Chapter 3 Individual Income Taxation

3.1 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.

3.2 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), and 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 separately. 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 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>

Table 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

---

1 This Part draws heavily from Rosen (1999), pp.258-260.
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 section, we assume they are defined in terms of average tax rates.

**The degree of progression**

Progression in the income tax schedule introduces disproportionality into the distribution of the tax burden and exerts a redistributive effect on the distribution of income. In order to explore these properties further, we need to be able to measure the degree of income tax progression along the income scale. Such measures are called measures of *structural progression* (sometimes, measures of *local progression*). There is more than one possibility, as we will see. Each such measure will induce a partial ordering on the set of all possible income tax schedules. We could not expect always to be able to rank a schedule \( t_2(x) \) unambiguously more, or less, structurally progressive than another schedule \( t_1(x) \): we must allow that a schedule could display more progression in one income range and less in another.

Nevertheless, the policymaker and tax practitioner, and indeed the man in the street, would like to be able to say which of any two alternative income tax systems is the more progressive *in its effects*. Is the federal income tax in the USA more redistributive than the personal income tax in Germany? This sort of question will take us from measures of structural progression to measures of *effective progression*. Measuring effective progression is a matter of reducing a tax schedule and income distribution pair to a scalar index number. The same schedule \( t(x) \) could be more progressive *in effect* when applied to distribution A than to distribution B. Trends in effective progression for a given country over time, as well as differences between the income taxes of different countries, can be examined using such index numbers.

Let us begin by defining as \( m(x) \) and \( a(x) \) respectively the marginal and average rates of tax experienced by an income \( x \):

\[
a(x) = \frac{t(x)}{x} \quad \text{and} \quad m(x) = t'(x)
\]  

(1)

Since

\[
\frac{d[t(x)]}{dx} = \frac{xt'(x) - t(x)}{x^2} = \frac{m(x) - a(x)}{x}
\]

(2)
for strict progression it is necessary and sufficient that

\[ m(x) > a(x) \text{ for all } x \]  

(3)

The strict inequality rules out the possibility that the tax could be proportional to income in any interval: we may relax it if we wish. Measures of structural progression quantify, in various ways, the excess of the marginal rate \( m(x) \) over the average rate \( a(x) \) at income level \( x \).

We introduce two particularly important measures here.

First, liability progression \( LP(x) \) is defined at any income level \( x \) for which \( t(x) > 0 \) as the elasticity of tax liability to pre-tax income:

\[ LP(x) = e^{t(x)x} = \frac{x t'(x)}{t(x)} = \frac{m(x)}{a(x)} > 1 \]

(4)

As we have already noted, for a strictly progressive income tax a 1 percent increase in pre-tax income \( x \) leads to an increase of more than 1 percent in tax liability. \( LP(x) \) measures the actual percentage increase experienced. A change of tax schedule which, for some \( x_0 \), casues an increase in \( LP(x_0) \) connotes, in an obvious sense, an increase in progression at that income level \( x_0 \). If the change in a strictly positive income tax involves an upward shift of the entire function \( LP(x) \), then the tax has become everywhere more liability progressive.

Second, residual progression \( RP(x) \) is defined at all income levels \( x \) as the elasticity of post-tax income to pre-tax income:

\[ RP(x) = e^{x-t(x)x} = \frac{x[1-t'(x)]}{x-t(x)} = \frac{1-m(x)}{1-a(x)} < 1 \]

(5)

The counterpart to the observation above is that a 1 percent increase in pretax income \( x \) leads to an increase of less than 1 percent in post-tax income. \( RP(x) \) measures the actual percentage increase in post-tax income. A reduction in \( RP(x) \) must clearly be interpreted as an increase in progression, according to this measure.

These measures were first proposed by Musgrave and Thin (1948). It is inconvenient that \( RP(x) \) should decrease when the tax becomes more progressive. We therefore make a minor change of definition in this book, which is not standard in all of the literature, replacing the measure \( RP(x) \) by its reciprocal:
\[
RP^*(x) = \frac{1}{RP(x)} = \frac{1-a(x)}{1-m(x)} > 1
\]  

\(RP^*(x)\) can be interpreted as the elasticity of pre-tax income to post-tax income. An increase in \(RP^*(x)\) makes the tax more residually progressive at \(x\).

Each of these measures quantifies in its own way the excess of the marginal rate of tax over the average rate of tax at the income level \(x\), and accordingly induces a different partial ordering on the set of all tax schedules. Equipped with these measures, we shall, in the next two sections of the chapter, be able to demonstrate that if the income tax becomes structurally more progressive, and the pre-tax income distribution does not change, then this implies enhanced deviation from proportionality (in the case of liability progression) and enhanced redistributive effect (in the case of residual progression).

### 3.3 The Individual or Family-Based Principle: Choice of Unit

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 2  Tax liabilities under a hypothetical tax system**

<table>
<thead>
<tr>
<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>12,200</td>
<td>30,000</td>
</tr>
<tr>
<td>Ricky</td>
<td>29,000</td>
<td>12,100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethel</td>
<td>15,000</td>
<td>5,100</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></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

---

2 This Part draws heavily on Rosen (1999, Chap.16, pp.363-5).
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 \[.10 \times 6,000 + .50 \times 23,000\].) 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 on $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’s 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 the Tax Reform Act of 1986 (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 adjusted gross income (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.”

### 3.4 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)\(^3\). Effectively, they tried to determine the tax system that

---

\(^3\) It is possible to include administrative costs in the tax analysis both in theory and in empirical investigations. It is rather important to take administration into account in tax analysis.
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,

\[ w = \frac{1}{1 - \varepsilon} \sum_h \left( \frac{u^h}{\varepsilon} \right)^{1 - \varepsilon} \]

for \( \varepsilon \neq 1 \)

\[ = \sum_h \log \left( \frac{u^h}{\varepsilon} \right) \]

for \( \varepsilon = 1 \)

(6)

where \( \varepsilon \) is a degree of inequality aversion.

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 in 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.

3.5 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.

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, $t$, were to change with every change in income, we would be looking for an optimal income schedule

$$ t = f(Y) \tag{7} $$

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 (7) to be complicated. In the case of the linear income tax, we had to find values

---

4 This part quotes Hillman (2003, Chap.7, pp.479-83.)
for income subsidy $S$ and $t$ related through the government budget. In the case of the tax schedule (7), 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 (7) 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 than 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 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.

3.6 The Mirrleesian Economy

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

1. Two commodities: a consumption good $x$ ($x \geq 0$) and a single labor service, $l$ ($0 \leq l \leq 1$).
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 $z(s) \equiv s l(s)$.

---

5 This part draws heavily from Myles (1995) Chap 2, pp.133-55.
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

\[
x(s) \leq c(z(s)) = z(s) - T(z(s))
\]  

(8)

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 Distribution of ability \( s \)**

7. All households have the same strictly concave utility function

\[
U = U(x, l)
\]  

(9)

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

\[
\text{Max } U(x, l) \text{ subject to } x(s) \leq c(sl(s))
\]  

(10)

In the absence of income taxation, a household of ability \( s \) would face the budget constraint
From (11), 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.

Figure 2 presents one case for the general shape of the pattern of pre-tax and post-tax income, \(z\) and \(Disp(z)\) respectively. The curve \(Disp(z)\) describes the case where marginal tax rates defined by \(T'(z)\) are first high for lower income, then rather low for middle-range income, and finally high again for high income. The shape of \(Disp(z)\) is very close to the marginal tax rate structure commonly observed in many countries.

\(Disp(z)\) 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 \(c(z)\) 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 \(s\).
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

\[ U_x > 0, \quad U_l < 0, \quad U_{xx} < 0 \quad (12) \]

and

\[ U_l(x, l) \to -\infty \quad \text{as} \quad l \to 1 \quad (13) \]

(15) 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.
To allow preferences and the budget constraint to be depicted on the same diagram, the utility function can be written

\[ U = U(x, l) = U(x, z/s) = u(x, z, s) \]  \hspace{1cm} (14)

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 3 Preference**

![Figure 3 Preference](image)

**Figure 4 Translation of indifference curves.**

![Figure 4 Translation of indifference curves](image)
Agent Monotonicity

The utility function (14) satisfies agent monotonicity if \(-\frac{u_z}{u_x}\), is a decreasing function of \(s\). Note that \(\Phi \equiv -\frac{u_z}{u_x}\) 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 \(-l \frac{U_l}{U_x}\) is an increasing function of \(l\) as \(-\frac{u_c}{u_x} = -\frac{U_l(x,z/s)}{sU_x(x,z/s)}\). Calculating \(\frac{\partial}{\partial l} \left[-l \frac{U_l}{U_x}\right]\) and \(\Phi_s\) shows

\[
\Phi_s = -\frac{1}{s^2} \frac{\partial}{\partial l} \left[-l \frac{U_l}{U_x}\right] \quad (15)
\]

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-url)
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,

$$u(x(s), z(s), s) \geq u(x(s'), z(s'), s')$$

for all $s$ 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 (10).

$$0 = u(s) - u(x(s), z(s), s) \leq u(s') - u(x(s), z(s), s')$$ \hspace{1cm} (16)

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

$$u'(s) = u(x(s), z(s), s).$$ \hspace{1cm} (17)

From the definition of $u(s)$ it follows that
\begin{align}
  u_s x'(s) + u_s z'(s) &= 0 
\end{align}

(18)

is equivalent to (17).

Condition (17) or (18) 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

\begin{align}
  u''(s) - u_{ss}(x(s), z(s), s) &\geq 0
\end{align}

(19)

Again using the definition of \( u(s) \),

\begin{align}
  u''(s) - u_{ss} x'(s) + u_{ss} z'(s) + u_{ss} 
\end{align}

(20)

which gives, by using (19).

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

(21)

Eliminating \( x'(s) \) using (18) provides the final condition

\begin{align}
  \left[ u_{ss} - u_{ss} \frac{u_s}{u_s} \right] z''(s) = - \Phi_z z'(s) \geq 0 
\end{align}

(22)

where \( \Phi_z \) is the marginal rate of substitution introduced in the discussion of agent monotonicity. With agent monotonicity \( \Phi_z < 0 \), 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.
3.7 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

$$Z = \int_0^\infty z(s)\gamma(s)\,ds$$

(23)

and aggregate demand, $X$, where

$$X = \int_0^\infty x(s)\gamma(s)\,ds$$

(24)

The optimal tax function is then chosen to maximize social welfare, where social welfare is given by the Bergson-Samuelson function.

$$W = \int_0^\infty w(u(s))\gamma(s)\,ds$$

(25)

with $W'' \leq 0$.

There are two constraints upon the maximization of (25). The first is that the chosen allocation must be productively feasible such that,

$$X \leq F(Z)$$

(26)

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 (26) can be written $X \leq \hat{F}(Z - Z^G) = F(Z)$.

Denoting the level of revenue required by $R(\equiv Z^G)$, the revenue constraint can be written

$$R \geq \int_0^\infty [z(s) - x(s)]\gamma(s)\,ds$$

(27)

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

### 3.8 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,

\[ T(sl) = -\tau + tsl \tag{28} \]

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

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

\[ x = \tau + \zeta \cdot sl. \tag{29} \]

Each household chooses consumption and labor supply to maximize utility \((3.2)\) subject to \((29)\).

The first-order conditions can be reduced to

\[ -\frac{U_l}{U_x} = \zeta s. \tag{30} \]

Labor supply and consumption demand functions can be written as,

\[ l = l(\zeta, \tau, s) \]
\[ x = \tau + \zeta sl(\zeta, \tau, s) \tag{31} \]

Substituting \((31)\) into the utility function, there determine the indirect utility function,
\[ U = U(\tau + \zeta s l(\zeta, \tau, s), l(\zeta, \tau, s)) = V(\zeta, \tau, s) \] (32)

with

\[
\frac{\partial V}{\partial \tau} = U_s, \quad \frac{\partial V}{\partial \zeta} = U_s l
\] (33)

where \( \frac{\partial V}{\partial \tau} \) 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, \( R \).

\[
\max_{\tau, \zeta} \int_0^\infty w(V(\zeta, \tau, s))\gamma(s)ds
\] (34)

subject to

\[
\int_0^\infty \left[-\tau + (1 - \zeta) l(\zeta, \tau, s)\right] \gamma(s)ds = R
\] (35)

Using (33) and defining the social marginal utility of income for a household of ability \( s \) by

\[
\beta(s) = w'(V(\zeta, \tau, s)) \frac{\partial V(\zeta, \tau, s)}{\partial \tau}
\] (36)

The necessary conditions for the choice of \( \tau \) and \( \zeta \) respectively are

\[
\int_0^\infty \beta \gamma(s)ds = \lambda \left[ H - \int_0^\infty (1 - \zeta) \frac{\partial \zeta}{\partial \tau} \gamma(s)ds \right]
\] (37)

and

\[
\int_0^\infty \beta \gamma(s)ds = \lambda \int_0^\infty \left[ z - (1 - \zeta) \frac{\partial \zeta}{\partial \zeta} \right] \gamma(s)ds
\] (38)

where \( H \) is the population size, \( H = \int_0^\infty \gamma(s)ds \).
Divide (38) by (37) and denote by a bar term of the form $x/H$.

\[
\frac{\int_0^\infty \beta x(s)ds}{\int_0^\infty \beta x(s)ds} = \frac{\bar{z} - \int_0^\infty (1 - \zeta) \frac{\partial x}{\partial \zeta} x(s)ds}{1 - \int_0^\infty (1 - \zeta) \frac{\partial x}{\partial \tau} x(s)ds}
\] (39)

The term on the left-hand side of (39) 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

\[
\frac{\partial \bar{A}}{\partial \zeta} \bigg|_{R_{\text{const}}} = \frac{-\bar{z} + \int_0^\infty (1 - \zeta) \frac{\partial x}{\partial \zeta} x(s)ds}{1 - \int_0^\infty (1 - \zeta) \frac{\partial x}{\partial \tau} x(s)ds}
\] (40)

Hence from (39) and (40),

\[
z(\beta) = \frac{\partial \bar{z}}{\partial \zeta} \bigg|_{R_{\text{const}}}
\] (41)

Since averaging over the population must give $\bar{z} = \bar{z}(z, \zeta)$, it follows from (41) that, holding revenue constant

\[
\frac{\partial \bar{z}}{\partial \zeta} \bigg|_{R_{\text{const}}} = \frac{\partial \bar{z}}{\partial \zeta} + \frac{\partial \bar{z}}{\partial \tau} \frac{\partial \zeta}{\partial \zeta} \bigg|_{R_{\text{const}}} = \frac{\partial \bar{z}}{\partial \zeta} - \frac{\partial \bar{z}}{\partial \tau} z(\beta)
\] (42)

Therefore (40) can be written in the form

\[
z(\beta) - \bar{z} = (1 - \zeta) \left[ \frac{\partial \bar{z}}{\partial \tau} z(\beta) - \frac{\partial \bar{z}}{\partial \zeta} \right] = -t \frac{\partial \bar{z}}{\partial \zeta} \bigg|_{R_{\text{const}}}
\] (43)

Recall that $t = 1 - \zeta$
where the derivative is taken with revenue constant.

Although the tax rule (44) 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 $\beta$ 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.

### 3.9 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,

$$R(s) = \int_{s_1}^{s_2} (z(s') - x(s'))\gamma(s')ds'$$

(45)

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, $\eta(s) \equiv \partial z/\partial s'$ 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

$$\int_{s_1}^{s_2} u(s)\gamma(s)ds$$

(46)

subject to

$$\frac{\partial R}{\partial \delta} (z(s) - x(s))\gamma(s)$$

(47)

$$R(s_1) = R(s_2) = 0$$

(48)

$$\frac{\partial u}{\partial \delta} = u_s(x(s), z(s), s)$$

(49)
\[
\frac{\partial z}{\partial s} = \eta(s) \quad (50)
\]

\[
\theta \left( \frac{\partial z}{\partial s} \right) = \theta(\eta(s)) \geq 0 \quad (51)
\]

The revenue constraint is captured by (47) and (48). To simplify, it is assumed that zero revenue is to be collected. The rate of change in revenue (47) is derived directly from (48).

The self-selection constraint is represented by (49) – (51); the first-order condition is (49), the second-order condition are (50) and (51).

Introducing the adjoint variables \( \lambda(s) \), \( \mu(s) \), \( \nu(s) \) and \( \kappa(s) \), the Hamiltonian for the optimization is

\[
H = u(s)\gamma(s) + \lambda(s)(z(s) - x(s))\gamma(s) + \mu(s)u_1(x(s), z(s), s) \\
+ \nu(s)\eta(s) + \kappa(s)\theta(s).
\quad (52)
\]

and the necessary conditions are

\[
\frac{\partial H}{\partial \eta} = \nu + \kappa\theta'(\eta) = 0 \quad (53)
\]

\[
\frac{\partial H}{\partial z} = \lambda \left( z(s) - x(s) \right) + \mu \frac{\partial u_1(x(s), z(s), s)}{\partial z} = -\nu' \quad (54)
\]

\[
\frac{\partial H}{\partial u} = \gamma(s) + \lambda \frac{\partial u_1(x(s), z(s), s)}{\partial u} + \mu \frac{\partial u_1(x(s), z(s), s)}{\partial u} = -\mu' \quad (55)
\]

\[
\frac{\partial H}{\partial R} = -\lambda' = 0 \quad (56)
\]

\[
\kappa \frac{\partial z}{\partial s} = 0, \quad \kappa \geq 0 \quad (57)
\]

with transversality conditions are
\( \mu(s_1) = \mu(s_2) = 0, \quad \nu(s_1) = \nu(s_2) = 0 \) \hspace{1cm} (58)

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

\[ \frac{\partial \hat{\alpha}}{\partial z} = \frac{u_x}{u_z} = \Phi \] \hspace{1cm} (59)

and

\[ \frac{\partial \hat{\alpha}}{\partial u} = \frac{1}{u_x}. \] \hspace{1cm} (60)

In addition

\[ \frac{\partial \phi}{\partial s} \right|_{x = x(s), z = z(s), s} = u_{xx} \frac{\partial \hat{\alpha}}{\partial z} + u_z = -u_x \Phi_s \] \hspace{1cm} (61)

and

\[ \frac{\partial \phi}{\partial u} |_{x = x(s), z = z(s), s} = u_{xx} \frac{\partial \hat{\alpha}}{\partial u} = \frac{u_{xx}}{u_x} \] \hspace{1cm} (62)

Now denoting \( \phi(s) = \kappa(s) \Theta(\eta(s)) \), (53)–(58) can be rewritten,

\[ -\mu u_x \Phi_s + \lambda (1 - \Phi) \gamma + \phi = 0 \] \hspace{1cm} (63)

\[ \mu' + \mu \frac{u_{xx}}{u_x} + \left[ 1 - \frac{\lambda}{u_x} \right] \gamma = 0 \] \hspace{1cm} (64)

\[ \frac{\partial \phi}{\partial s} = 0 \quad \phi \geq 0 \] \hspace{1cm} (65)

\[ \mu(s_1) = \mu(s_2) = 0, \quad \phi(s_1) = \phi(s_2) = 0 \] \hspace{1cm} (66)
The interpretation of these necessary conditions is as follows.

1. $\phi$ 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. $\phi$ is not zero for all $s$. If $\phi$ 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 $u_{zz} \geq 0$, 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.

3.10 Numerical Results

To generate numerical results, Mirrlees (1971) assumed that the social welfare function took the form
\[ w = \int_0^\infty \frac{1}{v} e^{-wU} \gamma(s) ds, \quad v > 0 \]
\[ = \int_0^\infty U \gamma(s) ds, \quad v = 0 \]  
(67)

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

The individual utility function was the Cobb-Douglas,

\[ U = \log x + \log(1 - l) \]  
(68)

and the skill distribution is log-normal,

\[ \gamma(s) = \frac{1}{s} \exp \left[ -\frac{(\log(s + 1))^2}{2} \right] \]  
(69)

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

### Table 3.1 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, v=0, \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, v=1, \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, v=1, \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)
The most important feature of the first two panel (a) and (b) in Table xx 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 3.1 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,

$$w = \min \{ U \}$$  \hspace{1cm} (70)

From the above table, it can be seen that increased concern for equity, $\nu$ 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.

<table>
<thead>
<tr>
<th>Level of $s$</th>
<th>Utilitarian</th>
<th>Maxi-min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average rate (%)</td>
<td>Marginal rate (%)</td>
</tr>
<tr>
<td>Median</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>Top decile</td>
<td>14</td>
<td>20</td>
</tr>
<tr>
<td>Top percentile</td>
<td>16</td>
<td>17</td>
</tr>
</tbody>
</table>

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

### 3.11 Numerical Results

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

---

6 This part draws from Myles (1955) Chap 2, pp.156-9.
Higher values of \( v \) represent greater concern for equity, with \( v=0 \) representing the utilitarian case.

The individual utility function was the Cobb-Douglas,

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

and the skill distribution is log-normal,

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

With a standard deviation \( \sigma \) (=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, \ v = 0, \ \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, \ v = 1, \ \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, \ v = 1, \ \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)
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,

\[ w = \min \{ U \} \]  

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.

<table>
<thead>
<tr>
<th>Level of s</th>
<th>Average rate (%)</th>
<th>Marginal rate (%)</th>
<th>Average rate (%)</th>
<th>Marginal rate (%)</th>
</tr>
</thead>
<tbody>
<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)

### 3.12 Voting over a Flat Tax

Having identified the properties of the optimal tax structure, we now consider the tax system that emerges from the political process. To do this, we consider people voting over tax schedules that have some degree of redistribution. Because it is difficult to model voting on nonlinear tax schemes given the high dimensionality of the problem, we will restrict attention to a linear tax structure as originally proposed by Romer (1975). We specify the model further with quasi-linear preferences to avoid unnecessary complications and to simplify the analysis of
the voting equilibrium.

Assume, as before, that individuals differ only in their level of skill. We assume that skills are distributed in the population according to a cumulative distribution function $F(s)$ that is known to everyone, with mean skill $\bar{s}$ and median $s_m$. Individuals work and consume. They also vote on a linear tax scheme that pays a lumpsum benefit $b$ to each individual financed by a proportional income tax at rate $t$. The individual utility function has the quasi-linear form

$$u\left( {x, \frac{z}{s}} \right) = x - \frac{1}{2} \left[ \frac{z}{s} \right]^2 , \quad (75)$$

and the individual budget constraint is

$$x = [1 - t]z + b . \quad (76)$$

It is easy to verify that in this simple model the optimal income choice of a consumer with skill level $s$ is

$$z(s) = [1 - t]s^2 . \quad (77)$$

The quasi-linear preferences imply that there is no income effect on labor supply (i.e., $z(s)$ is independent of the lump-sum benefit $b$). This simplifies the expression of the tax distortion and makes the analysis of the voting equilibrium easier. Less surprisingly a higher tax rate induces taxpayers to work less and earn less income.

The lump-sum transfer $b$ is constrained by the government budget balance condition

$$b = tE(z(s)) = t[1 - t]E(s^2) , \quad (78)$$

where $E(\cdot)$ is the mathematical expectation, and we used the optimal income choice to derive the second equality. This constraint says that the lump-sum benefit paid to each individual must be equal to the expected tax payment $tE(z(s))$. This expression is termed the Dupuit-Laffer curve and describes tax revenue as a function of the tax rate. In this simple model the Dupuit-Laffer curve is bell-shaped with a peak at $t = 1/2$ and no tax collected at the ends $t = 0$ and $t = 1$. We can now derive individual preferences over tax schedules by substituting (49) and (50) into (48). After re-arrangement (indirect) utility can be written
\[
v(t, b, s) = b + \frac{1}{2} [1 - t]^2 s^2.
\]  

(79)

Taking the total differential of (52) gives

\[
dv = db - [1 - t]s^2 dt.
\]  

(80)

so that along an indifference curve where \(dv = 0\),

\[
\frac{db}{dt} = [1 - t]s^2.
\]  

(81)

It can be seen from this that for given \(t\), the indifference curve becomes steeper in \((t, b)\) space as \(s\) increases. This monotonicity is a consequence of the single-crossing property of the indifference curves. The single-crossing property is a sufficient condition for the Median Voter Theorem to apply. It follows that there is only one tax policy that can result from majority voting: it is the policy preferred by the median voter (half the voters are poorer than the median and prefer higher tax rates, and the other half are richer and prefer lower tax rates). Letting \(t_m\) be the tax rate preferred by the median voter, then we have \(t_m\) implicitly defined by the solution to the first-order condition for maximizing the median voter’s utility. We differentiate (52) with respect to \(t\), taking into account the government budget constraint (51) to obtain

\[
\frac{\partial v}{\partial t} = [1 - 2t]E(s^2) - [1 - t]s^2.
\]  

(82)

Setting this expression equal to zero for the median skill level \(s_m\) yields the tax rate preferred by the median voter

\[
t_m = \frac{E(s^2) - s_m^2}{2E(s^2) - s_m^2}.
\]  

(83)

or, using the optimal income choice (50),

\[
t_m = \frac{E(z) - z_m}{2E(z) - z_m}.
\]  

(84)
This simple model predicts that the political equilibrium tax rate is determined by the position of the median voter in the income distribution. The greater is income inequality as measured by the distance between median and mean income, the higher the tax rate. If the median voter is relatively worse off, with income well below the mean income, then equilibrium redistribution is large. In practice, the income distribution has a median income below the mean income, so a majority of voters would favor redistribution through proportional income taxation. More general utility functions would also predict that the extent of this redistribution decreases with the elasticity of labor supply.

Empirical Fact in Japan

Table 5  Individual Inhabitants’ and Income Tax

<table>
<thead>
<tr>
<th></th>
<th>Inhabitants’ Tax</th>
<th>Income Tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax Recipient</td>
<td>Municipal governments on 1st January</td>
<td>National government</td>
</tr>
<tr>
<td>Tax Payer</td>
<td>Individual lives in a municipal government</td>
<td>Individual lives in Japan</td>
</tr>
<tr>
<td>Tax Method</td>
<td>With holding</td>
<td>Assessment</td>
</tr>
<tr>
<td>Tax Base</td>
<td>Last year’s income</td>
<td>This year’s income</td>
</tr>
<tr>
<td>Income Deduction</td>
<td>• Basic deduction</td>
<td>$330,000</td>
</tr>
<tr>
<td></td>
<td>• Spouse deduction</td>
<td>$330,000</td>
</tr>
<tr>
<td></td>
<td>• Special spouse deduction</td>
<td>max $330,000</td>
</tr>
<tr>
<td></td>
<td>• Family deduction</td>
<td>$330,000</td>
</tr>
<tr>
<td>Minimum Taxable Income (1998)</td>
<td>A couple with 2 children</td>
<td>Before a special tax deduction $3,031,000</td>
</tr>
<tr>
<td></td>
<td>After special tax deduction $4,373,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Before a special tax deduction $3,616,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>After special tax deduction $4,917,000</td>
<td></td>
</tr>
<tr>
<td>Tax Rate</td>
<td>Tax Rates</td>
<td>Taxable Income</td>
</tr>
<tr>
<td>Taxable Income</td>
<td>Prefectural</td>
<td>Municipal</td>
</tr>
<tr>
<td>Below $2,000,000</td>
<td>2%</td>
<td>3%</td>
</tr>
<tr>
<td>Below $7,000,000</td>
<td>2%</td>
<td>7%</td>
</tr>
<tr>
<td>Above $7,000,000</td>
<td>3%</td>
<td>12%</td>
</tr>
<tr>
<td>Capital Gain Income</td>
<td>Below $30,000,000</td>
<td>40%</td>
</tr>
<tr>
<td>Below $80,000,000</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>Above $80,000,000</td>
<td>3%</td>
<td>6%</td>
</tr>
<tr>
<td>In 1999-2001</td>
<td>Below $40,000,000</td>
<td>20%</td>
</tr>
<tr>
<td>Below $60,000,000</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>Above $60,000,000</td>
<td>2%</td>
<td>5.5%</td>
</tr>
<tr>
<td>Tax Exemption</td>
<td>- Adjustment for Double Taxation</td>
<td>- Adjustment for Double Taxation</td>
</tr>
<tr>
<td></td>
<td>- Dividends</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Foreign Tax</td>
<td></td>
</tr>
<tr>
<td>Tax Revenue (1996)</td>
<td>90,17 trillion yen</td>
<td>189,649 trillion yen</td>
</tr>
</tbody>
</table>
### 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>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>million</td>
<td>million</td>
<td>million</td>
</tr>
<tr>
<td>10</td>
<td>10.5</td>
<td>10.5</td>
<td>10</td>
<td>10</td>
<td>(~3)</td>
<td>10</td>
<td>(~3.3)</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>12</td>
<td>20</td>
<td>20</td>
<td>(~6)</td>
<td>20</td>
<td>(~9)</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>16</td>
<td>30</td>
<td>30</td>
<td>(~1)</td>
<td>30</td>
<td>(~18)</td>
</tr>
<tr>
<td>16</td>
<td>17</td>
<td>20</td>
<td>40</td>
<td>40</td>
<td>(~2)</td>
<td>40</td>
<td>(~30)</td>
</tr>
<tr>
<td>18</td>
<td>21</td>
<td>25</td>
<td>50</td>
<td>50</td>
<td>(2~)</td>
<td>50</td>
<td>(30~)</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></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tax Rate</td>
<td>34</td>
<td>45</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>50</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</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</td>
<td>¥3,821,000</td>
</tr>
</tbody>
</table>

### Table 7  Change in Income Tax Payers

<table>
<thead>
<tr>
<th>Year</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>4,757</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.

---

Lectures on Public Finance Part2_Chap3, 2012 version  P.35 of 43
### Table 8  The 1996 Declared Income Tax Burden Rate by Income Class

<table>
<thead>
<tr>
<th>Income class</th>
<th>Average Income</th>
<th>Average Income Deduction</th>
<th>Average Taxable Income</th>
<th>Average Calculated Tax</th>
<th>Average Tax Deduction</th>
<th>Average Tax Payment</th>
<th>Gross Income Tax Rate ((\frac{\text{①}}{\text{④}}))</th>
<th>Effective Income Tax Rate ((\frac{\text{①}}{\text{④}})/((\text{①} - \text{②})))</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 100</td>
<td>748</td>
<td>546</td>
<td>202</td>
<td>20</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>48</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>88</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>197</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>635</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>4,707</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>759</td>
<td>12.9%</td>
<td>17.7%</td>
</tr>
</tbody>
</table>

Source: *The 1996 Sample Survey of Declared Income Tax (Tax Bureau)*.

### 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)</th>
<th>Total Salary (100 million yen)</th>
<th>Tax paid (100 million yen)</th>
<th>Effective Tax Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Share</td>
<td>Share</td>
<td>Share</td>
<td>Share</td>
</tr>
<tr>
<td>- 100</td>
<td>3,228</td>
<td>7.2</td>
<td>479</td>
<td>1.2</td>
</tr>
<tr>
<td>100 – 200</td>
<td>4,818</td>
<td>10.7</td>
<td>3,419</td>
<td>8.7</td>
</tr>
<tr>
<td>200 – 300</td>
<td>6,818</td>
<td>15.2</td>
<td>6,230</td>
<td>15.9</td>
</tr>
<tr>
<td>300 – 400</td>
<td>7,780</td>
<td>17.3</td>
<td>7,328</td>
<td>18.7</td>
</tr>
<tr>
<td>400 – 500</td>
<td>6,530</td>
<td>14.5</td>
<td>6,244</td>
<td>15.9</td>
</tr>
<tr>
<td>500 – 600</td>
<td>4,964</td>
<td>11.1</td>
<td>4,802</td>
<td>12.3</td>
</tr>
<tr>
<td>600 – 700</td>
<td>3,273</td>
<td>7.3</td>
<td>3,215</td>
<td>8.2</td>
</tr>
<tr>
<td>700 – 800</td>
<td>2,384</td>
<td>5.3</td>
<td>2,372</td>
<td>6.1</td>
</tr>
<tr>
<td>800 – 900</td>
<td>1,604</td>
<td>3.6</td>
<td>1,604</td>
<td>4.1</td>
</tr>
<tr>
<td>900 – 1000</td>
<td>1,004</td>
<td>2.2</td>
<td>1,004</td>
<td>2.6</td>
</tr>
<tr>
<td>1000 – 1500</td>
<td>1,963</td>
<td>4.4</td>
<td>1,963</td>
<td>5.0</td>
</tr>
<tr>
<td>1500 – 2000</td>
<td>378</td>
<td>0.8</td>
<td>378</td>
<td>1.0</td>
</tr>
<tr>
<td>2000 – 2500</td>
<td>87</td>
<td>0.2</td>
<td>87</td>
<td>0.2</td>
</tr>
<tr>
<td>2500+</td>
<td>64</td>
<td>0.1</td>
<td>64</td>
<td>0.2</td>
</tr>
<tr>
<td>Total</td>
<td>44,896</td>
<td>100.0</td>
<td>39,189</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Note: Employees on December 31, 1996.

Source: *The 1996 Tax Bureau Private Sector Income Survey (Tax Bureau)*.

### Table 10  Tax Elasticity

<table>
<thead>
<tr>
<th>Years</th>
<th>Direct Tax (Income Tax)</th>
<th>Indirect Tax</th>
<th>Total Tax Revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual</td>
<td>Corporate</td>
<td>Individual</td>
</tr>
<tr>
<td>1985</td>
<td>1.00</td>
<td>1.31</td>
<td>0.64</td>
</tr>
<tr>
<td>1986</td>
<td>1.89</td>
<td>1.74</td>
<td>1.66</td>
</tr>
<tr>
<td>1987</td>
<td>3.64</td>
<td>2.58</td>
<td>4.78</td>
</tr>
<tr>
<td>1988</td>
<td>2.25</td>
<td>1.57</td>
<td>2.88</td>
</tr>
<tr>
<td>1989</td>
<td>1.66</td>
<td>2.73</td>
<td>0.58</td>
</tr>
<tr>
<td>1990</td>
<td>1.36</td>
<td>2.69</td>
<td>0.16</td>
</tr>
<tr>
<td>1991</td>
<td>0.02</td>
<td>0.06</td>
<td>0.14</td>
</tr>
<tr>
<td>1992</td>
<td>▲0.618</td>
<td>▲0.609</td>
<td>▲0.77</td>
</tr>
<tr>
<td>1993</td>
<td>▲0.414</td>
<td>2.14</td>
<td>▲1.71</td>
</tr>
<tr>
<td>1994</td>
<td>4.75</td>
<td>5.50</td>
<td>8.25</td>
</tr>
<tr>
<td>1995</td>
<td>1.10</td>
<td>▲2.00</td>
<td>6.50</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

   For the utility function \( U = \sum_{i=1}^{A} \frac{X_i^{1-e_i}}{1-1/e_i} - vL \), 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 \( (\partial x_i/\partial y) \) and cross price terms are zero. Derive the optimal tax structure where \( e_i \) 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 \( x = b + [1-t]w\ell \) in the presence of income taxation?

   Consider the budget constraint \( x = b + [1-t]w\ell \). Provide an interpretation of \( b \). How does the average rate of atx change with income? Let utility be given by \( U = x - \ell^2 \). How is the choice of \( \ell \) affected by increases in \( b \) and \( t \)? Explain these effects.

5. [Hindriks and Myles (2006) Chapter 15, Exercises 15.2]
   Assume that a consumer has preferences over consumption and leisure described by \( U = x[1 - \ell] \), where \( x \) is consumption and \( \ell \) is labor. For a given wage rate \( w \), which leads to a higher labor supply: an income tax at constant rate \( t \) or a lump-sum tax \( T \) that raises the same revenue as the income tax?

6. [Hindriks and Myles (2006) Chapter 15, Exercises 15.3]
   Let the utility function be \( U = \log(x) - \ell \). Find the level of labor supply if the wage rate, \( w \), is equal to 10. What is the effect of the introduction of an overtime premium that raises \( w \) to 12 for hours in excess of that worked at the wage of 10?

   Assume that utility is \( U = \log(x) - \log(1 - \ell) \). Calculate the labor supply function. Explain the form of this function by calculating the income and substitution effects of a wage increase.

   Show that a tax function is average-rate progressive (the average rate of tax rises with income) if \( MRT > ART \).

Which is better: a uniform tax on consumption or a uniform tax on income?

10. [Hindriks and Myles (2006) Chapter 15, Exercises 15.8]
Consider the utility function \( U = x - \ell^2 \).

a. For \( U = 10 \), plot the indifference curve with \( \ell \) on the horizontal axis and \( x \) on the vertical axis.

b. Now define \( z = s \ell \). For \( s = 0.5, 1, \) and \( 2 \) plot the indifference curves for \( U = 10 \) with \( z \) on the horizontal axis and \( x \) on the vertical.

c. Plot the indifference curves for \( s = 0.5, 1, \) and \( 2 \) through the point \( x = 20, z = 2 \).

d. Prove that at any point \((x, z)\) the indifference curve of a high-skill consumer is flatter than that of a low-skill.

11. [Hindriks and Myles (2006) Chapter 15, Exercises 15.9]
Consider an economy with two consumers who have skill levels \( s_1 = 1 \) and \( s_2 = 2 \) and utility function \( U = 10x^{1/2} - \ell^2 \). Let the government employ an income tax function that leads to the allocation \( x = 4, z = 5 \) for the consumer of skill \( s = 1 \) and \( x = 9, z = 8 \) for the consumer of skill \( s = 2 \).

a. Show that this allocation satisfies the incentive compatibility constraint that each consumer must prefer his allocation to that of the other.

b. Keeping incomes fixed, consider a transfer of 0.01 units of consumption from the high-skill to the low-skill consumer.

i. Calculate the effect on each consumer’s utility.

ii. Show that the sum of utilities increases.

iii. Show that the incentive compatibility constraint is still satisfied.

iv. Use parts i through iii to prove that the initial allocation is not optimal for a utilitarian social welfare function.

Assume that skill is uniformly distributed between 0 and 1 and that total population size is normalized at 1. If utility is given by \( U = \log(x)\log(1 - \ell) \) and the budget constraint is \( x = b + (1 - t)s \ell \), find the optimal values of \( b \) and \( t \) when zero revenue is to be raised. Is the optimal tax system progressive?

Consider an economy with two consumers of skill levels \( s_1 \) and \( s_2, s_2 > s_1 \). Denote the allocation to the low-skill consumer by \( x_1, z_1 \) and that to the high-skill consumer by \( x_2, z_2 \).

a. For the utility function \( U = u(x) - z/s \) show that incentive compatibility requires that \( z_2 = z_1 + [u(x_2) - u(x_1)] \).
b. For the utilitarian social welfare function \( W = u(x_1) - \frac{z_1}{s_1} + u(x_2) - \frac{z_2}{s_2} \), express \( W \) as a function of \( x_1 \) and \( x_2 \) alone.

c. Assuming \( u(x_h) = \log(x_h) \), derive the optimal values of \( x_1 \) and \( x_2 \) and hence of \( z_1 \) and \( z_2 \).

d. Calculate the marginal rate of substitution for the two consumers at the optimal allocation. Comment on your results.


Suppose two types of consumers with skill levels 10 and 20. There is an equal number of consumers of both types. If the social welfare function is utilitarian and no revenue is to be raised, find the optimal allocation under a nonlinear income tax for the utility function \( U = \log(x) - \ell \). Contrast this to the optimal allocation if skill was publicly observable.

15. [Hindriks and Myles (2006) Chapter 15, Exercises 15.16]

Tax revenue is given by \( R(t) = tB(t) \), where \( t \in [0,1] \) is the tax rate and \( B(t) \) is the tax base. Suppose that the tax elasticity of the tax base is \( \xi = -\frac{\mu}{1-\mu} \) with \( \gamma \in \left[\frac{1}{2}, 1\right] \).

a. What is the revenue-maximizing tax rate?

b. Graph tax revenue as a function of the tax rate both for \( \gamma = 1/2 \) and \( \gamma = 1 \). Discuss the implications of this Dupuit-Laffer curve.


Consider an economy populated by a large number of workers with utility function \( U = x^\alpha [1 - \ell]^{1-\alpha} \), where \( x \) is disposable income, \( \ell \) is the fraction of time worked \( (0 \leq \ell \leq 1) \), and \( \alpha \) is a preference parameter \( (0 < \alpha < 1) \). Each worker’s disposable income depends on his fixed “skill” as represented by wage \( w \) and a tax-transfer scheme \((t, B)\) such that \( x = B + [1-t]wt \), where \( t \in (0,1) \) is the marginal tax rate and \( B > 0 \) is the unconditional benefit payment.

a. Find the optimal labor supply for someone with ability \( w \). Will the high-skill person work more than the low-skill person? Will the high-skill person have higher disposable income than the low-skill person? Show that the condition for job market participation is \( w > \frac{[(1-\alpha)\ell B]}{1-t} \).

b. If tax proceeds are only used to finance the benefit \( B \), what is the government’s budget constraint?

c. Suppose that the mean skill in the population is \( \bar{w} \) and that the lowest skill is a fraction \( \gamma < 1 \) of the mean skill. If the government wants to redistribute all tax proceeds to finance the cash benefit \( B \), what condition should the tax-transfer scheme
(t, B) satisfy?

d. Find the optimal tax rate if the government seeks to maximize the disposable income of the lowest skill worker subject to everyone working.
References

Feenberg,