Chapter 3. Individual Income Taxation

Introduction

The individual income tax is the most important single tax in many countries. The basic principle of the individual income tax is that the taxpayer’s income from all sources should be combined into a single or global measure of income. Total income is then reduced by certain exemptions and deductions to arrive at income subject to tax. This is the base to which tax rate are applied when computing tax.

A degree and coverage of exemptions and deductions vary from country to country. A degree of progressivity of tax rates also varies.

Nevertheless the underlying principles of the tax system are common among countries and are worth reviewing.

The Income-Based Principle

Economists have argued that a comprehensive definition of income must be used that includes not only cash income but capital gains. A number of other adjustments have to be made to convert your “cash” income into the “comprehensive” income that, in principle, should form the basis of taxation.

This comprehensive definition of income is referred to as the Hicksian concept or the Haig-Simons concept. This concept measures most accurately reflects “ability to pay”.

1. Cash basis: In practice, only cash-basis market transactions are taxed. The tax is thus levied on a notion of income that is somewhat narrower than that which most economists would argue. Certain non-marketed (non-cash) economic activities are excluded, though identical activities in the market are subject to taxation (e.g. housewife’s work at home (vis-à-vis a maid’s work), own house (vis-à-vis rented house)).

Some non-cash transactions are listed in the tax code but are difficult to enforce.

Barter arrangements are subject to tax.

Unrealized capital gains is also not included in the income tax bases. Capital gains are taxed only when the asset is sold (not on an accrual basis).

2. Equity-based adjustments: Individuals who have large medical expenses or casualty losses are allowed to deduct a portion of those expenses from their income, presumably on the grounds that they are not in as good a position for paying taxes as someone with the same income without those expenses.
(3) **Incentive-based adjustments:** The tax code is used to encourage certain activities by allowing tax credits or deductions for those expenditures. Incentives are provided for energy conservation, for investment, and for charitable contributions.

(4) **Special Treatment of Capital Income:** The tax laws treat capital and wage income differently. The difficulty of assessing the magnitude of the returns to capital plays some role, while attempts to encourage savings as a source of domestic investment and growth.

**The Progressivity Principle**

Even the simplification of tax schedule prevails among countries, the premise remains that those with higher incomes not only should pay more but should pay a larger fraction of their income in taxes. In other words, progressivity is reflected in an increase not only in average rates but in marginal rates.

Defining progressive and regressive is not easy and, unfortunately, ambiguities in definition sometimes confuse public debate. A natural way to define these words is in terms of the **average tax rate**, the ratio of taxes paid to income. If the average tax rate increases with income, the system is **progressive**; if it falls, the tax is **regressive**.

Confusion arises because some people think of progressiveness in terms of the **marginal tax rate** – the change in taxes paid with respect to a change in income. To illustrate the distinction, consider the following very simple income tax structure. Each individual computes her tax bill by subtracting $3,000 from income and paying an amount equal to 20 percent of the remainder. (If the difference is negative, the individual gets a subsidy equal to 20 percent of the figure.) Table 3.1 shows the amount of tax paid, the average tax rate, and the marginal tax rate for each of several income levels. The average rates increase with income. However, the marginal tax rate is constant at 0.2 because for each additional dollar earned, the individual pays an additional 20 cents, regardless of income level. People could disagree about the progressiveness of this tax system and each be right according to their own definitions. It is therefore very important to make the definition clear when using the terms **regressive** and **progressive**. In the remainder of this book, we assume they are defined in terms of average tax rates.

Measuring how progressive a tax system is presents an even harder task than defining progressiveness. Many reasonable alternatives have been proposed, and we consider two simple ones. The first says that the greater the increase in average tax rates as income

---

1. This Part draws heavily from Rosen (1999), pp.258-260.
increases, the more progressive the system. Algebraically, let \( T_0 \) and \( T_1 \) be the tax liabilities at income levels \( I_0 \) and \( I_1 \), respectively (\( I_1 \) is greater than \( I_0 \)). The measurement of progressiveness, \( v_1 \), is

\[
v_1 = \frac{T_1 - T_0}{I_1 - I_0}
\]

(1)

Table 1  Tax Liabilities under a Hypothetical Tax System

<table>
<thead>
<tr>
<th>Income</th>
<th>Tax Liability</th>
<th>Average Tax Rate</th>
<th>Marginal Tax Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,000</td>
<td>-200</td>
<td>-0.10</td>
<td>0.2</td>
</tr>
<tr>
<td>3,000</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
</tr>
<tr>
<td>5,000</td>
<td>400</td>
<td>0.08</td>
<td>0.2</td>
</tr>
<tr>
<td>10,000</td>
<td>1,400</td>
<td>0.14</td>
<td>0.2</td>
</tr>
<tr>
<td>30,000</td>
<td>5,400</td>
<td>0.18</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Once the analyst has found the economic incidence of the tax as embodied in \( T_1 \) and \( T_0 \), the tax system with the higher value of \( v_1 \) is said to be more progressive.

The second possibility is to say that one tax system is more progressive than another if its elasticity of tax revenues with respect to income (i.e., the percentage change in tax revenues divided by percentage change in income) is higher. Here the expression to be evaluated is \( v_2 \), defined as

\[
v_2 = \frac{(T_1 - T_0)}{T_0} \times \frac{(I_1 - I_0)}{I_0}
\]

(2)

Now consider the following proposal: everyone’s tax liability is to be increased by 20 percent of the amount of tax he or she currently pays. This proposal would increase the tax liability of a person who formerly paid \( T_0 \) to \( 1.2 \times T_0 \), and the liability that was formerly \( T_1 \) to \( 1.2 \times T_1 \). Member of Congress A says the proposal will make the tax system more progressive, while member of Congress B says it has no effect on progressiveness whatsoever. Who is right? It depends on the progressivity measure. Substituting the expressions \( 1.2 \times T_0 \) and \( 1.2 \times T_1 \) for \( T_0 \) and \( T_1 \), respectively, in Equation (1), \( v_1 \) increases by 20 percent. The proposal thus increases progressiveness. On the other hand, if the same substitution is done in Equation (3.2), the value of \( v_2 \) is unchanged. (Both the numerator and denominator are multiplied by 1.2, which cancels out the effect.) The lesson here is that even very intuitively
appealing measures of progressiveness can give different answers\(^3\). Again, intelligent public debate requires that people make their definitions clear.

**The Individual or Family-Based Principle: Choice of Unit\(^4\)**

The basic unit of taxation can be either the individual or the family. Many countries adopt the individual-based principle while some adopt the family-based principle.

**Background**

To begin, it is useful to consider the following three principles:

1. The income tax should embody increasing marginal tax rates.
2. Families with equal incomes should, other things being the same, pay equal taxes.
3. Two individuals’ tax burdens should not change when they marry; the tax system should be *marriage neutral*.

<table>
<thead>
<tr>
<th></th>
<th>Individual Income</th>
<th>Individual Tax</th>
<th>Family Tax with Individual Filing</th>
<th>Joint Income</th>
<th>Joint Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lucy</td>
<td>1,000</td>
<td>100</td>
<td></td>
<td>12,200</td>
<td>30,000</td>
</tr>
<tr>
<td>Ricky</td>
<td>29,000</td>
<td>12,100</td>
<td></td>
<td>12,600</td>
<td>30,000</td>
</tr>
<tr>
<td>Ethel</td>
<td>15,000</td>
<td>5,100</td>
<td></td>
<td>10,200</td>
<td>30,000</td>
</tr>
<tr>
<td>Fred</td>
<td>15,000</td>
<td>5,100</td>
<td></td>
<td>12,600</td>
<td>30,000</td>
</tr>
</tbody>
</table>

Although a certain amount of controversy surrounds the second and third principles, it is probably fair to say they reflect a broad consensus as to desirable features of a tax system. While agreement on the first principle is weaker, increasing marginal tax rates seem to have wide political support.

Despite the appeal of these principles, a problem arises when it comes to implementing them: In general, no tax system can adhere to all three simultaneously. This point is made most easily with an arithmetic example. Consider the following simple progressive tax schedule: a taxable unit pays in tax 10 percent of all income up to $6,000, and 50 percent of all income in excess of $6,000. The first two columns of Table 2 show the incomes and tax liabilities of four individuals, Lucy, Ricky, Fred, and Ethel. (For example, Ricky’s tax liability is $12,100 [ .10 x

\(^3\) Note also that \(v_1\) and \(v_2\), in general, depend on the level of income. That is, even a single tax system does not usually have a constant \(v_1\) and \(v_2\). This further complicates discussions of the degree of progressiveness.

\(^4\) This Part draws heavily on Rosen (1999), Chap.16, pp.363-5.
$6,000 + .50 \times $23,000). \) Now assume that romances develop – Lucy marries Ricky, and Ethel marries Fred. In the absence of joint filing, the tax liability of each individual is unchanged. However, two families with the same income ($30,000) will be paying different amounts of tax. (The Lucy-Rickys pay $12,200 while the Ethel-Freds pay $10,200, as noted in the third column.) Suppose instead that the law views the family as the taxable unit, so that the tax schedule applies to joint income. In this case, the two families pay equal amounts of tax, but now tax burdens have been changed by marriage. Of course, the actual change in the tax burden depends on the difference between the tax schedules applied to individual and joint returns. This example has assumed for simplicity that the schedule remains unchanged. But it does make the main point: given increasing marginal tax rates, we cannot have both principles 2 and 3.

What choice has the United States made? Over time, the choice has changed. Before 1948, the taxable unit was the individual, and principle 2 was violated. In 1948, the family became the taxable unit, and simultaneously income splitting was introduced. Under income splitting, a family with an income of $50,000 is taxed as if it were two individuals with incomes of $25,000. Clearly, with increasing marginal tax rates, this can be a major advantage. Note also that under such a regime, and unmarried person with a given income finds his or her tax liability reduced substantially if he or she marries a person with little or no income. Indeed, under the 1948 law, it was possible for an individual’s tax liability to fall drastically when the person married – a violation of principle 3.

The differential between a single person’s tax liability and that of a married couple with the same income was so large that Congress created a new schedule for unmarried people in 1969. Under this schedule, a single person’s tax liability could never be more than 20 percent higher than the tax liability of a married couple with the same taxable income. (Under the old regime, differentials of up to 40 percent were possible.)

Unfortunately, this decrease in the single/married differential was purchased at the price of a violation of principle 3 in the opposite direction: it was now possible for person’ tax liabilities to increase when they married. In effect, the personal income tax levied a tax on marriage. In 1981, congress attempted to reduce the “marriage tax” by introducing a new deduction for two-earner married couples. Two-earner families received a deduction equal to 10 percent of the lower earning spouse’s wage income, but no more than $3,000. However, the two-earner deduction was eliminated by TRA86. It was deemed to be unnecessary because lower marginal tax rates reduced the importance of the “marriage tax.”

Nevertheless, a substantial penalty still exists, and it tends to be highest when both spouses have similar earnings. Under certain conditions, for example, when two individuals with $25,000 AGIs marry, their joint tax liability can increase by more than $700. On the other
hand, when there are considerable differences in individuals’ earnings, the tax code provides a bonus for marriage. If two people with $10,000 and $50,000 AGIs marry, their joint tax liability can decrease by $1,100. In cases like these, the law provides a “tax dowry.”

**The Annual Measure of Income Principle**

Income tax is usually based on *annual income*, not *lifetime income*. Economic theory is usually based on the lifetime utility maximization (e.g. life-cycle hypothesis), with the current annual income taxation, consumption smoothing does not avoid tax distortion. Because of the progressive nature of our tax system, the individual with the fluctuating income has to pay more taxes over his lifetime than the individual with a steady income.

**The Basic Framework**

(1) **The Criteria for Optimality**

Musgrave in his classic text book (*The Theory of Public Finance* (1959), McGraw-Hill) provides three criteria for appropriate taxation;

(a) the need for taxes to be *fair*;
(b) the need to *minimize administrative costs*; and
(c) the need to *minimize disincentive effects*.

The approach of the optimal taxation literature is to use economic analysis to combine the criteria into one, implicitly deriving the relative weights that should be applied to each criterion. This is done by using the concepts of individual utility and social welfare.

Economists found it very difficult to model the relationship between tax rates and administrative costs. They usually ignored administrative costs in their analysis and concentrated on criteria (a) and (c). Effectively, they tried to determine the tax system that will provide the best compromise between *equality* (fairness) and *efficiency* (incentives).

(2) **The Specification of Social Welfare**

As we discussed before, social welfare function can take many forms as policy makers have different policy objectives and welfare criteria.

If our idea of a fair tax system is one that reduces inequality of utility, our social welfare function must place more weight on utility gains of poor people than those of rich people. This is achieved by using the following formulae,
\[ w = \frac{1}{1 - \varepsilon} \sum_h \left( u^h \right)^{1 - \varepsilon} \quad \text{for } \varepsilon \neq 1 \]
\[ = \sum_h \log(u^h) \quad \text{for } \varepsilon = 1 \]  

(3) The Modeling of Disincentives

In case of optimal income taxation in a model where labor supply response is the only disincentive problem, the utility function for each individual is used both to predict how that person will alter his/her labor supply when taxes are changed and to evaluate the resulting level of individual utility. The changes in labor supply will then be used to calculate the change in tax revenue, while the changes in utility will be used to calculate the change in social welfare. The optimal tax system will be the one where it is impossible to increase social welfare without reducing overall tax revenue.

The requirement to raise a specific amount of tax revenue is obviously fundamental. It has two important implications. First, it means that the solution to the optimal tax problem depends on the size of the revenue requirement. Second, it means that the tax changes that are considered should be revenue-neutral.

Why does it matter that a higher tax rate with higher personal allowances will reduce labor supply? After all, the objective is to maximize social welfare, not the size of the national income. The answer is that, by choosing to work less on average, workers will have lower incomes and thus will pay less taxes. Thus a change that would have been revenue-neutral for a fixed level of labor supply will, as a result of the reduction in work, produce a revenue loss. It is this revenue loss that represents the ‘excess burden’ of taxation. It requires an increase in tax rates to offset it – an increase that will reduce social welfare and counteract, at least in part, the gain is social welfare from the reduction in inequality that is produced by the increase in tax progressivity.

(4) Problems of Application

The usefulness of the optimal tax results depends on the realism of the economic models. This is not to say that the presence of any unrealistic assumption invalidates the results. Rather, any practical application of theoretical analysis requires an evaluation of whether any violation of the assumption can be expected to alter the results significantly.

Examples of unrealistic assumptions:

(a) neglect of administrative and compliance costs of tax collection.

(b) Assumption of perfect competition

(c) neglect of heterogeneity of households in terms of their composition and preferences

What, then, the modern theory of income taxation ought to be concerned?
1. It must capture the efficiency/equity trade off involved in income taxation.

2. The structure of the income tax must be compatible with the revelation of the ability of households.

In a simpler term, the fundamental policy issue is whether it would be a good idea to increase the rate of income tax and use the proceeds to fund an increase in tax allowances, thus reducing after-tax income inequality.

**Optimal Income taxation**

A linear income tax schedule is one possibility for choosing an income tax. More generally, however, the structure of personal income taxation need not be linear with a constant rate of taxation, but rather marginal tax rates can vary with different levels of income. This is the case in practice. Usually marginal tax rates increase with income.

Suppose that we set out to determine the general optimal income tax structure that maximizes a social welfare function. We are then looking for a relationship between the tax rate and earned income that could in principle be progressive or regressive. If the rate of taxation were to change with every change in income, we would be looking for an optimal income schedule

\[ t = f(Y) \]  \hspace{1cm} (4)

If a linear income tax schedule were to maximize social welfare, this would be revealed as the solution to the general problem of optimal personal income taxation.

We can expect the quest to identify a general structure of optimal income taxation as expressed by (4) to be complicated. In the case of the linear income tax, we had to find values for \( S \) and \( t \) related through the government budget. In the case of the tax schedule (4), we are looking for a relationship between the tax rate and income that maximizes social welfare. That is, the solution to the general optimal taxation problem is a function that tells us how to set the tax rate \( t \) for all values of income \( Y \). We have seen that there are cases for both progressive and regressive taxes, and we therefore do not know beforehand whether the relationship expressed in (4) will indicate progressive or regressive taxation. We do know that finding an optimal income tax schedule requires a compromise between the progressive taxation sought for reasons of social justice (through the equal-sacrifice principle) and the efficiency and tax-base benefits of regressive taxation.

Through the substitution effect between work (and effort) and leisure, progressive taxation increases the leaks in the redistributive bucket. We can return to our example of the dentist
who responds to increasingly greater marginal tax rates by stopping work and heading for the
golf course. The substitution effect places an excess burden of taxation on the dentist, but also we saw that the dentist is led by progressive taxes to take personal utility in the form of leisure rather in the form of earned income. Because leisure cannot be taxed, the beneficiaries of the income transfers financed by the dentist’s taxes have reason to want the dentist to keep working and earning taxable income. In deciding on a tax schedule, an important question is therefore, how do taxpayers respond in their work and leisure decisions to the degree of progressivity or regressivity in the income tax schedule? The answer to this question determines the efficiency losses (through the excess burden of taxation) that are required to be incurred for the sake of social justice defined as a more equal post-tax income distribution. The answer tells us how far inside the efficient frontier a society has to go to approach greater post-tax income equality.

In a choosing a social welfare function, a society can stress efficiency or social justice (expressed as a preference for post-tax equality). We have seen that a society’s choice of social welfare function correspondingly expresses the society’s aversion to risk and determines whether social insurance is complete (with Rawls) or incomplete (with Bentham and other formulations of social welfare).

The extent of inefficiency, or the leak in the bucket of redistribution through the response of taxpayers to progressivity or regressivity in the income tax schedule, is an empirical matter. We need to be able to observe labor-supply behavior to determine how people respond to taxes. The choice of the social welfare function to be maximized is an ideological issue. Some economists and political decision makers stress the desirability of social justice with little concern for efficiency (they are followers of Rawls) and want highly progressive income taxes. Others (who are closer to Bentham) stress the desirability of efficiency and want low marginal income tax rates or flat tax rates.

Although labor-supply behavior is empirically determined, different people often have different views or priorities about how labor-supply decisions respond to taxes. For economists and political decision makers who take the view that people more or less “contribute according to their ability”, work and effort substitution responses to taxes are low, and efficiency losses through excess burdens of taxation are not a deterrent to highly progressive income taxes. Such economists and political decision makers might then see their way free to choose a social welfare function close to Rawls, with resulting high tax rates and high progressivity in the tax schedule. Economists and political decision makers who interpret the evidence as an indication that incentives to work and exert effort are important stress the efficiency losses from taxation and recommend income tax structures with low tax rates and low levels of progressivity. In particular, the latter group of economists and political decision

---

5 This part quotes Hillman (2003, Chap.7, pp.479-83.)
makers often recommends a linear income tax schedule, or a schedule with a small number of tax brackets with low rates of taxation and low progressivity.

**The Mirrlees Problem**

In formal economic analysis, economists derive an optimal income tax schedule by maximizing a social welfare function after choosing a utility function that represents the preferences of the population for work (or consumption) and leisure. Individuals are generally viewed as having the same preferences or utility functions regarding work and leisure but as having different abilities to earn income. The income tax affects how individuals apply their different abilities to productive effort, and so determines the value of the output that people produce.

James A. Mirrlees (of the University of Cambridge, who received the Nobel Prize in economics in 1996) pointed out (1971) that, although a government might want to maximize a social welfare function by taxing individuals’ abilities to earn income, individuals’ abilities are not observable\(^6\). Governments therefore confront a problem of asymmetric information when seeking to tax ability: the information is asymmetric because individuals know their personal abilities, but governments do not. A government can therefore only observe personal incomes earned through the application of personal ability and can only tax personal welfare when all people have the same utility function but differ in their unobserved abilities to earn income\(^7\). The government in Mirrlees’s optimal income tax problem also could not observe how many hours a person worked or how much effort a person put into work. The government did, however, know the distribution of abilities in the population. The question posed by Mirrlees is therefore: if a government were to know the common utility function of the population and the distribution of taxpayers’ abilities but could only observe the incomes that people earn yet wished to maximize social welfare, how progressive or regressive a personal income tax would the government wish to choose? When this question is answered, much hinges on the choice of the utility function proposed to describe the labor supply behavior of the population. The trade-off between efficiency and social justice depends on the efficiency losses through the excess burden of taxation, which through the substitution effect depends on the utility function that is chosen to represent the labor-market behavior of the population. A utility function could be chosen according to which free time or leisure is little valued. Taxes then do not

---

\(^6\) If individuals’ abilities were innate (and are not acquired or enhanced through personal decisions to invest in education), the taxation of ability would provide a tax base with no efficiency loss through an excess burden of taxation because, as we have noted, there are no substitution responses available to individuals to avoid a tax on innate ability. We have also observed that a tax on ability would be unfair, by taxing a personal life-style decision not to use ability to earn market income.

\(^7\) An additional incentive-compatibility constraint ensures that individual decisions to work maximize personal utility
much affect efficiency, and the optimal income tax can focus on achieving equality in post-tax income distribution and can be quite progressive. If a utility function is chosen for which people do significantly value free time or leisure and respond to high marginal taxes by significantly reducing work hours or work effort, the optimal income tax will not be all that progressive and may be regressive. Mirrlees found this ambiguity: no generally valid answer could be given to the question whether income taxes should be progressive or regressive at different levels of income.

Mirrlees’s optimal income tax solution does however have some definitive attributes. The income tax should provide incentives for higher-ability people to work at least as much as lower-ability people. We can see this by considering the behavior of one person with high ability and another with lower ability. The higher-ability person can always copy the work behavior of the lower-ability person, thereby earning the same pre-tax income and paying the same taxes as the lower-ability person. The higher-ability person who is behaving in a way that makes him or her seem to be a lower-ability person earns the same (observable) income as the lower-ability person but works fewer (non-observable) hours or exerts less (non-observable) effort than the lower-ability person. To avoid this outcome, and to induce higher-ability people to exert more effort, the optimal income tax should provide the higher-ability person with a higher level of post-tax utility than a person who has less ability. That is, it is efficient to give the higher-ability person an incentive through the tax structure to earn more income than a lower-ability person. The tax structure should not therefore discourage more able people from working more than less able people. That is, the optimal income structure should provide an incentive for high-ability people to reveal their superior ability through the greater pre-tax incomes they earn.

Simulations

The inconclusiveness of the solution to the general problem formulated by Mirrlees had led to emphasis on attempts to identify the ideal structure of an income tax from simulations of a model of an economy. As we expect, simulations that presuppose high substitution between free time and income-earning activity show that society is better off with income tax structures that have low income tax rates and low progressivity. The simulations also sometimes reveal

---

8 The analysis of optimal income taxation has also provided conclusions about marginal tax rates at the bottom and top of the income distribution. A zero marginal tax rate at the highest income is optimal because this does not deter the highest-income person from working more. If this person is deterred from working because of a positive marginal tax rate, the income tax structure cannot be efficient because, when the highest-ability person works more (by free choice), he or she is better off. At the same time, not tax revenue is lost to the government by the decision to work more. If everybody is working, the marginal tax rate is also zero for the lowest-ability person in the population (but, if some people are not working, the marginal tax rate at the bottom can be positive).
that a flat tax (a constant marginal tax rate as in the linear income tax) is the preferred way of taxing income because of the minimal disincentives for productive activity.

When simulations are based on low substitution between free time and hours worked or work effort, the simulation results, on the other hand, show that high progressivity in income taxes does not significantly impede work effort, and that social welfare is maximized by marginal tax rates that can be quite high at higher incomes\(^9\). Mirrlees’s own simulations suggested that the ideal income tax structure is not very progressive and that the highest marginal rate should not exceed 30 to 35 percent.

The different results and views leave us to make a personal choice about how we view the relation between taxes and incentives to earn income and about the importance we assign to efficiency compared to social equality when choosing an income tax. We may believe that personal incentives are important in determining how hard people work and the time they put into work. In that case, we would not accept that highly progressive income tax rates and high rates of taxation are in the best interests of a society. An emphasis on efficiency and the effects of taxation on the tax base would lead us to want to choose low and moderately progressive personal income taxes or the linear income tax with an amendment that adds one or two low-tax brackets.

**Social mobility and the structure of the income tax**

A society with high social mobility places no impediments to people born into low-income families progressing to the upper end of the income distribution. Income tax structures with high tax rates and high progressivity obstruct social mobility by impeding the accumulation of wealth of people born into poor families. The traditionally wealthy families pay the same high and progressive income taxes, but these families have income from their endowed personal wealth. High tax rates and high progressivity in the income tax structure therefore favor preservation of the status of the traditional wealthy families. That is, high tax rates and high progressivity in income taxes preserve social inequality over time by stopping low-income people from becoming rich because of the high taxes that are paid and the disincentives to exert the effort that is required to achieve upward social mobility. The desire not to impede social mobility is another reason for moderate taxation in an income tax schedule.

---

\(^9\) For example, Emmanuel Saez (2001) suggested top marginal tax rates between 50 and 80 percent.
The Mirrleesian Economy

The basic features of the Mirrleesian optimal income tax model can be described as follows. Imagine an economy where individuals have the innate ability to transform working-time into a single consumption good, which is called income. Each individual’s utility is a numerical function which depends only on his net income and the quantity of labor he supplies.

Thus his preferences are personal. Individuals are regarded as identical except that they vary in their ability to supply labor. They can be grouped by productivity types. For the same number of hours worked, a more able person naturally can produce more income. Each individual decides how much labor to supply, calculating what will maximize his utility.

All there labor supply decisions taken together determine the output of the economy. A redistributive system is feasible if the output of the economy. A redistributive system is feasible if the output of the economy is sufficient to provide for public and private expenditures. Furthermore, it is supposed that the government or the tax authority does not know the productivity type of each individual. It cannot monitor the number of working hours a person chooses to work, but can only observe a person’s income. For this reason the only policy the government can execute is to impose a tax schedule. The government chooses the income tax schedule which maximizes its social welfare function, knowing the manner in which individual of any productivity type will respond.

There are four main elements in Mirrlees’s formulation:

The social welfare function, the individual supply or behavioral response function, the distribution of ability, and the production structure.

The social welfare function embodies society’s values about issues of equity and fairness. Investigations have been carried out into how alternative social welfare functions affect the results, focusing on the impact of giving different social weights to the utilities of the rich and the poor. More egalitarian preferences are represented by assigning the utilities of poor people more weight than the utilities of the rich. An extreme case is the utility-based version of the Rawlsian difference principle – maximum – which attaches no weight at all to anybody’s well-being apart from that of the worst-off individual.

The most critical element of the income tax problem is that the taxpayer can alter his behavior according to the tax system he faces. In particular, he chooses how much labor to supply. The labor supply function provides essential information about the costs of a redistribution tax system.

The skill of an individual is what determines his ability to produce income. Since factors affecting the ability to earn income (e.g. inheritance, intelligence, motivation, luck) are not equally distributed, the pre-tax distribution of income is usually far more equal. In the
standard model it is assumed that the labor services supplied by different individuals differ in productivity according to their different abilities. A common “efficiency” unit is therefore used, and it is assumed that an individual with ability \( n \) works \( y \) hours and supplies labor services at the amount of \( my \), as measured in efficiency units. His gross income is then \( z = wy \) where \( w \) is the real wage rate per efficiency unit (in terms of the consumption good).

The most important features of the tax design problem.

1. There must be an unequal distribution of income in the before tax equilibrium of the economy. Otherwise there is no point introducing equity motivation for taxation.
2. The income distribution must be generated endogenously by the model, with households differing in the income they earn, and the income tax must affect the labor supply decisions of the households in order to introduce efficiency considerations.
3. The economy must be sufficiently flexible that no prior restrictions are placed on the tax functions that may be solutions.

Assumptions on the economy

1. The economy is competitive.
2. Households differ only in the level of skill in employment. A household’s level of skill determines their hourly wage and hence income.
3. The skill level is private information which is not known to the government.
4. The only tax instrument of the state is an income tax. An income tax is employed both because lump-sum taxes are infeasible and because it is assumed that it is not possible for the state to observe separately hours worked and income per hour. Therefore, since only total income is observed, it has to be the basis for the tax system.

The Basic Structure of the Economy: the Model

1. Two commodities: a consumption good \( x (x \geq 0) \) and a single labor service, \( l \quad (0 \leq l \leq 1) \).
2. Each household is characterized by their skill level, \( s \). The value of \( s \) gives the relative effectiveness of the labor supplied per unit of time. If a household of ability \( s \) supplies \( l \) hours of labor, they provide a quantity \( sl \) of effective labor. For simplicity, the marginal product of labor is equal to a worker’s ability \( s \). The total productivity of a worker during the \( l \) hours at work is equal to \( sl \).
3. Denote the supply of effective labor of a household with ability \( s \) by \( z(s) = sl(s) \).
4. The price of the consumption good is normalized at 1.
5. $z(s)$ is the household’s pre-tax income in units of consumption. Denoting the tax function by $T(z)$ and the consumption function by $c(z)$, a household that earns $z(s)$ units of income can consume

$$x(s) \leq c(z(s)) = z(s) - T(z(s)) \quad (4)$$

6. The ability parameter $s$ is continuously distributed throughout the population with support $s$ ($s$ can be finite with $s=[s_1,s_2]$ or infinite with $s=[0,\infty]$). The cumulative distribution of $s$ is given by $\Gamma(s)$, so there are $\Gamma(s)$ households with ability $s$. The corresponding density function is denoted $\gamma(s)$.

**Figure 1**

7. All households have the same strictly concave utility function

$$U = U(x, l) \quad (5)$$

Each household makes the choice of labor supply and consumption demand to maximize utility subject to the budget constraint.

$$\text{Max } U(x, l) \text{ subject to } x(s) \leq c(sl(s)) \quad (6)$$

In the absence of income taxation, a household of ability $s$ would face the budget constraint

$$x \leq sl \quad (7)$$

From (7), it is obvious that the budget constraint differs with ability.

For simplicity, all households face the same budget constraint. This can be achieved by setting the analysis in $(z,x)$ space. In this space, the pre-tax budget constraint is given by the 45° line for households of all abilities.

---

10 This part draws heavily from Myles (1955) Chap 2, pp.133-55.
Figure 2 presents one case for the general shape of the pattern of pre-tax and post-tax income, $z$ and $T(z)$ respectively. The curve $T(z)$ describes the case where marginal tax rates defined by $1 - T'(z)$ are first high for lower income, then rather low for middle-range income, and finally high again for high income. The shape of $T(z)$ is very close to the marginal tax rate structure commonly observed in many countries.

$T(z)$ is sometimes referred to as a tax structure which sets the floor for poverty and the ceiling for wealth. At the same time it is very favorable for middle income groups. It can, however, be argued that when the possible disincentive effects of high marginal tax rates are taken into account, the pattern described above is no longer desirable. The problem is how to seek a compromise between the disincentive effects of marginal tax rates and their effects in achieving a more equal distribution of economic welfare.

Let $c(z(s))$ denote consumption, and $l(s)$ hours worked by an individual whose ability (productivity) is $s$ and whose gross income is $z = sl(s)$. Let $c(z(s))$ be the consumption schedule imposed by the government, with respect to which each individual must make his or her choice. Further, it is assumed that the indifference curves between $c$ and $l(s)$ for all individuals are the same. By expanding each $(c,l)$ curve horizontally by the factor $s$, indifference curves between $c$ and $sl$ can be drawn. If $c$ is non-inferior, these curves are strictly flatter at each point the greater the value of $s$. This means that an individual with a greater $s$ is more able to substitute labor for consumption. Under this assumption it is clear that $c$ and $z$ must be increasing functions of $s$.

**Figure 2**
The Structure of Utility

The households have identical preferences over consumption and leisure. The utility function is continuously differentiable, strictly increasing in consumption and strictly decreasing in leisure. In addition, it satisfies

\[ U_x > 0, \quad U_l < 0, \quad U_{xx} < 0 \]  

(8)

and

\[ U_l (x,l) \rightarrow -\infty \text{ as } l \rightarrow 1 \]  

(9)

(5) implies that each household will endeavour to avoid corner solutions with \( l=1 \) (no one wants to work all day long!!). The indifference curves of the utility function are illustrated bellows in which utility increases to the north west.

Figure 3

To allow preferences and the budget constraint to be depicted on the same diagram, the utility function can be written

\[ U = U(x,l) = U(x, z/s) = u(x,z,s) \]  

(10)
The indifference curves of \( u(x,z,s) \), drawn \((z,x)\)-space are dependent upon the ability level of the household since it takes a high-ability household less labor time to achieve any given level of income.

In fact, the indifference curves are constructed from those in \((l,x)\)-space by multiplying by the relevant value of \( s \). This construction for the single indifference curve \( I_0 \) and households of three different ability levels.

**Figure 4**

Agent Monotonicity

The utility function (7) satisfies agent monotonicity if \(-\frac{u_x}{u_z} \), is a decreasing function of \( s \).

Note that \( \Phi \equiv -\frac{u_x}{u_z} \) is the marginal rate of substitution between consumption and pre-tax income and that agent monotonicity requires \( \Phi_s \equiv \frac{\partial \Phi}{\partial s} < 0 \).

An equivalent definition of agent monotonicity is that \(-\frac{U_l}{U_x} \) is an increasing function of \( l \) as \(-\frac{u_z}{u_x} = -\frac{U_l(x,z/s)}{sU_x(x,z/s)} \). Calculating \( \partial \left[ -\frac{U_l}{U_x} \right] / \partial l \) and \( \Phi_s \) shows

\[
\Phi_s = -\frac{1}{s^2} \partial \left[ \frac{U_l}{U_x} \right] / \partial l
\]

Agent monotonicity is equivalent to the condition that, in the absence of taxation,
consumption will increase as the wage rate increases. A sufficient condition for agent
monotonicity is that consumption is not inferior, i.e. it does not decrease as lump-sum income
increases.

The marginal rate of substitution is the gradient of the indifference curve, agent monotonicity
implies that at any point in \((z,x)\)-space the indifference curve of a household of ability \(s^1\) passing
through that point is steeper than the curve of a household of ability \(s^2\) if \(s^2 > s^1\). Agent
monotonicity implies that any two indifference curves of households of different abilities only
cross once. In other words, the indifference curve of an \(s\)-ability individual through the point
\((x,z)\) in consumption-labor space rotates strictly clockwise as \(s\) increases.

**Figure 5**

![Figure 5](image)

Mirrlees proved a theorem which shows, when the consumption function is a differentiable
function of labor supply, agent monotonicity implies that gross income is an increasing function
of ability (in other words, if agent monotonicity holds and the implemented tax function has
pre-tax income increasing with ability, then the second-order condition for utility maximization
must hold!). This is important as to identify one’s ability by watching gross income.

**Self-selection**

Let \(x(s)\) and \(z(s)\) represent the consumption and income levels that the government intends a
household of ability \(s\) to choose. The household of ability \(s\) will choose \((x(s), z(s))\) provided
that this pair generates at least as much utility as any other choice. This condition must apply
to all consumption-income pairs and to all households. Formally we can write,

\[
\text{The self-selection constraint is satisfied if } \ u(x(s), z(s), s) \geq u(x(s'), z(s'), s') \text{ for all } s \text{ and } s'.
\]
In case of linear taxation, it does not need to consider the self-selection constraints since the behavior of the household can be determined as a function of the two parameters that describe the tax function; the lump-sum payment and the marginal rate of tax.

In case of non-linear taxation, the self-selection constraints must be included. This is achieved by noting that the satisfaction of the self-selection constraint is equivalent to achieving the minimum of a certain minimization problem. If the sufficient conditions for the minimization are satisfied by the allocation resulting from the tax function, then the self-selection constraint is satisfied. The idea is to induce the more able group to ‘reveal’ that they have a high income, not the reverse.

To derive the required minimization problem, let \( u(s) = u(x(s), z(s), s) \) represent the maximized level of utility for a consumer of ability \( s \) resulting from (6).

\[
0 = u(s) - u(x(s), z(s), s) \leq u(s') - u(x(s), z(s), s') \tag{12}
\]

so that \( s' = s \) minimizes \( u(s') - u(x(s), z(s), s') \). Hence

\[
u'(s) = u_s(x(s), z(s), s). \tag{13}
\]

From the definition of \( u(s) \) it follows that

\[
u_x'(s) + u_z'(s) = 0 \tag{14}
\]

is equivalent to (13).

Condition (13) or (14) is the necessary (the first order) condition for the self-selection constraint to be satisfied.

The second-order condition for the self-selection constraint is found from the second derivative of \( u(s') - u(x, z, s') \) with respect to \( s' \) to be

\[
u''(s) - u_{ss}(x(s), z(s), s) \geq 0 \tag{15}
\]

Again using the definition of \( u(s) \),
\begin{align}
  u^{**}(s) - u_{ss}x'(s) + u_{sz}z'(s) + u_{sx}z''(s) + u_{sz}z''(s)
\end{align} \quad (16)

which gives, by using (15).

\begin{align}
  u_{ss}x'(s) + u_{sz}z'(s) & \geq 0
\end{align} \quad (17)

Eliminating \( x'(s) \) using (14) provides the final condition

\begin{align}
  \left[ u_{sz} - u_{ss} \frac{u_{sz}}{u_{sx}} \right] z'(s) = -\frac{\Phi_s}{u_{sx}} z''(s) \geq 0
\end{align} \quad (18)

where \( \Phi_s \) is the marginal rate of substitution introduced in the discussion of agent monotonicity. With agent monotonicity \( \Phi_s < 0 \), so that satisfaction of the second-order condition for self-selection is equivalent to \( z'(s) \geq 0 \). Any tax function that leads to an outcome satisfying (14) and \( z'(s) \geq 0 \) will therefore satisfy the self-selection constraint.

**Characterization of Optimal Tax Function**

It will clearly not be possible to calculate the function without precisely stating the functional forms of utility, production and skill distribution. What will be achieved is the derivation of a set of restrictions that the optimal function must satisfy.

**The General Problem**

Using the individual demand and supply functions and integrating over the population, it is possible to define total effective labor supply \( Z \), by

\begin{align}
  Z &= \int_0^\infty z(s)\gamma(s)ds
\end{align} \quad (19)

and aggregate demand, \( X \), where

\begin{align}
  X &= \int_0^\infty x(s)\gamma(s)ds
\end{align} \quad (20)
The optimal tax function is then chosen to maximize social welfare, where social welfare is given by the Bergson-Samuelson function.

\[ W = \int_{0}^{\infty} w(u(s)) y'(s) ds \]  

(21)

with \( W'' \leq 0 \).

There are two constraints upon the maximization of (21). The first is that the chosen allocation must be productively feasible such that,

\[ X \leq F(Z) \]  

(22)

where \( F \) is the production function for the economy.

This definition of productive feasibility can incorporate the government revenue requirement, expressed as a quantity of labor consumed by the government \( Z^G \), by noting that (22) can be written \( X \leq \hat{F}(Z - Z^G) = F(Z) \).

Denoting the level of revenue required by \( R(\hat{Z}^G) \), the revenue constraint can be written

\[ R \geq \int_{0}^{\infty} [z(s) - x(s)] y'(s) ds \]  

(23)

The second constraint is that it must satisfy the self-selection constraint which has already been discussed.

**Linear taxation**

With linear taxation the marginal rate of tax is constant and there is an identical lump-sum tax or subsidy for all households.

The advantages of this restriction is that it ensures that the budget sets of all households are convex so that optimal choices will be unique when preferences are strictly convex. In addition, the tax system is described by just two parameters: the marginal tax rate and the lump-sum subsidy.

The linear tax structure corresponds to proposals for negative income tax schemes, in which all households below a given income level receive a subsidy from the tax system.
Under a linear tax system a household with ability \(s\) supplying \(l\) units of labor will pay tax of amount,

\[
T(sl) = -\tau + ts l
\]  

(24)

where \(t\) is the marginal rate of tax and \(\Delta\) is a lump-sum subsidy if positive and a tax if negative.

Denoting \((1-t)\) by \(\Delta\), the consumption function of the household is

\[
x = \Delta + \Delta sl.
\]  

(25)

Each household chooses consumption and labor supply to maximize utility (5) subject to (25).

The first-order conditions can be reduced to

\[
- \frac{U_i}{U_x} = \zeta s.
\]  

(26)

Labor supply and consumption demand functions can be written as,

\[
\begin{align*}
&l = l(\zeta, \tau, s) \\
x = \tau + \zeta sl(\zeta, \tau, s)
\end{align*}
\]  

(27)

Substituting (27) into the utility function, there determine the indirect utility function,

\[
U = U(\tau + \zeta sl(\zeta, \tau, s), l(\zeta, \tau, s)) = V(\zeta, \tau, s)
\]  

(28)

with

\[
\frac{\partial V}{\partial \tau} = U_x, \quad \frac{\partial V}{\partial \zeta} = U_s l
\]  

(29)

where \(\partial V/\partial \tau\) is equal to the marginal utility of income.

The government’s optimization problem is to choose the parameters of the tax system to maximize social welfare subject to raising the required revenue, \(R\).
\[
\max_{\tau, \zeta} \int_{0}^{\infty} w(V(\zeta, \tau, s)) \gamma(s) ds
\] (30)

subject to

\[
\int_{0}^{\infty} \left[ -\tau + (1 - \zeta) s l(\zeta, \tau, s) \right] \gamma(s) ds = R
\] (31)

Using (29) and defining the social marginal utility of income for a household of ability \( s \) by

\[
\beta(s) = w'(V(\zeta, \tau, s)) \frac{\partial V(\zeta, \tau, s)}{\partial \tau}
\] (32)

The necessary conditions for the choice of \( \beta \) and \( \gamma \) respectively are

\[
\int_{0}^{\infty} \beta \gamma(s) ds = \lambda \left[ H - \int_{0}^{\infty} (1 - \zeta) \frac{\partial \zeta}{\partial \tau} \gamma(s) ds \right]
\] (33)

and

\[
\int_{0}^{\infty} \beta \gamma(s) ds = \lambda \int_{0}^{\infty} \left[ z - (1 - \zeta) \frac{\partial \zeta}{\partial \zeta} \right] \gamma(s) ds
\] (34)

where \( H \) is the population size, \( H = \int_{0}^{\infty} \gamma(s) ds \).

Divide (34) by (33) and denote by a bar term of the form \( x/H \).

\[
\frac{\int_{0}^{\infty} \beta \gamma(s) ds}{\int_{0}^{\infty} \beta \gamma(s) ds} = \frac{\int_{0}^{\infty} (1 - \zeta) \frac{\partial \zeta}{\partial \zeta} \gamma(s) ds}{1 - \int_{0}^{\infty} (1 - \zeta) \frac{\partial \zeta}{\partial \tau} \gamma(s) ds}
\] (35)

The term on the left-hand side of (35) is now denoted \( z(\beta) \) and can be interpreted as the welfare-weighted average labor supply. From totally differentiating the government revenue constraint whilst holding \( R \) constant, it can be found that
\[
\frac{\partial A}{\partial \zeta} \bigg|_{R_{\text{const}}} = -\bar{z} + \int_0^l \left(1 - \zeta\right) \frac{\partial \bar{z}}{\partial \zeta} \gamma(s) ds \\
1 - \int_0^l \left(1 - \zeta\right) \frac{\partial \bar{z}}{\partial \tau} \gamma(s) ds
\]

(36)

Hence from (35) and (36),

\[
z(\beta) = -\left. \frac{\partial A}{\partial \zeta} \right|_{R_{\text{const}}}
\]

(37)

Since averaging over the population must give \( \bar{z} = \bar{z}(z, \zeta) \), it follows from (37) that, holding revenue constant

\[
\left. \frac{\partial \bar{z}}{\partial \zeta} \right|_{R_{\text{const}}} = \left. \frac{\partial \bar{z}}{\partial \tau} \right|_{R_{\text{const}}} \partial \tau - \left. \frac{\partial \bar{z}}{\partial \tau} \right|_{R_{\text{const}}} \partial \tau
\]

(38)

Therefore (36) can be written in the form

\[
z(\beta) - \bar{z} = (1 - \zeta) \left[ \frac{\partial \bar{z}}{\partial \tau} z(\beta) - \frac{\partial \bar{z}}{\partial \tau} \right] = -t \left. \frac{\partial \bar{z}}{\partial \tau} \right|_{R_{\text{const}}}
\]

(39)

Recall that \( t = 1 - \zeta \)

\[
t = \frac{\bar{z} - z(\beta)}{-\left. \frac{\partial \bar{z}}{\partial \zeta} \right|_{R_{\text{const}}}}
\]

(40)

where the derivative is taken with revenue constant.

Although the tax rule (40) only provides an implicit expression for \( t \), it can be used to assess the effects of various parametric changes. A reduction in the optimal tax would occur, with \( t \) a decreasing function of \( s \) and \( z \) on increasing function of \( s \), if the welfare weights were increased on the high-\( s \) households so that equity was given less weight.
Non-linear taxation

With non-linear taxation, the self-selection constraint must be taken fully into account.

The optimal structure of income taxation is characterized by applying Pontryagin’s maximum principle.

The revenue function,

\[ R(s) = \int_{s_1}^{s_2} (z(s') - x(s'))\gamma(s')ds' \]  \hspace{1cm} (41)

The level of utility \( u(s) \), pre-tax income \( z(s) \) and the tax payments of households of ability \( s \) are taken as the state variables and the derivative of gross income, \( \eta(s) \equiv \frac{\partial z}{\partial s} \) is taken as the control variable. The level of consumption can then be found by solving \( u(s) = u(x(s), z(s), s) \).

Adopting a utilitarian objective, the control variable is chosen to maximize

\[ \int_{s_1}^{s_2} u(s)\gamma(s)ds \]  \hspace{1cm} (42)

subject to

\[ \frac{\partial R}{\partial z} \left( z(s) - x(s) \right)\gamma(s) \]  \hspace{1cm} (43)

\[ R(s_1) = R(s_2) = 0 \]  \hspace{1cm} (44)

\[ \frac{\partial u}{\partial z} = u_x \left( x(s), z(s), s \right) \]  \hspace{1cm} (45)

\[ \frac{\partial z}{\partial z} = \eta(s) \]  \hspace{1cm} (46)

\[ \theta \left( \frac{\partial z}{\partial z} \right) = \theta(\eta(s)) \geq 0 \]  \hspace{1cm} (47)

The revenue constraint is captured by (43) and (44). To simplify, it is assumed that zero revenue is to be collected. The rate of change in revenue (43) is derived directly from (41).

The self-selection constraint is represented by (45) - (47); the first-order condition is (45), the second-order condition are (46) and (47).
Introducing the adjoin variables $\Xi(s), \Psi(s), \Omega(s)$ and $\Delta(s)$, the Hamiltonian for the optimization is

$$
H = u(s)\gamma(s) + \lambda(s)(z(s) - x(s))\gamma(s) + \mu(s)u_x(x(s),z(s),s) + \nu(s)\eta(s) + \kappa(s)\theta(s).
$$

and the necessary conditions are

$$
\frac{\partial H}{\partial \eta} = v + \kappa\theta'(\eta) = 0 \quad (49)
$$

$$
\frac{\partial H}{\partial z} = \lambda\left[\frac{\partial}{\partial z}(z(s) - x(s))\right] + \mu\frac{\partial}{\partial z}u_x(x(s),z(s),s) = -v' \quad (50)
$$

$$
\frac{\partial H}{\partial u} = \gamma(s) + \lambda\left[\frac{\partial}{\partial u}(z(s) - x(s))\right] + \mu\frac{\partial}{\partial u}u_x(x(s),z(s),s) = -\mu' \quad (51)
$$

$$
\frac{\partial H}{\partial R} = -\lambda' = 0 \quad (52)
$$

$$
\kappa \frac{\partial \delta}{\delta \delta} = 0, \quad \kappa \geq 0 \quad (53)
$$

with transversality conditions are

$$
\mu(s_1) = \mu(s_2) = 0, \quad \nu(s_1) = \nu(s_2) = 0 \quad (54)
$$

To derive the form of these conditions that will be used below, note that from the identity $u(x(s),z(s),s)$ it follows that

$$
\frac{\partial \delta}{\partial \delta} = -\frac{u_x}{u_z} = \Phi \quad (55)
$$

and
\[
\frac{\partial x}{\partial u} = \frac{1}{u_x}.
\]  

(56)

In addition

\[
\frac{\partial u_z(x(s), z(s), s)}{\partial z} = u_{sx} \frac{\partial x}{\partial z} + u_z = -u_z \Phi_x
\]  

(57)

and

\[
\frac{\partial u_z(x(s), z(s), s)}{\partial u} = u_{sx} \frac{\partial x}{\partial u} = \frac{u_{sx}}{u_x}
\]  

(58)

Now denoting \( \phi(s) = \kappa(s) \Phi(\eta(s)) \), (49) - (54) can be rewritten,

\[
- \mu u_x \Phi_x + \lambda (1 - \Phi) \gamma - \phi' = 0
\]  

(59)

\[
\mu' + \mu \frac{u_{sx}}{u_x} + \left[ 1 - \frac{\lambda}{u_x} \right] \gamma = 0
\]

(60)

\[
\phi \frac{\partial z}{\partial s} = 0 \quad \phi \geq 0
\]

(61)

\[
\mu(s_1) = \mu(s_2) = 0, \quad \phi(s_1) = \phi(s_2) = 0
\]

(62)

The interpretation of these necessary conditions is as follows.

(1) If \( \Box \) is zero for all \( s \). The second order condition for the satisfaction of the self-selection constraint is not binding and pre-tax income is a strictly increasing function of ability. The first-order approach is identical to the second-order approach.

(2) If \( \Box \) is not zero for all \( s \). If \( \Box \) is positive over \([s_0, s_1]\), all households with abilities falling in this interval earn the same pre-tax income. These households are bunched at a single income level. Furthermore they must have the same level of consumption. Note that although pre-tax income and consumption are identical, utility is increasing with \( s \) over the bunched households since those with higher \( s \) have to work less to obtain the common level of income.
There are several theoretical results on optimal income tax.

**Theorem 1** (Mirrlees (1971))

If there exists an ability level \( s_0 \in s \) such that \( l(s_0) = 0 \), then \( l(s) = 0 \) for any \( s < s_0 \).

**Proof** see Myles (1995, pp.148-9) for proof

This result implies that, without specifying the tax function, that the optimal tax system may generate unemployment in the sense that it results in low ability households choosing to do no work. As these households are productive whenever their ability level is non-zero, and output would increase if they did work, this carries important implications for the relation between optimal taxes and the achievement of maximum potential output.

**Theorem 2** (Mirrlees (1971))

The marginal tax rate is always less than or equal to 1.

**Proof**

Self-selection requires that \( u_x x' + u_z z' = 0 \) and \( z' \geq 0 \), therefore \( x' \geq 0 \). By definition, \( x(s) = c(z(s)) \) so that \( x' = c' z' \). Hence \( c' \geq 0 \). Since \( c(z) = z - T(z) \), the fact that \( c' \geq 0 \) implies that \( c' = 1 - T' \geq 0 \iff T' \leq 1 \), so the marginal tax rate is always less than or equal to 1.

**Theorem 3** (Seade (1982))

Assuming agent monotonicity, if leisure is not an inferior good and \( u_{zx} \geq 0 \), then the marginal tax rate is positive.

**Proof**


**Theorem 4** (Seade (1982))

Let the upper bound on ability \( s_2 \) be finite. Then the marginal rate of tax must be 0 for a household of ability \( s_2 \).

**Proof**


In short, since there is no household beyond ability \( s_2 \), there is no point to set the marginal rate of tax positive at \( s_2 \).

**Theorem 5** (Seade (1977))
For a population with bounded ability, any income tax schedule with a positive marginal rate at the top of the scale can be replaced by one that leaves all households better off, including them to earn more income but paying the same tax.

**Proof**


**Theorem 6** (Seade (1977))

If there is no bunching at the lowest income, the optimal marginal rate for the household of lowest ability is zero.

**Proof**

See Myles (1995, pp.154). This is a Mirror image of Theorem 4.

The above results of the optimal non-linear tax have been derived. The optimal marginal rate of taxation must be between 0 and 1. At the highest and lowest abilities, the tax rate must be zero. The latter finding shows that the optimal tax function cannot be progressive. In other words, it may be optimal to force some households to choose to undertake no work. In this case, it is the lowest ability households that will not work. Pre-tax income and consumption must both be increasing functions of ability.

**Numerical Results**

To generate numerical results, Mirrlees (1971) assumed that the social welfare function took the form

\[
\begin{align*}
    w &= \int_0^\infty \frac{1}{\nu} e^{-\nu U} \gamma(s) ds, \quad \nu > 0 \\
    &= \int_0^\infty U \gamma(s) ds, \quad \nu = 0
\end{align*}
\]  

(63)

Higher values of \( \nu \) represent greater concern for equity, with \( \nu = 0 \) representing the utilitarian case.

The individual utility function was the Cobb-Douglas,

\[ U = \log x + \log(1 - l) \]  

(64)
and the skill distribution is log-normal,

\[
\gamma(s) = \frac{1}{s} \exp \left[ -\frac{(\log(s + 1))^2}{2} \right]
\]  

(65)

With a standard deviation \( \bar{\sigma} \) (=0.39 from Lydall (1968)). An implicit assumption is that the skill distribution can be inferred directly from an observed income distribution.

Table 3  Optimal Tax Schedule

<table>
<thead>
<tr>
<th>Income</th>
<th>Consumption</th>
<th>Average tax (%)</th>
<th>Marginal tax (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) ( z^G=0.013, \nu=0, \bar{\sigma}=0.39 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>0.03</td>
<td>---</td>
<td>23</td>
</tr>
<tr>
<td>0.05</td>
<td>0.07</td>
<td>-34</td>
<td>26</td>
</tr>
<tr>
<td>0.10</td>
<td>0.10</td>
<td>-5</td>
<td>24</td>
</tr>
<tr>
<td>0.20</td>
<td>0.18</td>
<td>9</td>
<td>21</td>
</tr>
<tr>
<td>0.30</td>
<td>0.26</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>0.40</td>
<td>0.34</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>0.50</td>
<td>0.43</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>(b) ( z^G=0.003, \nu=1, \bar{\sigma}=0.39 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>0.05</td>
<td>---</td>
<td>30</td>
</tr>
<tr>
<td>0.05</td>
<td>0.08</td>
<td>-66</td>
<td>34</td>
</tr>
<tr>
<td>0.10</td>
<td>0.12</td>
<td>-34</td>
<td>32</td>
</tr>
<tr>
<td>0.20</td>
<td>0.19</td>
<td>7</td>
<td>28</td>
</tr>
<tr>
<td>0.30</td>
<td>0.26</td>
<td>13</td>
<td>25</td>
</tr>
<tr>
<td>0.40</td>
<td>0.34</td>
<td>16</td>
<td>22</td>
</tr>
<tr>
<td>0.50</td>
<td>0.41</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>(c) ( z^G=0.013, \nu=1, \bar{\sigma}=1 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>0.10</td>
<td>---</td>
<td>50</td>
</tr>
<tr>
<td>0.10</td>
<td>0.15</td>
<td>-50</td>
<td>58</td>
</tr>
<tr>
<td>0.25</td>
<td>0.20</td>
<td>20</td>
<td>60</td>
</tr>
<tr>
<td>0.50</td>
<td>0.30</td>
<td>40</td>
<td>59</td>
</tr>
<tr>
<td>1.00</td>
<td>0.52</td>
<td>48</td>
<td>57</td>
</tr>
<tr>
<td>1.50</td>
<td>0.73</td>
<td>51</td>
<td>54</td>
</tr>
<tr>
<td>2.00</td>
<td>0.97</td>
<td>51</td>
<td>52</td>
</tr>
<tr>
<td>3.00</td>
<td>1.47</td>
<td>51</td>
<td>49</td>
</tr>
</tbody>
</table>

Source: Mirrlees (1971)

---

This part draws from Myles (1955) Chap 2, pp.156-9.
The most important feature of the first two panels (a) and (b) in Table 3 is the low marginal rates of tax, with the maximal rate being only 34%. There is also limited deviation in these rates. The marginal rates become lower at high incomes but do not reach 0 because the skill distribution is unbounded. The average rate of tax is negative for low incomes so that low-income consumers are receiving an income supplement from the government.

The panel (c) of Table 3 show the effect of increasing the dispersion of skills (changing standard deviation from 0.39 to 1.00). This raises the marginal tax rates but there remain fairly constant across the income range. Kanbur and Tuomala (1994) find that an increased dispersion of skills raises the marginal tax rate at each income level and that it also has the effect of moving the maximum tax rate up the income range, so that the marginal tax rate is increasing over the majority of households.

Atkinson (1975) considered the effect of changing the social welfare function to the extreme maxi-min form,

$$w = \min \{U\}$$

(66)

From the above table, it can be seen that increased concern for equity, $v$ going from 0 to 1, increased the optimal marginal tax rates. The natural question is “can strong equity considerations such as maxi-min SWF lead to high marginal rates?”. The result is given the below table. The maxi-min criterion leads to generally higher taxes. However they are again highest at low incomes and then decline. Absolute rate is lower than expected in all cases.

<p>| Table 4 Optimal tax Schedule: Utilitarian vs. Maxi-min |
|----------------------------------------|----------------------------------------|----------------------------------------|----------------------------------------|</p>
<table>
<thead>
<tr>
<th>Level of $s$</th>
<th>Utilitarian</th>
<th>Maxi-min</th>
<th>Utilitarian</th>
<th>Maxi-min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average rate (%)</td>
<td>Marginal rate (%)</td>
<td>Average rate (%)</td>
<td>Marginal rate (%)</td>
<td></td>
</tr>
<tr>
<td>Median (50%)</td>
<td>6</td>
<td>21</td>
<td>10</td>
<td>52</td>
</tr>
<tr>
<td>Top decile (10%)</td>
<td>14</td>
<td>20</td>
<td>28</td>
<td>34</td>
</tr>
<tr>
<td>Top percentile (1%)</td>
<td>16</td>
<td>17</td>
<td>28</td>
<td>26</td>
</tr>
</tbody>
</table>

Source: Atkinson and Stiglitz (1980, Table 13-3, p.421)
Empirical Fact in Japan

Table 5  Individual Inhabitants’ and Income Tax

<table>
<thead>
<tr>
<th>Tax Recipient</th>
<th>Inhabitants’ Tax</th>
<th>Income Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Municipal governments on 1st January</td>
<td>¥330,000</td>
<td>¥380,000</td>
</tr>
<tr>
<td>National government</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tax Payer</th>
<th>Individual lives in a municipal government</th>
<th>Individual lives in Japan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax Method</td>
<td>With holding</td>
<td>Assessment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tax Base</th>
<th>Last year’s income</th>
<th>This year’s income</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Income Deduction     | •Basic deduction ¥330,000                                      | •Basic deduction ¥380,000                        |
|                      | •Spouse deduction max ¥330,000                                 | •Spouse deduction max ¥380,000                   |
|                      | •Special spouse deduction                                     | •Special spouse deduction max ¥380,000          |
|                      | •Family deduction ¥330,000                                     | •Family deduction ¥380,000                       |

<table>
<thead>
<tr>
<th>Minimum Taxable Income (1998)</th>
<th>A couple with 2 children Before a special tax deduction ¥3,031,000</th>
<th>A couple with 2 children Before a special tax deduction ¥3,616,000</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>After special tax deduction ¥4,373,000</td>
<td>After special tax deduction ¥4,917,000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Taxable Income</th>
<th>Tax Rates</th>
<th>Taxable Income</th>
<th>Tax Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below ¥2,000,000</td>
<td>2%</td>
<td>Prefectural: 3%</td>
<td>Total: 5%</td>
</tr>
<tr>
<td>Below ¥7,000,000</td>
<td>2%</td>
<td>Municpal: 7%</td>
<td>10%</td>
</tr>
<tr>
<td>Above ¥7,000,000</td>
<td>3%</td>
<td>Municpal: 12%</td>
<td>15%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Capital Gain Income</th>
<th>Preference: 2%</th>
<th>Municipal: 4%</th>
<th>Total: 6%</th>
<th>Preference: 2%</th>
<th>Municipal: 5.5%</th>
<th>Total: 7.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below ¥40,000,000</td>
<td>2%</td>
<td>4%</td>
<td>6%</td>
<td>Below ¥40,000,000</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>Below ¥80,000,000</td>
<td>2%</td>
<td>5.5%</td>
<td>7.5%</td>
<td>Below ¥80,000,000</td>
<td>2%</td>
<td>5.5%</td>
</tr>
<tr>
<td>Above ¥80,000,000</td>
<td>3%</td>
<td>6%</td>
<td>9%</td>
<td>Above ¥80,000,000</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>In 1999-2001</td>
<td>Prefectural: 2%</td>
<td>Municipal: 4%</td>
<td>Total: 6%</td>
<td>Prefectural: 2%</td>
<td>Municipal: 4%</td>
<td>Total: 6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tax Exemption</th>
<th>- Adjustment for Double Taxation</th>
<th>- Adjustment for Double Taxation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>•Dividends</td>
<td>•Dividends</td>
</tr>
<tr>
<td></td>
<td>•Foreign Tax</td>
<td>•Foreign Tax</td>
</tr>
<tr>
<td></td>
<td>- Tax Incentives</td>
<td>- Tax Incentives</td>
</tr>
<tr>
<td></td>
<td>•Housing Acquisition</td>
<td>•Housing Acquisition</td>
</tr>
<tr>
<td></td>
<td>•Research &amp; Development</td>
<td>•Research &amp; Development</td>
</tr>
</tbody>
</table>

| Tax Revenue (1996) | 90,17 trillion yen | 189,649 trillion yen |
Table 6  Historical Changes in Income Tax Rates Schedule

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax Rate</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>% million</td>
<td>% million</td>
<td>% million</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10.5</td>
<td>10.5</td>
<td>10</td>
<td>10 (→3)</td>
<td>10 (→3.3)</td>
<td>10 (→3.3)</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>12</td>
<td>16</td>
<td>20</td>
<td>20 (→6)</td>
<td>20 (→9)</td>
<td>20 (→9)</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>16</td>
<td>20</td>
<td>40 (→1)</td>
<td>30 (→18)</td>
<td>30 (→18)</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>20</td>
<td>40 (→2)</td>
<td>40 (→2)</td>
<td>40 (→30)</td>
<td>40 (→30)</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>21</td>
<td>25</td>
<td>50</td>
<td>50 (2→)</td>
<td>50 (30→)</td>
<td>50 (30→)</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>25</td>
<td>30</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>30</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>35</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>45</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>50</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum Inhabitant’s Tax Rate</td>
<td>18%</td>
<td>18%</td>
<td>18%</td>
<td>16%</td>
<td>15%</td>
<td>15%</td>
<td>13%</td>
</tr>
<tr>
<td>Combined Maximum Tax Rate</td>
<td>93%</td>
<td>88%</td>
<td>78%</td>
<td>76%</td>
<td>65%</td>
<td>65%</td>
<td>50%</td>
</tr>
<tr>
<td>Number of Tax Brackets</td>
<td>19 (13)</td>
<td>15 (14)</td>
<td>12 (14)</td>
<td>6 (7)</td>
<td>5 (3)</td>
<td>5 (3)</td>
<td>4 (3)</td>
</tr>
<tr>
<td>Minimum Taxable Income</td>
<td>¥1,707,000</td>
<td>¥2,357,000</td>
<td>¥2,615,000</td>
<td>¥2,619,000</td>
<td>¥3,198,000</td>
<td>¥3,539,000 (¥3,616,000 after the 1998 amendment)</td>
<td>¥3,821,000</td>
</tr>
</tbody>
</table>
Table 7  Change in Income Tax Payers

<table>
<thead>
<tr>
<th></th>
<th>Salary Earners</th>
<th>Declared Earners</th>
<th>Total</th>
<th>Business</th>
<th>Agriculture</th>
<th>Small Business</th>
<th>Others</th>
<th>Applicants for Refund</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td>3,665</td>
<td>737</td>
<td>4,402</td>
<td>228</td>
<td>31</td>
<td>67</td>
<td>411</td>
<td>599</td>
</tr>
<tr>
<td>1986</td>
<td>3,728</td>
<td>770</td>
<td>4,498</td>
<td>231</td>
<td>32</td>
<td>70</td>
<td>437</td>
<td>654</td>
</tr>
<tr>
<td>1987</td>
<td>3,767</td>
<td>771</td>
<td>4,538</td>
<td>235</td>
<td>25</td>
<td>70</td>
<td>441</td>
<td>699</td>
</tr>
<tr>
<td>1988</td>
<td>3,909</td>
<td>780</td>
<td>4,689</td>
<td>245</td>
<td>24</td>
<td>70</td>
<td>441</td>
<td>696</td>
</tr>
<tr>
<td>1989</td>
<td>3,961</td>
<td>796</td>
<td>4,758</td>
<td>242</td>
<td>23</td>
<td>67</td>
<td>464</td>
<td>659</td>
</tr>
<tr>
<td>1990</td>
<td>4,158</td>
<td>855</td>
<td>5,013</td>
<td>250</td>
<td>25</td>
<td>67</td>
<td>513</td>
<td>663</td>
</tr>
<tr>
<td>1991</td>
<td>4,333</td>
<td>856</td>
<td>5,189</td>
<td>252</td>
<td>24</td>
<td>68</td>
<td>512</td>
<td>699</td>
</tr>
<tr>
<td>1992</td>
<td>4,403</td>
<td>858</td>
<td>5,261</td>
<td>247</td>
<td>24</td>
<td>69</td>
<td>518</td>
<td>735</td>
</tr>
<tr>
<td>1993</td>
<td>4,473</td>
<td>843</td>
<td>5,316</td>
<td>230</td>
<td>22</td>
<td>67</td>
<td>524</td>
<td>781</td>
</tr>
<tr>
<td>1994</td>
<td>4,478</td>
<td>822</td>
<td>5,300</td>
<td>223</td>
<td>27</td>
<td>62</td>
<td>510</td>
<td>867</td>
</tr>
<tr>
<td>1995</td>
<td>4,484</td>
<td>802</td>
<td>5,286</td>
<td>213</td>
<td>19</td>
<td>60</td>
<td>510</td>
<td>864</td>
</tr>
<tr>
<td>1996</td>
<td>4,537</td>
<td>824</td>
<td>5,361</td>
<td>213</td>
<td>20</td>
<td>60</td>
<td>531</td>
<td>883</td>
</tr>
<tr>
<td>1997</td>
<td>4,627</td>
<td>827</td>
<td>5,454</td>
<td>206</td>
<td>16</td>
<td>60</td>
<td>545</td>
<td>909</td>
</tr>
<tr>
<td>1998</td>
<td>4,703</td>
<td>848</td>
<td>5,551</td>
<td>210</td>
<td>20</td>
<td>61</td>
<td>557</td>
<td>---</td>
</tr>
</tbody>
</table>

Source: Ministry of Finance.

Table 8  The 1996 Declared Income Tax Burden Rate by Income Class

<table>
<thead>
<tr>
<th>Income class</th>
<th>Average Income Deduction (million yen)</th>
<th>Average Taxable Income (million yen)</th>
<th>Average Calculated Tax (million yen)</th>
<th>Average Tax Deduction (million yen)</th>
<th>Average Tax Payment (million yen)</th>
<th>Gross Income Tax Rate (%)</th>
<th>Effective Income Tax Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>– 100</td>
<td>748</td>
<td>546</td>
<td>202</td>
<td>0</td>
<td>18</td>
<td>2.4%</td>
<td>8.9%</td>
</tr>
<tr>
<td>100 – 200</td>
<td>1,540</td>
<td>995</td>
<td>545</td>
<td>56</td>
<td>0</td>
<td>3.1%</td>
<td>8.8%</td>
</tr>
<tr>
<td>200 – 300</td>
<td>2,476</td>
<td>1,459</td>
<td>1,016</td>
<td>103</td>
<td>1</td>
<td>3.5%</td>
<td>8.7%</td>
</tr>
<tr>
<td>300 – 500</td>
<td>3,870</td>
<td>1,796</td>
<td>2,074</td>
<td>233</td>
<td>6</td>
<td>5.1%</td>
<td>9.5%</td>
</tr>
<tr>
<td>500 – 1000</td>
<td>6,899</td>
<td>2,020</td>
<td>4,879</td>
<td>700</td>
<td>14</td>
<td>9.2%</td>
<td>13.0%</td>
</tr>
<tr>
<td>1000+</td>
<td>23,005</td>
<td>1,893</td>
<td>21,112</td>
<td>4,789</td>
<td>32</td>
<td>20.5%</td>
<td>22.3%</td>
</tr>
<tr>
<td>Total Average</td>
<td>5,864</td>
<td>1,572</td>
<td>4,292</td>
<td>795</td>
<td>8</td>
<td>12.9%</td>
<td>17.7%</td>
</tr>
</tbody>
</table>

### Table 9  Number of Salary Earners, Total Salary and Tax in 1996

<table>
<thead>
<tr>
<th>Income Class</th>
<th>Number of Salary Earners (thousands) Share</th>
<th>Total Salary (100 million yen) Share</th>
<th>Tax paid (100 million yen) Share</th>
<th>Effective Tax Rate Share (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 100</td>
<td>3,228 7.2 479 1.2</td>
<td>23,993 1.2</td>
<td>2,789 0.1</td>
<td>213 0.2 7.6</td>
</tr>
<tr>
<td>100 – 200</td>
<td>4,818 10.7 3,419 8.7</td>
<td>72,423 3.5</td>
<td>54,310 2.8</td>
<td>1,345 1.3 2.5</td>
</tr>
<tr>
<td>200 – 300</td>
<td>6,818 15.2 6,230 15.9</td>
<td>173,522 8.4</td>
<td>158,778 8.0</td>
<td>5,021 4.9 3.2</td>
</tr>
<tr>
<td>300 – 400</td>
<td>7,780 17.5 7,328 18.7</td>
<td>272,122 13.2</td>
<td>256,351 13.0</td>
<td>8,471 8.2 3.3</td>
</tr>
<tr>
<td>400 – 500</td>
<td>6,530 14.5 6,244 15.9</td>
<td>292,908 13.2</td>
<td>263,861 13.4</td>
<td>8,898 8.7 3.4</td>
</tr>
<tr>
<td>500 – 600</td>
<td>4,964 11.1 4,802 12.3</td>
<td>272,721 13.2</td>
<td>263,861 13.4</td>
<td>8,898 8.7 3.4</td>
</tr>
<tr>
<td>600 – 700</td>
<td>3,273 7.3 3,215 8.2</td>
<td>211,900 10.2</td>
<td>208,179 10.6</td>
<td>7,429 7.2 3.6</td>
</tr>
<tr>
<td>700 – 800</td>
<td>2,384 5.3 2,372 6.1</td>
<td>178,002 8.6</td>
<td>177,094 9.0</td>
<td>7,707 7.5 4.4</td>
</tr>
<tr>
<td>800 – 900</td>
<td>1,604 3.6 1,604 4.1</td>
<td>135,837 6.6</td>
<td>135,837 6.9</td>
<td>7,168 7.0 5.3</td>
</tr>
<tr>
<td>900 – 1000</td>
<td>1,004 2.2 1,004 2.6</td>
<td>95,168 4.6</td>
<td>95,168 4.8</td>
<td>6,137 6.0 6.4</td>
</tr>
<tr>
<td>1000 – 1500</td>
<td>1,963 4.4 1,963 5.0</td>
<td>232,610 11.2</td>
<td>232,610 11.8</td>
<td>21,461 20.9 9.2</td>
</tr>
<tr>
<td>1500 – 2000</td>
<td>378 0.8 378 1.0</td>
<td>64,247 3.1</td>
<td>64,247 3.3</td>
<td>9,475 9.2 14.7</td>
</tr>
<tr>
<td>2000 – 2500</td>
<td>87 0.2 87 0.2</td>
<td>20,098 1.0</td>
<td>20,098 1.0</td>
<td>3,956 3.8 19.7</td>
</tr>
<tr>
<td>2500+</td>
<td>64 0.1 64 0.2</td>
<td>23,254 1.1</td>
<td>23,254 1.2</td>
<td>6,294 6.1 27.1</td>
</tr>
<tr>
<td>Total</td>
<td>44,896 100.0 39,189 100.0</td>
<td>2,068,805 100.0</td>
<td>1,972,750 100.0</td>
<td>102,797 100.0 5.2</td>
</tr>
</tbody>
</table>

Note: Employees on December 31, 1996.

### Table 10  Tax Elasticity

<table>
<thead>
<tr>
<th>Year</th>
<th>Direct Tax (Income Tax)</th>
<th>Total Tax Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual</td>
<td>Corporate</td>
</tr>
<tr>
<td>1985</td>
<td>1.00</td>
<td>1.31</td>
</tr>
<tr>
<td>1986</td>
<td>1.89</td>
<td>1.74</td>
</tr>
<tr>
<td>1987</td>
<td>3.64</td>
<td>2.58</td>
</tr>
<tr>
<td>1988</td>
<td>2.25</td>
<td>1.57</td>
</tr>
<tr>
<td>1989</td>
<td>1.66</td>
<td>2.73</td>
</tr>
<tr>
<td>1990</td>
<td>1.36</td>
<td>2.69</td>
</tr>
<tr>
<td>1991</td>
<td>0.02</td>
<td>0.06</td>
</tr>
<tr>
<td>1992</td>
<td>□ 6.18</td>
<td>□ 6.09</td>
</tr>
<tr>
<td>1993</td>
<td>□ 4.14</td>
<td>2.14</td>
</tr>
<tr>
<td>1994</td>
<td>4.75</td>
<td>5.50</td>
</tr>
<tr>
<td>1995</td>
<td>1.10</td>
<td>□ 2.00</td>
</tr>
</tbody>
</table>
Figure 6  Before Tax Household Income Distribution in Japan

Note: Monthly average income during September through November.

Figure 7  Before Tax Household Income Distribution in Japan (log normal transformation)

Note: Monthly average income during September through November.
Figure 8  After Tax Household Income Distribution in Japan

Note: Monthly average income during September through November.

Figure 9  After Tax Household Income Distribution in Japan (log normal transformation)

Note: Monthly average income during September through November.
Exercises

1. (Atkinson and Stiglitz (1980), p.376) For the utility function \( U = \sum_{i=1}^{n} A_i \frac{X_i^{1-\frac{1}{\epsilon_i}}}{1-1/\epsilon_i} - vL \), where \( L \) units of labor, wage is \( w \) thus the budget constraint for the household is \( \sum q_iX_i = wL = y \). Show that the income terms \( \frac{\partial X_i}{\partial y} \) and cross price terms are zero.

Derive the optimal tax structure where \( \epsilon_i \) are (positive) constants.

2. Recently many countries adopt indirect taxation (e.g. VAT) and shift its weight from direct taxation (e.g. individual income). Could you justify this shift of tax reform?

3. Stiglitz once argued that “it can be shown, that if one has a well-designed income tax, adding differential commodity taxation is likely to add little, if anything.” Would you agree with him? Or in what circumstances does the use of commodity taxation allow a higher level of social welfare to be achieved in the presence of income taxation?
References on Personal Income Tax


