does not work for is the highest earner.

To come up with more specifics regarding the optimal income tax, Mirrlees used numerical simulations. The simulations involve choosing a social welfare function, labor supply elasticity, distribution of abilities, and revenue requirement, then solving for the optimal tax structure. With a representative set of assumptions, including strict utilitarian social welfare function, Cobb-Douglas individual utility, a fairly large elasticity, log-normal ability distribution, and 20% of GDP required as revenue, a few representative results emerge. First, the optimal tax structure has approximately constant marginal rates that decrease slightly as income rises. The tax structure allows for transfers to those at the bottom, so that in some cases, the optimal transfer will encourage low wage earners to actually drop out of the labor force. If a Rawlsian social welfare function is used instead, optimal marginal tax rates increase, but the most significant difference is in marginal tax rates for the near-poor, which are much higher than in the utilitarian case. An important conclusion to draw from these results is that intuition is rarely a good guide to translating preferences about equality into a tax system.

Optimal Linear Income Tax The results of the Mirrlees model suggest that a linear income tax will be close to optimal, and it is much simpler to analyze.

For a single-bracket income tax, labor supply elasticity and preferences regarding equality will dictate whether a very progressive system with a large grant and high tax rate (Figure 17a) will be preferred to a less progressive system with a lower rate to finance a smaller grant (Figure 17b). For a two-bracket income tax, there are four choice variables: the tax rate for each bracket, the amount of the grant, and the income level that divides the brackets. In most simulations, a system with decreasing marginal rates (Figure 18b) is preferable to one in which marginal tax rates rise with income (Figure 18a), because lower top marginal rates encourage labor supply, which finances a larger grant. It is important to note that a system with decreasing marginal rates can still be a progressive tax system if the average tax rate is increasing with income. Most developed economies have decreasing overall marginal rates when the whole tax/transfer system, including welfare phaseout rates, is considered.

The rate schedule is not the only policy instrument available to the government. Administration, enforcement, and base broadening can lower the elasticity of taxable income itself and allow the collection of more revenue for a given deadweight loss.

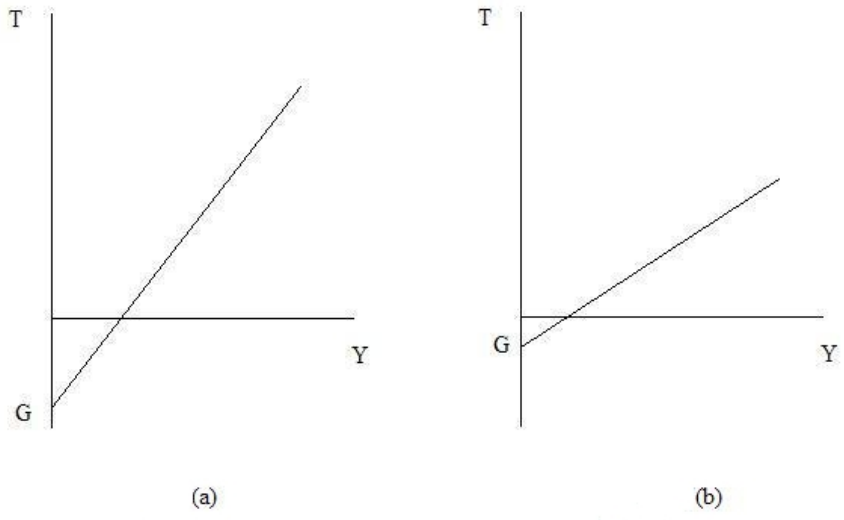
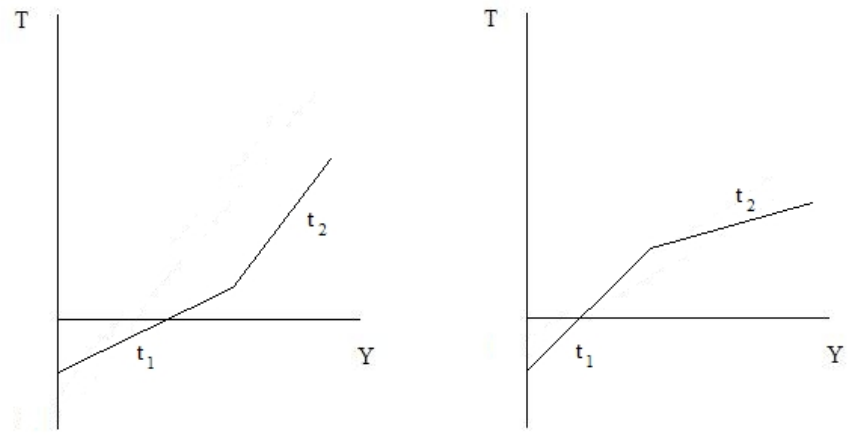


Figure 17



(a) $t_2 > t_1$ (b) $t_2 < t_1$

Figure 18

Tagging and Differential Consumption Taxes

If the government has information on immutable characteristics correlated with ability, such as age or disability, incorporating this into the tax structure can increase efficiency. "Tagging" these groups increases efficiency because money is not wasted on grants to those with high ability and low income. However, those who are low ability and not included in the tagged group are excluded from the transfer.

Differential commodity taxation may also have efficiency benefits. If the tax system is restricted to a linear income tax when a non-linear structure is preferred, identifying goods consumed only by those who would in different brackets and taxing them accordingly can substitute for the non-linear income tax. If the non-linear income tax is available, it is preferred to differential commodity taxation since it distorts only the labor/leisure decision and not relative consumption decisions as well. The argument for differential commodity taxation is not an argument for a luxury tax, however. There is no efficiency gain to taxing goods that high-ability people buy simply because they have higher income. There would only be an efficiency gain to taxing goods bought by high-ability people if their purchases are based on differing, immutable preferences correlated with ability (such differences in preferences were ruled out by the assumptions in the basic model). If high-ability people like opera and low ability people like wrestling, taxing opera and subsidizing wrestling might be indicated. It may also be efficient to tax goods that are complimentary to leisure as a way to

lessen the labor/leisure distortion. For example, taxing raw food, which requires time out of market production to prepare, and subsidizing restaurant meals, then distributing the proceeds to the poor could be a Pareto improvement.