Chapter Outline

**Designing Broad-Based Taxes: The Economic Objectives** 171

- Ability to Pay: Theoretical Considerations 173
  - Two Preliminary Considerations 173

**Horizontal Equity** 174
- From Horizontal Equity to the Ideal Tax Base 174
- The Three Principles of Tax Design 174
  - People Bear the Tax Burden 174
  - Individuals Sacrifice Utility 174
- The Ideal Tax Base as the Best Surrogate 175
  - Measure of Utility 175

**Haig-Simons Income** 175
- The Sources and Uses of Income 175
  - Sources of Income 175
  - Uses of Income 176
- Real versus Nominal Income 176
- Other Tax Bases 176
- Criticisms of Haig-Simons Income 176
  - A Flawed Surrogate Measure of Utility? 176
  - A Better Alternative to Haig-Simons Income? 177
- Consumption or Expenditures as the Preferred Alternative 178
  - Consumption versus Income Taxes: An Example 179
  - Haig-Simons Income versus Expenditures: Musgrave’s Perspective 180
- Horizontal Equity and the Interpersonal Equity Conditions 180

**Vertical Equity** 182
- Progressive, Proportional, and Regressive Taxes 182
- Vertical Equity and the Interpersonal Equity Conditions 183

Sacrifice Principles of Vertical Equity 183
- Minimize Aggregate Sacrifice 183
- Equal Sacrifice 184
- Young’s Prescription for Vertical Equity 184
- Young’s Six Principles of Taxation 184
- Proportional Taxation 186
- Progressive Taxation 186
- Vertical Equity in the United States 186

**Reflections on the Haig-Simons Criterion in Practice:** 187

- The Federal Personal Income Tax 187
  - Personal Income 187
  - Capital Gains 188
- The Taxation of Personal Income: The Tax Loopholes 188
- Tax Loopholes, Tax Capitalization, and Horizontal Equity 189
- Tax Loopholes, Vertical Equity, and Inefficiency 190
- The Taxation of Capital Gains: Inflation Bias and Realization 191

**The Inflationary Bias against Income from Capital** 192

**Taxing Realized Gains: Auerbach’s Retrospective Taxation Proposal** 193
- A Two-Period Example 193
  - Option 1: Sell and Invest Risk-Free Asset 194
  - Option 2: Hold for Two Periods 194
- The Vickrey Proposal 194
- The Auerbach Proposal 194
- Capital Gains Taxation: A Postscript 195

**The Taxation of Human Capital** 196
- Summary 197

References 198

---

**DESIGNING BROAD-BASED TAXES: THE ECONOMIC OBJECTIVES**

Economists have proposed five economic objectives that governments should strive for in designing broad-based taxes:

1. Ease of administration and taxpayer compliance
2. Minimize deadweight loss
3. Promote long-run economic growth
4. Maintain flexibility
5. Honor society’s norms of fairness or equity

The first objective takes precedence in the sense that if a tax does not meet both parts of this objective, it simply will not be used. Ease of administration refers to the ability of a department of revenue to collect the taxes due easily and economically, at a small fraction of the cost of the revenues raised. Ease of taxpayer compliance refers to the taxpayers’
properties of taxes: the second to static efficiency and the third to dynamic efficiency. Regarding static efficiency, we saw in Chapter 2 that buyers and sellers must face the same market prices to achieve the pareto-optimal conditions. Taxes distort markets by driving a wedge between the prices faced by buyers and sellers, thereby generating deadweight efficiency losses. The goal of tax design is to minimize the deadweight efficiency losses for any given amount of revenues collected. The dynamic efficiency problem is that taxes may also reduce incentives to save and invest, to the detriment of long-run economic growth. The goal is to maintain incentives for saving and investment to the fullest extent possible. A related problem is to ensure that tax policy keeps the economy as close as possible to the Golden Rule of Accumulation, the capital/labor ratio that maximizes consumption per person for any given rate of growth.

The flexibility objective is usually associated with the macroeconomic stabilization goal of smoothing the business cycle to keep the economy close to its production—possibilities frontier. Taxes are the main instrument of fiscal policy. As such, they must be flexible enough to be adjusted up or down as needed to smooth the business cycle.

The final objective calling for equity in taxation is a reminder that taxes must be consistent with society’s norms in its quest for end results and process equity.

This chapter addresses only the final objective of achieving equity in taxation, for two reasons. One is that the pursuit of equity is a fundamental problem for a market economy that even a first-best perspective cannot assume away. The second is that the other tax design objectives are either less compelling or inapplicable in a first-best environment. The static and dynamic efficiency objectives, although very important to the design of broad-based taxes, are necessarily the second-best objectives and will be considered in Part III of the text. So, too will the first objective. Ease of administration is generally not a serious issue for any of the broad-based taxes in the United States. All the major U.S. taxes are collected fairly easily and at very low cost. In contrast, taxpayer compliance is an important issue for some of the taxes, with the most serious problems resulting from private information. Taxpayers who are unwilling to pay their taxes may be able to hide information about themselves from the tax authorities. Private information is inherently a second-best issue, however. Finally, the macro flexibility issues are beyond the scope of this chapter.

We have seen in the previous chapters that first-best public sector theory does not provide much guidance to policymakers charged with designing broad-based taxes that the public will view as fair. The prescription for distributive equity in taxation (and transfer payments) is entirely contained within the interpersonal equity conditions for a social welfare maximum, yet these conditions beg the prior question of what the social welfare function should be. We also considered the benefits-received principle of taxation. It appears to have great appeal as a principle of tax equity in the United States, but its role in first-best theory is strictly as an efficiency principle. In any event, the benefits-received principle can only be narrowly applied to certain resource-using expenditures whose pattern of benefits is clearly defined. It cannot serve as the basis for designing broad-based taxes.

As it happens, tax practitioners have not been much bothered by the difficulties surrounding the social welfare function or the limitations of the benefits-received principle. Attempts to design fair broad-based taxes are almost always grounded in another principle of tax equity called the ability-to-pay principle, which dates from the beginnings of modern economics, having first been proposed by Adam Smith in the late 1700s and then further developed by John Stuart Mill in the early 1800s. The only established principle of tax equity before Smith and Mill was the benefits-received principle, which had originated in the fourteenth and fifteenth centuries under feudalism. The feudal lords would pay a tribute (tax) to the Crown in return for protection from foreign enemies. Smith and Mill recognized the need for another principle of tax equity for general taxes that were not so clearly tied to particular benefits received by the taxpayers.

The remainder of this chapter focuses on the ability-to-pay principle, indicating how to proceed from the principle to the design of broad-based taxes. The U.S. federal personal income tax will serve as the primary application throughout the chapter. Of all the broad-based taxes, it is this tax that is most closely grounded in the ability-to-pay principle.

The Smith—Mill ability-to-pay principle and the Bergson—Samuelson interpersonal equity conditions of first-best theory are also compared and contrasted. The older ability-to-pay principle would appear to bear a close kinship to the newer interpersonal equity conditions. The taxes and transfers implied by the interpersonal equity conditions surely depend on individuals’ economic well-being, that is, on their ability to pay. Even so, the two principles are not as closely related as one might think.
They derive from fundamentally different views of taxation and, as such, they do not necessarily imply that the government should collect the same tax revenues from individuals or even use the same taxes.

**ABILITY TO PAY: THEORETICAL CONSIDERATIONS**

Smith and Mill recognized the limitations of the benefits-received principle as public expenditures became more varied and their benefits more diffused throughout the population. They reacted by introducing the concept of taxes as a necessary evil, a sacrifice that individuals have to make for the common good to support desired public expenditures. Given their perspective, they saw the fundamental question of tax equity as being one of how the government should ask people to sacrifice for the commonwealth, the common good. Their answer was that people should be asked to sacrifice in accordance with their ability to pay. In addition, the pattern of sacrifice should honor the two principles of horizontal equity and vertical equity. Horizontal equity says that equals should be treated equally. Two persons judged to have equal ability to pay should bear the same tax burden. Vertical equity allows for the unequal treatment of the unequals; that is, two persons with unequal abilities to pay can properly be asked to bear unequal tax burdens. This new Smith—Mill ability-to-pay principle was a sacrifice principle, pure and simple. Taxpayers should not expect a quid pro quo from general or broad-based taxes, in direct contrast to taxes paid according to benefits received.1 Ability-to-pay principle was viewed as a default principle, to be used whenever the narrower benefits-received principle could not be applied.

The ability-to-pay principle quickly gained virtual unanimous acceptance as the appropriate equity norm for broad-based tax design. Its intellectual origins were familiar, dating from Aristotle and perhaps even further back, but a huge gap remained in applying the principles of horizontal and vertical equities to the actual design of a tax.

The requirements of horizontal and vertical equities beg two important and difficult questions. The first is the definition of equality: In what sense are two persons equal or unequal for the purposes of taxation? Both principles require an answer to this question. The second is the fundamental question in applying vertical equity: How unequally should unequal be treated under the tax laws? This is a part of the broader question of end-results equity, or distributive justice, related to the distribution of income.

The quest for horizontal equity in taxation has typically been associated with the goal of defining the ideal tax base. A person’s tax liability is computed by multiplying the tax rate and the tax base. Therefore, two persons with the same value of the tax base necessarily pay the same tax and are treated equally in terms of taxation. The ideal tax base applies to vertical equity as well, since it defines the extent to which people are judged to be unequal for purposes of taxation.

Once the ideal tax base has been determined, the quest for vertical equity is then concerned with the design of the tax structure, which has two main components. First is the pattern of rates to be applied to different levels of the tax base. The second is the pattern of allowable exemptions, deductions, credits, and other adjustments to the tax base in computing the tax liability. These adjustments are justified in terms of promoting certain social goals that the government deems important. Two examples under the federal personal income tax are the personal exemptions that prevent the poor from having to pay taxes, which are permitted in the name of equity, and the deduction of interest payments on mortgages, which are permitted to encourage homeownership.

**Two Preliminary Considerations**

Two points should be noted before turning to the ideal tax base and tax structure. The first point is the fundamental difference in perspective between the Bergson—Samuelson interpersonal equity conditions of first-best theory and the Smith—Mill ability-to-pay principle. The taxes called for by the interpersonal equity conditions are inherently viewed as a good in and of themselves, since the interpersonal equity conditions are one of the two sets of first-order conditions necessary for maximizing social welfare. They promote social welfare by helping society reach the best distribution of income or utility on the utility—possibilities frontier. Taxes are not at all the necessary evil that Smith and Mill saw them to be. This sharp difference in perspective helps to explain why these two theoretical principles do not necessarily imply the same taxes, even if the taxes required by the interpersonal equity conditions are levied on the basis of ability to pay. We will return to this point after developing the implications of the ability-to-pay principle for the design of taxes.

The second point is that the ability-to-pay principle can properly be considered as a part of first-best theory. Ability-to-pay as a sacrifice principle relates specifically to the goal of distributive justice. Second-best tax theory is concerned, first and foremost, with the efficiency costs of distorting taxation. In a many-person second-best environment, efficiency considerations must be tempered by the equity implications of alternative distorting taxes, so that second-best theory has an interest in ability-to-pay principles. But the principles themselves have nothing whatsoever to do with the questions of efficiency. Hence, ability-to-pay principles

---

1. Smith (1904), Mill (1921). For an excellent history of the development of ability-to-pay principles, see Musgrave (1959), Chapter 5.
are analyzed most conveniently in a first-best environment, one in which efficiency and equity issues are separable. This is precisely what happened in the professional literature.

Careful distinctions between first-best and second-best analysis are a fairly recent phenomenon, but it is clear that early ability-to-pay theorists were implicitly assuming a first-best environment. We have two clues on this. The first is that the ability-to-pay literature generally ignores efficiency considerations altogether. This would be impossible in a second-best framework. The second is that ability-to-pay theory has traditionally equated tax payments and tax burdens. This, too, implies a first-best environment, for reasons that can only be sketched at this point in the text.

Tax incidence theory, the subject matter of Chapter 16, distinguishes between the burden of a tax (who sacrifices as a result of the tax) and the impact of a tax (who physically pays the tax—writes the check—to the government). We were careful earlier when defining horizontal and vertical equities to refer to “tax burdens.” This is not always done. The two principles are often defined in terms of “tax payments” as follows:

- **Horizontal equity**: Equals should pay equal taxes.
- **Vertical equity**: Unequals should pay unequal taxes.

The difference is significant. Tax incidence theory shows that under certain conditions in a first-best policy environment, lump-sum tax payments are an appropriate measure of individual welfare losses, or burdens, using standard willingness-to-pay criteria such as Hicks’ Compensating or Equivalent Variations. With distorting taxes, however, the tax payments are never entirely accurate measures of welfare loss. These points are fairly subtle and will be discussed in detail in Chapters 13 and 16. What matters here in terms of the ability-to-pay principles is that equal tax payments may yield unequal burdens with distorting taxes simply because of the distortions. Alternatively, unequal tax payments may entail equal burdens. Hence, once the possibility of distorting taxation is recognized, horizontal and vertical equities must be more broadly defined in terms of tax burdens, as we have done. Conversely, equating tax payments and tax burdens must imply both a first-best policy environment and lump-sum taxation.

We will adopt a first-best framework and equate tax payments and tax burdens to focus strictly on the equity issues involved with the ability-to-pay principles. This is at best an uneasy convenience, however. The problem is that the ability-to-pay principles lead to choices of broad-based taxes that are almost certainly not lump sum, so that it is impossible to ignore distortions entirely. In particular, the federal personal income tax contains a number of second-best distortions whose equity implications can only be understood in terms of the broader tax-burden interpretation of horizontal and vertical equities. Thus, we will occasionally stray from the first-best assumptions.

**HORIZONTAL EQUITY**

From Horizontal Equity to the Ideal Tax Base

Mainstream public sector economists do not agree on which tax base best satisfies the principle of horizontal equity. They do agree, however, on the proper way to think about what the ideal tax base should be. The line of reasoning from horizontal equity to the ideal tax base always relies on the same three principles of tax design. The disagreement occurs in applying the third principle, which describes the final step to the tax base.

**The Three Principles of Tax Design**

**People Bear the Tax Burden**

The first principle of tax design is that people ultimately bear the burden of any tax no matter what is actually taxed. For example, corporate income taxes and sales taxes are levied on business firms in the United States, but the fact that a business firm pays $X million in taxes is of little consequence. The interesting questions in terms of tax equity are which people finally bear the burden of these taxes. Is some or the entire burden “passed forward” to the consumers of the final product through higher prices, “passed back” to the labors employed by the firm through lower wages, borne by the stockholders of the firm, or borne by third parties not directly associated with the firm? Social well-being is directly related to individuals’ utility functions, not to production relationships, and any tax eventually burdens people in their roles as consumers or as suppliers of factors, or both.

**Individuals Sacrifice Utility**

The second principle of tax design is that individuals ultimately sacrifice utility when they pay general taxes, so that the ideal tax base would be individual utility levels. In 1976, Martin Feldstein clarified what horizontal equity must mean to mainstream, neoclassical economists.

Feldstein’s Horizontal Equity Principle: Two people with the same utility before tax must have the same utility after tax.

This is the only sensible economic interpretation of equal treatment of equals under a sacrifice principle of taxation.

Feldstein also proposed a minimum condition for the unequal treatment of unequals—no reversals—that has also gained universal acceptance among neoclassical economists.
Feldstein’s Vertical Equity Principle (No Reversals): *If person i has greater utility than another person j before tax, then person i must have greater utility than person j after tax*.

Feldstein’s two principles can only be guaranteed if utility is the tax base.

**The Ideal Tax Base as the Best Surrogate Measure of Utility**

Taxing utility is impossible, of course, but it still serves as a goal to strive for. Therefore, in lieu of taxing utility, the third principle of tax design is that the tax base should be the best practical surrogate measure of utility. Under this “ideal” tax base, the best surrogate for utility, two persons with equal values of the tax base are equals and should pay the same tax. This is as close as the tax practitioner can come to Feldstein’s principle of equal utility before tax: equal utility after tax in the quest for horizontal equity.

Mainstream economists agree on the three principles, but they have not reached a consensus on what constitutes the best surrogate measure of utility. The two main contenders are income and consumption.

**Haig—Simons Income**

Neither Smith nor Mill was able to produce a convincing argument for an ideal tax base from their ability-to-pay principles. The first-proposed tax base that caught on appeared over 100 years later, in the 1920s and the 1930s. Robert Haig of Columbia and Herbert Simons of Chicago, following the line of reasoning above, independently concluded that a certain broad-based measure of income was the ideal tax base (Simons, 1938; Haig, 1921). Their proposal was almost universally adopted, and “Haig-Simons income” remained essentially unchallenged among mainstream economists as the best surrogate measure of utility until the 1960s, when consumption began to gain favor as a better surrogate measure. The majority of mainstream economists today may view consumption as the better choice.

Haig and Simons argued that purchasing power is the best surrogate measure of utility. This led them to propose income defined as the increase in purchasing power during the year as the ideal tax base for a tax levied annually. Using standard national income accounting terminology, Haig—Simons income can be defined as:

Haig—Simons income = consumption + the increase in net worth.

Consumption is the additional purchasing power actually taken, and the increase in net worth is additional potential purchasing power that has been deferred for future consumption. Net worth can be increased either by new saving or by an increase in the value of the individual’s assets existing at the beginning of the year, the individual’s capital gains. Therefore,

- Haig—Simons income = consumption + saving + capital gains
- or
- Haig—Simons income = personal income + capital gains

Haig—Simons income is also called the **accretion standard** or, more commonly, the **comprehensive tax base**, a label so widely used now that it is often just referred to by the initials CTB.

Having determined that Haig—Simons income is the best surrogate measure of utility, horizontal equity is then defined as follows:

**Horizontal equity**: Two persons with identical amounts of Haig—Simons income are equals and should pay the same tax. Similarly, two persons with different amounts of Haig—Simons income are unequals and should pay different taxes by the principle of vertical equity. The difference in their taxes depends on the tax structure applied to Haig—Simons income.

**The Sources and Uses of Income**

All components of Haig—Simons income are equivalent in terms of increasing purchasing power, so that the sources of income should not affect the amount of tax paid. The uses of the income are also irrelevant to the tax payment. Therefore, distinctions of the following kind should not matter in computing a person’s tax liability, although they happen to matter under the U.S. federal personal income tax (violations of the Haig-Simons standard under the federal personal income tax are noted in brackets).

**Sources of Income**

1. Whether income is derived from personal income or capital gains. (Capital gains are taxed at a substantially lower rate.)
2. Whether personal income is earned (wages, rents, etc.) or unearned (transfer payments). (Many transfer payments are untaxed, such as public assistance.)

---

2. Feldstein (1976). Feldstein’s no-reversals principle is more than an equity principle. It also has important efficiency implications in a second-best world of imperfect information in which the government might not know how well off certain people are. Some people would have a powerful incentive to hide private information about themselves, if the tax laws permitted reversals of utility. We will return to this point in Chapter 15 when analyzing optimal second-best taxes.

3. Notice that the Haig—Simons definition uses personal income rather than disposable income because the former includes personal income taxes, which are originally part of the tax base.

4. The U.S. Internal Revenue Service refers to the tax as the individual income tax. Economists, however, typically refer to taxes levied on individuals as personal taxes and we adopt the economists’ convention in this text.
3. Whether income is received in cash or in kind. (Many fringe benefits received by employees are untaxed, such as employer contributions to pensions and insurance.)

4. Whether earned income derives from labor, capital, or land. (Interest income on many forms of saving for retirement is exempt from income tax, such as the interest on Individual Retirement Accounts (IRAs).)

**Uses of Income**

1. Whether income is consumed or saved. Both consumption and saving increase utility. In terms of tax policy, the only relevant consideration is the increase in purchasing power, whether realized currently as consumption or postponed through saving. (Income used to purchase IRAs and some other retirement accounts is deductible from income in computing taxable income.)

2. Within capital gains, whether a gain is realized by selling an asset or simply accrues in value without a sale. Allowing gains to accrue is merely one particular form of saving. Also, capital losses should be fully offset against other income. (Capital gains are taxed only when realized, and there is only partial offset of capital losses.)

3. Consumption choices are also irrelevant, since all consumption decisions are viewed as voluntary and thereby utility increasing. These include contributions to private charities and tax payments to other governments to pay for the services they offer. (The following expenditures are deductible from income in computing taxable income (sometimes above some minimum level): medical expenses, contributions to charities and other nonprofit institutions such as colleges, state and local income and property taxes, and interest on a first mortgage.)

The only legitimate deduction from Haig–Simons income is the expenditures necessary for earning income in the first place, so-called *business expenses*. Presumably income used in this manner does not represent an increase in utility-enhancing purchasing power.5

**Real versus Nominal Income**

Haig–Simons income should be indexed for inflation so that inflation alone does not affect a taxpayer’s real tax liability. Real income, not nominal income, is the better surrogate measure of the increase in purchasing power during the year. This point is important for an income tax since it taxes income from capital, which can differ greatly in real and nominal terms. Indexing for inflation matters for all sources of income when a tax uses a set of graduated rates that increase with income, as the federal personal income tax does. (The tax rates varied in seven steps from 10% to 39.6% in 2013). Inflation itself can move a taxpayer into a higher tax bracket and increase the real tax liability. (Only some components of the personal income tax, such as the personal exemptions and the income defining the tax rate brackets, are indexed for inflation.)

**Other Tax Bases**

A final point is that *all* tax bases other than Haig–Simons income are inappropriate because they are not the best surrogate measures of utility. These include: all broad-based taxes such as sales taxes, gift and estate (inheritance) taxes, and value-added taxes; selective excise taxes (except when required by the benefits-received principle); and taxes on specific sources of income, such as the payroll (Social Security) tax and the corporation income tax. Also inappropriate is taxing wealth in any form, such as local property taxes. The increase in purchasing power during the year, not accumulated purchasing power, is the appropriate annual tax base. The flaw with all these other taxes is that they cannot guarantee that two persons with the same Haig–Simons income before tax bear the same tax burden as required for horizontal equity. In fact, equals in terms of Haig–Simons income are very likely to be treated unequally under these other taxes.

**Criticisms of Haig–Simons Income**

Although Haig–Simons income is a reasonable choice for a tax base under the ability-to-pay principle, it could not be expected to gain unanimous acceptance among economists and policymakers, and it has not. Haig–Simons income is vulnerable to both negative and positive attacks. The negative attack is that Haig–Simons income may be a terrible surrogate measure of utility, in which case it loses its appeal as the ideal tax base. The positive attack is simply the belief that there is a better alternative to Haig–Simons income as the ideal tax base. The increasing support among neoclassical economists for consumption or expenditures as the ideal tax base is an argument of this kind. Finally, economists who do not accept the neoclassical perspective are likely to believe that some tax base other than Haig–Simons income is the better alternative.

**A Flawed Surrogate Measure of Utility?**

The negative view that Haig–Simons income may be a poor surrogate measure of utility is worth some discussion because the same argument can be applied to all proposed
tax bases under the ability-to-pay principle. Haig—Simons income does not necessarily suffer relative to other tax bases on these matters. We will simply use it to illustrate the nature of these attacks.

Haig—Simons income is a perfect surrogate measure of utility if people have the same tastes, abilities, and opportunities; otherwise, it may be a very poor surrogate. This is easily seen by means of the simple labor-leisure model in which people exchange hours of leisure for income at a constant hourly wage, \( w \). The budget constraint is

\[
Y = w(24 - \text{leisure})
\]

where \( Y \) is income, \( w \) is the hourly wage, and there are 24 h in the day. Labor is the only source of income.

The two panels in Fig. 11.1 illustrate the difficulties with Haig—Simons income (wage income here) when tastes and opportunities differ. Tastes differ in the left-hand panel. One person is a leisure lover with indifference curves given by \( I_{LL} \). The other person is a work lover (relatively speaking) with indifference curves given by \( I_{WL} \). They face the same wage rate, \( w \), the slope of the budget line. The diagram is meant to indicate that they have the same utility before tax because they reach the same numbered indifference curve, \( I^2 \). Therefore, they should have the same utility after tax by the principle of horizontal equity. But they have different incomes, \( Y_{LL} \) and \( Y_{WL} \), so that they would pay different taxes with Haig—Simons income as the tax base. Consequently, their after-tax utilities may well differ, in violation of horizontal equity.

Opportunities differ in the right-hand panel. The two persons, 1 and 2, have the same tastes but face different wages, \( w_1 \) (the steeper slope) and \( w_2 \). The person facing the higher wage, \( w_1 \), is assumed to take all the additional purchasing power as increased leisure, to sharpen the point about income as a surrogate measure of utility. Person 1 is clearly better off, but they both earn the same income and therefore pay the same tax. Unequals are treated equally, in possible violation of both horizontal and vertical equities.

The failure of Haig—Simons income as a surrogate measure of utility in these examples is that it captures only one of the two variables that confer utility. The narrowness of income would not matter if the two persons were identical in every respect. It would then be a perfect surrogate for utility. These points are not peculiar to (Haig—Simons) income; they apply as well to anything chosen as the tax base. Income, consumption, or any component of income or consumption serves as a perfect surrogate for utility when people are identical in every respect, provided it is something purchased or earned by everyone (as opposed to an either-or choice of, say, a house or an apartment, which otherwise identical people may choose with indifference). Conversely, any one item that generates utility can be wide of the mark as a utility surrogate when tastes, abilities, and/or opportunities differ, because then all utility-generating items may matter in comparing utility.

A Better Alternative to Haig—Simons Income?

Is there a better alternative to Haig—Simons income as the ideal tax base for broad-based taxes? Many economists would say that there is.

To begin with, nonmainstream economists would not necessarily accept the three principles of tax design above as the path to the ideal tax base. Marxist economists, for example, would surely opt for differential treatment of wage and profit income for reasons that have nothing to do with surrogate measures of utility. As another example, Nicholas Kaldor is credited with the first serious proposal for a consumption or expenditures tax. He favored consumption not because of its relation with individual utility but from a broader social perspective. Kaldor agreed that consumption and saving are both self-serving choices by individuals designed to increase their
utility, either now or in the future, but he argued that society can meaningfully distinguish between the two, as follows. When individuals consume, they use up scarce resources for their own personal satisfaction, sacrificing others’ well-being. In contrast, when individuals save, they provide funds for investment that leads to a more productive economy, to the potential future benefit of all citizens. Therefore, Kaldor (1955) argued that society can properly discriminate against consumption in taxation even if taxes are based on a sacrifice principle, providing sacrifice is viewed from a social rather than an individual perspective.

Consumption or Expenditures as the Preferred Alternative

The growing support among neoclassical economists for consumption or expenditures as the ideal tax base is in part based on Kaldor’s argument. The only twist is that Kaldor’s argument is seen today as a dynamic efficiency argument, not an equity argument. Simple, stylized, overlapping generations (OLG) models with perfect foresight that tracks the economy out for 100 periods and more find that replacing an income tax with a consumption tax leads to huge steady-state increases in output per person. Some models report increases on the order of 10–20%. The increased output results from the increase in saving, investment, and productivity under the consumption tax, exactly as Kaldor argued. This is seen as a powerful efficiency argument in favor of a consumption tax.

Many neoclassical economists add to this efficiency argument an equity argument that follows the standard three-step argument to an ideal tax base. They accept Feldstein’s principle of horizontal equity—equal utility before tax, equal utility after tax—and the notion that the ideal tax base is the best surrogate measure of utility. But, they part company with the traditional Haig–Simons conclusion because they believe that the proponents of Haig–Simons income have the time frame wrong.

The break with the traditional view began in the 1960s following the development of Friedman’s Permanent Income Hypothesis and Modigliani–Brumberg’s life-cycle hypothesis (LCH), which themselves broke from the traditional Keynesian view of the consumption decision. The new theories viewed consumers as determining their consumption decisions over a longer period of time than a single year, indeed, over an entire lifetime in the case of the LCH.

The newer mainstream view of the ability-to-pay principle was that taxation should also be viewed in a lifetime context. Haig–Simons income is flawed as the ideal tax base because it relates only to a single year. Feldstein’s equal utility before tax/equal utility after tax is the correct principle, but it should be applied to lifetime utility, appropriately discounted to present value: Two persons with equal present value of lifetime utility before tax should have equal present value of lifetime utility after tax. Therefore, the ideal tax base is the best surrogate measure of (discounted) lifetime utility.

The lifetime perspective argues for consumption, not income, as the ideal tax base by the following line of reasoning. The act of consumption is most closely related to the generation of utility. The Haig–Simons proponents have to think in terms of purchasing power because they adopt an annual perspective in which some purchasing power can be saved for future consumption. This is unnecessary in a lifetime perspective, however, because all income is eventually consumed (counting bequests to heirs as the final act of consumption). People receive income over their lifetimes in three forms: labor market earnings, inheritance, and other transfers from individuals and government. They eventually consume all their income (again, counting the final bequest) such that the lifetime budget constraint holds: The present value of lifetime income equals the present value of lifetime consumption.

From a lifetime perspective, therefore, the best surrogate for the present value of lifetime utility is the present value of lifetime consumption. Consequently, horizontal equity requires that two persons with identical present value of lifetime consumption before tax should have the same present value of lifetime consumption after tax. If taxes were levied on a lifetime basis, it would not matter whether consumption or income was the tax base, because the present value of lifetime consumption and income are equal. But taxes are levied on an annual basis, which matters. Only an annual consumption tax can guarantee that two persons with the same present value of lifetime consumption before tax have the same present value of lifetime consumption after tax.

An annual income tax breaks the equality between lifetime (discounted) consumption before and after tax because it effectively taxes saving twice. The income out of which the saving occurs is taxed, and any returns to the saving are also taxed. In other words, the pattern of consumption and saving matters in determining after-tax lifetime (discounted) consumption under an annual income tax, but not under an annual consumption tax. The following simple example illustrates this point.

---

6. Income from capital is not a source of lifetime income, at least not in an expected present value sense. Income from capital is expected to grow at the same rate of return, r, that is used as the discount rate to compute the present value of income. Therefore, any savings out of three sources of lifetime income cannot grow in expected present-value terms. Saving only changes the timing of consumption, not the overall present value of consumption, from a lifetime perspective.

7. In fact, most people lead virtually self-contained economic lives. The vast majority of people inherit very little wealth, which is the same as saying that most people bequeath very little wealth to their heirs.
Consumption versus Income Taxes: An Example

Suppose that two persons each live for two periods and earn a fixed amount of income Y in each period. Person 1 consumes the entire amount of income each period. Person 2 saves all of the first-period income and consumes everything in the second period. The savings earn a rate of interest, r, the same rate that they use to discount their second-period income and consumption to present value.

The present value of lifetime consumption before tax, which is equal for both people, is:

\[ \text{PV}_{\text{Consumption}} = \frac{Y + Y(1 + r)}{(1 + r)} \]

Under an annual consumption tax levied at rate \( t_c \), the present value of lifetime taxes is the same for both of them: \( \text{PV}_{\text{taxes}} = t_c[Y + Y(1 + r)] \). The only difference is that person 1 pays the tax in two installments and person 2 pays the tax all at once in the second period. Their present values of consumption are the same after tax, as required for horizontal equity.

The present value of taxes differs under an income tax at rate \( t_i \), however, as illustrated by the bottom half of Table 11.1. Notice first, that the discount rate changes from \((1 + r)\) to \((1 + r(1 - t_i))\) under an income tax because interest income is taxed. The double taxation of saving occurs because the income of person 2 is taxed in the first period, so that only \( Y(1 - t_i) \) is available as saving for second-period consumption, and then the interest on the saving is taxed again (assumed to be taxed in the second period here). The taxing of the interest income is what drives a wedge between the present value of taxes for the two persons. Horizontal equity is thus violated under an annual income tax. The two persons have equal present value of consumption before tax but unequal present value of consumption after tax.

The simple example also illustrates two ways to make an income tax equivalent to a consumption tax. One possibility is to allow taxpayers to deduct saving from income in computing their taxable income. This is an expenditures tax, which would be levied exactly as the personal income tax but with a deduction allowed for saving in computing taxable income. Since income is taxed only if consumed, an expenditures tax is the same as a consumption tax. In terms of the bottom half of Table 11.1, the deduction of saving removes the first-period tax from person 2 and also removes the tax on the interest income until it is consumed. With accumulating interest untaxed until consumed, the relevant discount factor reverts to \((1 + r)\), and the income tax with the savings deduction is fully equivalent to the consumption tax (assuming \( t_i = t_c \)).

The second possibility is to remove the double taxation of saving by allowing the taxpayer to deduct all interest income in computing taxable income (in general, any returns to saving/income from capital, whatever its form). This deduction also causes the discount rate to be \((1 + r)\) and removes the second tax term in second period for person 2. The income tax and consumption tax are once again equivalent.

Note, finally, that an expenditures tax is equivalent to a tax on wage income in this simple example because it is an income tax in which all income from capital is deductible. An expenditures tax and a wage tax are not equivalent in actual economies, however. The difference is that a wage tax is paid only during the working years, whereas an expenditures tax is paid in all years of life, including the retirement years. Neoclassical OLG models show that switching from a wage tax to an expenditures tax increases saving and investment because it hits the retired elderly particularly hard. They paid the wage tax while working and now they have to pay a tax on the consumption during retirement that they are financing from their accumulated savings while working. They also have the highest marginal propensity to consume of all the cohorts. The equivalent taxes in an OLG framework are an expenditures tax and a wage tax that includes a one-time capital levy on the retired elderly generation. The capital accumulated at the time of retirement equals the expected present value of

<table>
<thead>
<tr>
<th>Table 11.1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Person 1</td>
</tr>
<tr>
<td>Person 2</td>
</tr>
<tr>
<td>Tax payments (income tax)</td>
</tr>
<tr>
<td>Period 1</td>
</tr>
<tr>
<td>Period 2</td>
</tr>
</tbody>
</table>
consumption until death, counting any bequest as a final act of consumption.

**The Tax Reform Act of 1986: Income Taxation versus Expenditures Taxation**

The Tax Reform Act of 1986 (TRA86) was the largest single reform of the federal personal income tax ever undertaken, and the last reform of the tax base of any consequence. It made significant changes in the definition of taxable income and in the graduated rate structure applied to taxable income. The Reagan administration considered the possibility of replacing the income tax with an expenditures tax when preparing its initial proposal to Congress. The tax at that time was a mixture of the two kinds of taxes: essentially an income tax but with many features of an expenditures tax. The most important expenditures tax features were the treatment of various forms of pension savings such as IRAs and contributions to employer-sponsored pension plans. Contributions to these accounts and plans are deductible from income when made, and the accrued interest income until retirement is also excluded from taxable income. The pension incomes are taxed when received during retirement. This is exactly how savings of all kinds would be treated under an expenditures tax (provided that the pension income is consumed).

The administration decided to stay with the income tax, in large part because of the administrative headaches involved in switching from an income to an expenditures tax. A particular sticking point was what to do about the elderly. They had already been double-taxed on their nonpension forms of saving. Under the income tax, they are not taxed again when they draw down their savings for consumption during retirement. If an expenditures tax were substituted for an income tax, the elderly would be taxed a third time as they consumed their savings. In truth, the large dynamic efficiency gains of switching from an income to an expenditures tax in an OLG framework come at an enormous cost to one group, the elderly, at the time of the switch. Burdening the elderly in this way was naturally considered grossly unfair, yet it was not clear how to protect the elderly (and near-elderly) during the changeover.

---

8. The expenditures tax treatment of pension savings was retained, however, and still exists today.


**Haig–Simons Income versus Expenditures: Musgrave’s Perspective**

Richard Musgrave believes that economists should call a halt to the income tax versus expenditures tax debate regarding horizontal equity. In his view, either Haig–Simons income or expenditures is an acceptable tax base. Neither one is a perfect surrogate measure of utility, but nothing else is either; and continuing to debate, which is the better utility surrogate, is pointless. Musgrave believes that vertical equity is far more important than horizontal equity in any event. Distributive justice is less affected by the choice of Haig–Simons income or expenditures as the tax base than by the tax structure applied to either.

According to Musgrave, the most useful way to interpret the call for horizontal equity in taxation is in a legalistic sense, the same way it is applied in other economic contexts. Equal treatment of equals should simply mean that the tax laws must never discriminate against people in inappropriate ways, such as on the basis of sex, race, or religion. Both Haig–Simons income and expenditures are admissible tax bases by this test. Therefore, Musgrave’s position is that the federal government should simply choose one of them as the tax base and then worry about the appropriate tax structure (Musgrave, 1990).

**Horizontal Equity and the Interpersonal Equity Conditions**

Neoclassical economists would presumably want a tax designed in accordance with ability-to-pay principles to bear a fairly close relationship with the interpersonal equity conditions, since the interpersonal equity conditions are the ultimate guidelines for end-results equity in first-best public sector theory. Unfortunately, the ability-to-pay principle and the interpersonal equity conditions do not lead to the same pattern of taxation in general, even though the interpersonal equity conditions pay attention to peoples’ economic circumstances, their ability to pay. The differences between them begin with the quest for horizontal equity.

Under the ability-to-pay principle, two persons are necessarily treated equally if they have the same tastes, abilities, and opportunities. Equal treatment under the interpersonal equity conditions also requires that people be identical over these three attributes but adds a fourth attribute as well: They must have the same marginal social welfare weights at equal levels of Haig–Simons income. Two persons with equal utility before tax necessarily have

---

10. For the purposes of this discussion, assume that Haig–Simons income is chosen as the ideal tax base to satisfy horizontal equity and is the item redistributed lump sum to satisfy the interpersonal equity conditions.
equal utility after tax under the interpersonal equity conditions only if they are equal across all the four attributes. Furthermore, if two persons with equal utility before tax are not identical over the first three attributes, then they are not necessarily treated the same under the two principles even if they have equal marginal social welfare weights.

To illustrate, compare the interpersonal equity conditions and Feldstein’s equal-utility-before-tax, equal-utility-after-tax criterion of horizontal equity within the context of a two-person, two-good exchange economy with fixed endowments of the two goods. Let $X_{ij} = \text{consumption of good } j \text{ by person } i$, for $i, j = 1, 2$. The first-order conditions for a social welfare maximum in this economy are

\begin{align*}
\text{Pareto optimality:} \quad \frac{\partial u^1}{\partial x_{11}} = \frac{\partial u^2}{\partial x_{22}} \tag{11.1} \\
\text{Interpersonal equity:} \quad \frac{\partial w}{\partial u^1} \frac{\partial u^1}{\partial x_{11}} = \frac{\partial w}{\partial u^2} \frac{\partial u^2}{\partial x_{22}} \tag{11.2} \\
\frac{\partial w}{\partial u^1} \frac{\partial u^1}{\partial x_{12}} = \frac{\partial w}{\partial u^2} \frac{\partial u^2}{\partial x_{22}} \tag{11.2}
\end{align*}

If the two otherwise-identical people have unequal social welfare weights, $\partial w/\partial u^1 \neq \partial w/\partial u^2$, evaluated at equal utility levels, the equal-utility-before-tax, equal-utility-after-tax criterion is inconsistent with Eqn (11.2), in general. This possibility can arise under an affirmative action policy that corrects for past injustices, as exists in the United States for women and minorities. The social welfare function can incorporate such a policy through the marginal social welfare weights, whereas the ability-to-pay principle cannot because it depends only on individuals’ utilities.

Suppose the social welfare weights are equal so that Eqn (11.2) becomes

\begin{align*}
\frac{\partial u^1}{\partial x_{11}} = \frac{\partial u^2}{\partial x_{21}} \tag{11.3} \\
\frac{\partial u^1}{\partial x_{12}} = \frac{\partial u^2}{\partial x_{22}}
\end{align*}

Even Eqn (11.3) differs from the horizontal equity criterion if people’s tastes and/or initial endowments are unequal. If the two consumers happen to enjoy the same level of utility at an initial pareto optimum before the government redistributes one of the goods to satisfy the interpersonal equity conditions, there is no guarantee that they will enjoy equal utility levels after the socially optimum tax and transfer has been effected. The following simple model in which the two persons have different tastes provides a counterexample. Let

\begin{align*}
U^1 &= X_{11}(3 + X_{12} + 27) \\
U^2 &= \frac{1}{2}X_{21}(1 + 2X_{22}) \\
X_{11} + X_{12} &= 10 \\
X_{21} + X_{22} &= 10 \\
\text{The reader can verify that an initial equal-utility pareto optimum occurs at} \\
&(X_1, X_2) \\
&\text{Person 1 (4, 2.4)} \\
&\text{Person 2 (6, 7.6)}
\end{align*}

The social welfare optimum, satisfying both Eqns (11.1) and (11.3), occurs at

\begin{align*}
(X_1, X_2) \\
&\text{Person 1 (5, 15/4)} \\
&\text{Person 2 (5, 25/4)}
\end{align*}

with unequal utilities. The difference occurs because the interpersonal equity requires equal after-tax marginal utilities, whereas the Feldstein criterion requires equal after-tax utility levels. Even after ignoring differences in social welfare weights, these two rules are consistent only if preferences and endowments are identical.

The unsettling conclusion is that the ability-to-pay principle of taxation is unlikely to be consistent with the interpersonal equity conditions of social welfare maximization. There are three differences between the two that cannot be fully reconciled.

The most important difference is that the interpersonal equity conditions add a new piece of information, the social welfare function, that is missing from the ability-to-pay principle. This alone is enough to generate inconsistencies between the two principles. The presence of the social welfare function also underscores their fundamentally different views of broad-based taxes: as promoters of social welfare on the one hand and as a necessary evil on the other hand.

A second difference is that horizontal equity under the ability-to-pay principle involves a before and after comparison of individuals’ utility levels: equal utility before tax, equal utility after tax. The interpersonal equity conditions, in contrast, are concerned only with individuals’ positions after tax (and transfer), and the comparison is in terms of margins, not levels: equal social marginal utilities. Moreover the equal-utility-level-before-tax, equal-utility-level-after-tax requirement is vulnerable as an equity principle because it is indifferent as to how two persons arrived at their equal utilities beforehand. One could be a respected
entrepreneur and the other a criminal, a distinction the social welfare function could take into account.

A final difference between them is that the quest for horizontal equity under the ability-to-pay principle is concerned with determining the ideal tax base, whereas the choice of the tax base is irrelevant under the interpersonal equity conditions. As we saw in Chapter 2, if pareto optimality holds and the interpersonal equity conditions are satisfied for any one good or factor, then the interpersonal equity conditions are automatically satisfied for all goods and factors, as required for a social welfare maximum. Any good or factor can be chosen for lump-sum redistribution; that is, any tax base will do. The only concern of the interpersonal equity conditions is vertical equity, the choice of the tax structure to be applied to whatever tax base is chosen. In summary, the ability-to-pay principles and the interpersonal equity conditions are quite different principles of taxation.

The question remains whether the ability-to-pay principle is a useful addition to neoclassical tax theory, given that the interpersonal equity conditions of social welfare maximization are the neoclassical statement of distributive equity. Might it not be better for policymakers to announce their preferred social welfare function, design a tax (and transfer) system that roughly corresponds to the requirements of the interpersonal equity conditions, and let citizens judge whether they are willing to accept the policymakers’ social welfare function? What is gained by adding a completely different set of equity principles to the design of tax policy? These questions are in the spirit of Musgrave’s suggestion to worry much more about the tax structure than the choice of an ideal tax base.

The practical answer appears to be that people are generally satisfied with the ability-to-pay principles. The Bergson—Samuelson social welfare function has had an enormous impact on the economic theory of the public sector but almost no impact at all on the design of broad-based taxes so far as equity itself is concerned. The only impact of social welfare analysis has been on the level of the tax rates, and then only when efficiency considerations are intermingled with equity considerations in a second-best environment. The interaction of efficiency and equity principles in taxation will be discussed in Chapters 14 and 15. The next step in this chapter is to consider the principle of vertical equity.

**VERTICAL EQUITY**

Once the ideal tax base has been determined, the quest for vertical equity centers on the design of the tax structure. Should the tax be levied at a single rate—a flat tax—or should the rates be graduated, rising with income? Should some minimum amount of income be exempt from taxation (assuming Haig—Simons income is the tax base)? Should taxpayers be allowed to deduct certain items of income or expenditure in computing their taxable income? The answers to these questions determine exactly how unequally unequals are treated under the tax laws, which is the central issue of vertical equity.

**Progressive, Proportional, and Regressive Taxes**

Actual policy discussions almost never get much further than the debate over whether taxes should be progressive, proportional, or regressive, three very broad indexes of vertical equity. Economists have devised various methods of defining these terms, but the most common definition is in terms of the average tax burden across individuals. Let

\[ Y_i = \text{value of the ideal tax base for individual } i. \]

\[ T_i = \text{burden of the ideal tax on individual } i. \]

The average tax burden on individual \( i \) is the ratio \( T_i/Y_i \). Rank order individuals on the basis of \( Y_i \) and ask how the average tax burden varies as \( Y_i \) increases.

- The tax is *progressive* if \( T_i/Y_i \) increases as \( Y_i \) increases.
- The tax is *proportional* if \( T_i/Y_i \) remains constant as \( Y_i \) increases.
- The tax is *regressive* if \( T_i/Y_i \) decreases as \( Y_i \) increases.

A number of points are worth stressing in applying this measure. The numerator should be the tax burden rather than the tax payment if the two differ, because the implicit standard is the relative loss in utility from the tax. By the same token, although the measure can be applied to any tax, the denominator should always be the ideal tax base for the purposes of assessing the vertical equity of the tax. The ideal tax base is the surrogate measure of an individual’s utility and not anything else that might happen to be taxed. Additionally, the time frame should correspond to the time frame used to determine the ideal tax base. For example, proponents of Haig—Simons income as the ideal tax base should use it for the \( Y_i \) and the annual tax burden of a particular tax for the \( T_i \). Proponents of consumption or expenditures should use the expected present value of lifetime consumption or income for the \( Y_i \) and the expected present value of the lifetime tax burden for a particular tax for the \( T_i \).

A final point is that the three broad characterizations of vertical equity are not very limiting. Suppose, for example, that Haig—Simons income is chosen as the ideal tax base and society decides that it wants to collect more taxes from the rich than the poor under the ability-to-pay principle. A wide range of tax structures—progressive, regressive, or proportional—can satisfy the vertical equity criterion of unequal treatment of unequals and collect more taxes from higher-income individuals. For example, a tax structure that applies a 10% rate to an income of $50,000 and a 5% rate to an income of $200,000 is regressive. Yet, it collects more
tax from the richer individual, in broad concordance with the ability-to-pay principle.

About all one can say with confidence for the United States is that there appears to be an overwhelming consensus in favor of progressive or proportional taxes over regressive taxes. Studies of the overall U.S. tax system tend to show that the burden of all taxes is roughly proportional over all but the lowest income levels, within which they are slightly progressive. The U.S. tax system does not appear to redistribute much purchasing power in and of itself. We will return to this point in Chapter 17.

**Vertical Equity and the Interpersonal Equity Conditions**

In principle, the interpersonal equity conditions solve the problem of achieving vertical equity in tax design as part of determining the optimal distribution of income (assuming, again, that Haig—Simons income is the ideal tax base). Suppose that \( Y^B = (Y^B_1, \ldots, Y^B_h, \ldots, Y^B_H) \) is the vector of Haig—Simons incomes across individuals before tax and transfer, and \( Y^A = (Y^A_1, \ldots, Y^A_h, \ldots, Y^A_H) \) is the vector of Haig—Simons incomes across individuals after taxing and transferring to satisfy the interpersonal equity conditions. The difference between the corresponding elements in \( Y^B \) and \( Y^A \) defines the exact rate of tax (or transfer) to apply to each individual.

As usual, however, the interpersonal equity conditions are not very helpful to the tax practitioner. In addition to the uncertainties surrounding the social welfare function, the pattern of taxation may require that different tax rates be applied to people with the same \( Y^B \) if, say, the social welfare function incorporates a policy of affirmative action. Taxing different people differently on some basis other than their incomes may well be illegal in the United States. It also violates Musgrave’s interpretation of horizontal equity as a proscription against taxation on the basis of inappropriate personal characteristics, a compelling proscription in matters of taxation.

Finally, we saw in Chapter 4 that attempts to apply the social welfare function typically assume: (1) equal marginal social welfare weights at equal incomes; (2) everyone has the same tastes; and (3) diminishing private marginal utility of income. The implication of the interpersonal equity conditions under these three assumptions is that everyone should have the mean level of income after tax and transfer. Hardly anyone accepts this view of vertical equity, perhaps because it is so difficult to ignore the efficiency implications of leveling everyone’s income to the mean. And, indeed, the mainstream position is that the efficiency implications of any tax should be incorporated into a social-welfare-maximizing framework to determine the optimal structure of the tax.

**Sacrifice Principles of Vertical Equity**

Public sector economists had long worked on the problem of vertical equity from the sacrifice perspective of the ability-to-pay principle, but without much success until 1988. This line of research had pretty much died out by the 1980s. The main suggestions for vertical equity in the tax literature at that time dated from the late 1800s to the early 1900s. Then, in 1988, H. Peyton Young achieved a substantial breakthrough. Building on one of the earlier principles, Young used the methods of cooperative game theory to develop specific recommendations for the tax structure. Young’s game-theoretic approach appears to be a promising avenue for future research.11

The two long-standing principles of vertical equity in taxation before Young wrote were minimum aggregate sacrifice and equal sacrifice.

**Minimize Aggregate Sacrifice**

The call to minimize the aggregate sacrifice from taxation came from the utilitarian school led by Jeremy Bentham, who believed that the economic goal of society should be to maximize aggregate happiness or utility. Their social welfare function was the straight sum of individual utilities. With broad-based taxes viewed as a necessary sacrifice for the common good, the corresponding utilitarian goal for tax policy was to minimize the aggregate sacrifice from collecting the taxes. Under the assumptions of identical tastes and diminishing marginal utility of income, aggregate sacrifice is minimized by levying taxes in a top—down, highly progressive manner until the required total tax revenue is collected.

To see this, suppose there are three groups of consumers having pretax incomes \( Y_1, Y_2, \) and \( Y_3, \) such that \( Y_1 < Y_2 < Y_3. \) Incomes are equal within each group. Assume further that their pretax marginal utilities of income are, respectively,

\[
\frac{\partial U^1}{\partial Y^1} = 10, \quad \frac{\partial U^2}{\partial Y^2} = 9, \quad \frac{\partial U^3}{\partial Y^3} = 8
\]

reflecting diminishing marginal utility.

If the government wants to collect a given amount of tax revenue, the minimum aggregate sacrifice principle requires that the government tax people in the third group until either their marginal utility rises to nine or the required tax revenue has been collected. If the former applies, then the government taxes both the second and third groups until either their marginal utility rises to 10 or the required tax revenue has been collected. If still more revenue needs to be collected, then the government taxes all three groups, maintaining equality on the margin, until the revenue

requirement has been met. This pattern of tax collections is highly progressive in terms of the tax burdens.

**Equal Sacrifice**

The other main suggestion called for equal sacrifice in terms of utility, the only debate being whether the government should require equal absolute sacrifice or equal proportional sacrifice. Letting $Y_h$ be pretax income and $T_h$ be the tax for person $h$, the two candidates are

Equal absolute sacrifice:
$$ U(Y_h) - U(Y_h - T_h) = c \quad \text{all } h = 1, \ldots, H $$

Equal proportional sacrifice:
$$ \frac{U(Y_h) - U(Y_h - T_h)}{U(Y_h) - T_h} = k \quad \text{all } h = 1, \ldots, H $$

The equal-proportional-sacrifice variation was a modern restatement of Aristotle’s belief that proportional taxation was the just way to raise tax revenues.

Neither the utilitarian nor equal-sacrifice versions of vertical equity ever gained much standing among economists as a prescription for the design of a tax structure. One problem at the outset was the cardinality of the measures. The utilitarian prescription relies on diminishing marginal utility, which is neither a necessary nor sufficient condition for diminishing marginal rates of substitution, the condition for a well-behaved consumer indifference map.

Even if marginal utility is diminishing with respect to one utility index, there exists an admissible monotonic transformation of the utility function that leaves demands (and factor supplies) unchanged and implies either constant or increasing marginal utility. That is, given a utility index, $\phi(X)$ and its transformation, $F[\phi(X)]$, $F > 0$,

$$ \frac{\partial^2 F[\phi(X)]}{\partial X_i^2} = F \frac{\partial^2 \phi}{\partial X_i^2} + \left( \frac{\partial \phi}{\partial X} \right)^2 F'' \geq 0 \quad (11.4) $$

is consistent with $\frac{\partial^2 \phi}{\partial X_i^2} < 0$ for $F'' > 0$.

The same problem plagues the equal sacrifice principles. Equal absolute or proportional sacrifice with respect to $\phi(X)$ does not necessarily imply equal absolute or proportional sacrifice with respect to $F[\phi(X)]$. Needless to say, economists are skeptical of any economic principles based on cardinal utility measures.

Finally, suppose the government picked one cardinal representation of the utility index that satisfies diminishing marginal utility for each person in order to design a tax structure. Unfortunately, the pattern of taxes implied by any of the sacrifice principles could be just about anything. Even the utilitarian tax program need not be progressive. Using a simple general equilibrium model with one good and one factor, Efriam Sadka was able to show that lump-sum taxes consistent with the utilitarian social welfare function would not necessarily be progressive, where factor income is used as the basis of comparison. Whether the taxes are progressive or not turns on a number of parameters, including the elasticity of the consumers’ indifference curves between the factor and the good, third derivatives of the utility function, and the like. Certainly no conclusions can be drawn a priori (Sadka, 1976).

**Young’s Prescription for Vertical Equity**

H. Peyton Young revived the equal sacrifice ability-to-pay principle of vertical equity by introducing a new and thoroughly modern view of the problem of tax design. Young reasoned that if society views broad-based taxes as a necessary evil, a sacrifice made for the common good, then the levying of these taxes ought to be viewed as a cooperative game played by all members of the society. The design of the tax structure becomes the standard exercise in cooperative game theory of establishing a set of sharing rules for splitting up the profits or costs of the game. In this instance, the design problem is to posit a set of sacrifice principles that society could agree to in the levying of a broad-based tax and see what the principles imply for the tax structure. Arrow used the same cooperative game theory approach in proving his General Impossibility Theorem for social decisions in a democratic society.

Young posited six principles that he thought a democratic society could agree to in the levying of a broad-based tax. He then proved that they imply equal sacrifice in terms of one of two utility functions commonly used in the theory of risk taking. They also imply very simple tax systems.

We will assume that Haig–Simons income has been chosen as the tax base in demonstrating his result. Also everyone is assumed to have the same tastes; individuals vary only in the amount of income they have. We saw that the same-tastes assumption is necessary when selecting a tax base as a surrogate measure of utility. It is also necessary in order to say anything definite about vertical equity.

**Young’s Six Principles of Taxation**

Young proposed the following six principles as the base for an equitable tax structure:

1. **The consistency principle**—If a method of taxation is considered to be fair for the entire group of taxpayers, then it must also be considered fair for any subgroup of the taxpayers. The force of this principle is to ensure that people cannot alter their tax liabilities simply by joining different subgroups. As such, it satisfies the requirement of coalition stability for solutions of cooperative games. The consistency principle is automatically satisfied if the tax is levied on individuals, since
different subgroupings or coalitions of taxpayers cannot possibly alter individual tax liabilities.\textsuperscript{12}

2. **Monotonicity**—If the government is forced to increase total tax revenues, then everyone’s tax liability must increase. This is the strong version of the principle. The weak version is that if total tax revenues increase, then no individual’s tax liability can decrease. The monotonicity principle captures the spirit of ability to pay as a sacrifice principle, namely that the taxpayers are all in this game together. Notice that the strong version might not be satisfied by the utilitarian aggregate minimum sacrifice principle with its highly progressive, top–down tax collections.

3. **The composition principle**—The method used to raise a given amount of tax revenue must also be used to raise any increment in tax revenue. In other words, society should stick with the method that it believes is fair. This principle is satisfied by surtaxes, which raise additional revenue by requiring taxpayers to pay an additional percentage of their existing tax liability.

Feldstein’s principles of horizontal and vertical equalities constitute (4) and (5):

4. **Horizontal equity**—Two persons with equal utility before tax should have equal utility after tax.

5. **Vertical equity**—No utility reversals. For any two persons, the one with higher utility before tax must have higher utility after tax.

These two principles can also be stated in terms of Haig–Simons income since it is assumed to be an appropriate surrogate measure of utility.

6. **Scale invariance or the homogeneity principle**—Suppose everyone’s incomes and the revenue requirement increase by a scalar $0$. Then, everyone’s tax liability must increase by $0$. This principle is standard in income distribution theory, where it is applied to measures of income inequality. The idea is that an index of inequality should be invariant to scalar increases or decreases in everyone’s income. It applies to relative tax burdens in this context.

The results of cooperative game theory rely on accepting the underlying principles, which may or may not be persuasive. If a democratic society were to accept Young’s six principles of tax design, however, then the results are rather striking. Young proved that the first five principles hold if and only if the tax collections imply equal sacrifice with respect to some utility function, without specifying what that function should be. By adding the homogeneity principle, Young’s six principles hold if and only if tax collections imply equal sacrifice with respect to one of the following two utility functions:

$$U^h = a \ln(Y_h) + b \quad \text{or} \quad U^h = aY^p + b \quad a, P < 0$$

These are the utility functions commonly used in the theory of risk taking because they exhibit constant relative risk aversion (CRRA), meaning that the elasticity of marginal utility with respect to income is constant (as the reader can easily verify). Equal sacrifice under these two utility functions in turn implies very simple tax functions: the first, a proportional tax and the second, a progressive tax.

An important point to note before demonstrating Young’s results is that the distinction between equal absolute sacrifice and equal proportional sacrifice is irrelevant to modern economic theory. The reason is that equal absolute sacrifice with respect to some utility function, say $U$, is equivalent to equal proportional sacrifice with respect to the function $e^U$, which is a valid monotonic transformation of $U$ and would have no effect on individual choice. To see this, assume equal absolute sacrifice exists with respect to $U$, such that $U(Y_h) - U(Y_h - T_h) = C$. Equal proportional sacrifice with respect to $e^U$ is

$$\left[ e^{U(Y_h)} - e^{U(Y_h - T_h)} \right] / e^{U(Y_h - T_h)} = K$$

Simplifying Eqn (11.5) and rearranging terms, equal proportional sacrifice implies

$$e^{[U(Y_h) - (U(Y_h - T_h)]} = K = 1 = K'$$

which can only hold if $U(Y_h) - U(Y_h - T_h)$ is constant.

We will consider the sufficient conditions to see what Young’s principles imply for the tax structure.\textsuperscript{13} The first task is to show that each of Young’s first five principles hold if the tax collections satisfy equal sacrifice with respect to some utility function. (Equal absolute sacrifice is easier to work with.) Therefore, suppose $U(Y_h) - U(Y_h - T_h) = C$, for $h = 1, \ldots, H$, and consider each of the first five principles.

1. **Consistency**—This holds by definition assuming that the tax base is each individual’s Haig–Simons income.

2. **Monotonicity**—The strong version of monotonicity must hold under equal absolute sacrifice assuming positive marginal utility of income. Let total tax collections rise and assume that person $i$ is taxed more. Then $U(Y_i) - U(Y_i - T_i) > C$. To maintain equal absolute sacrifice, everyone else must pay more tax to increase their

\textsuperscript{12} It is not satisfied by the U.S. federal personal income tax, however, because the IRS cannot decide if it wants to tax on an individual or a family basis. As a result, taxpayers within a family have the options of filing as individuals or pooling their incomes and filing jointly as members of a family. The individual and joint filing income cut-offs at which the different graduated rates apply differ, which means that taxpayers’ liabilities can vary if they marry or divorce. Young’s principle would permit only individual filing and thereby avoids the marry/divorce problem.

\textsuperscript{13} The necessary conditions are much more difficult to prove and will be left to a reading of Young’s paper.
difference between $U(Y_h)$ and $U(Y_h - T_h)$ and restore equal sacrifice.

3. **Composition**—Assume that $U(Y_h) - U(Y_h - T_{1h}) = C$ for given total tax collections $T_1$. Suppose that tax collections rise to $T_2$ and equal absolute sacrifice is maintained for the increment of taxes between $T_1$ and $T_2$: $U(Y_h - T_{1h}) - U(Y_h - T_{1h} - T_{2h}) = C^*$. Adding the two results: $U(Y_h) - U(Y_h - T_{1h} - T_{2h}) = C + C^*$. Equal absolute sacrifice is also maintained for the new higher tax collections $T_2$.

4. **Feldstein’s horizontal equity principle**—Two people with equal utility before tax should have equal utility after tax.

5. **Feldstein’s vertical equity principle**—No utility reversals. For any two people, the person with higher utility before tax must have higher utility after tax.

These two principles must hold under equal absolute sacrifice as long as the marginal utility of income is positive. Regarding horizontal equity, if $U(Y_h) = U(Y_j)$ and $U(Y_h) - U(Y_j - T_h) = U(Y_j) - U(Y_j - T_j)$, then $U(Y_h - T_h) = U(Y_j - T_j)$. Regarding the principle of no reversals, if $U(Y_h) > U(Y_j)$ and $U(Y_h) - U(Y_h - T_h) = U(Y_j) - U(Y_j - T_j)$, then $U(Y_h - T_h) > U(Y_j - T_j)$. Therefore, equal absolute sacrifice with respect to any valid utility function $U$ satisfies each of Young’s first five principles of taxation.

Now add the scale invariance or homogeneity principle, which generates Young’s two proposed tax structures. The sufficient conditions on the tax structures involve two steps. First, determine the tax structure implied by equal absolute sacrifice with respect to the two CRRA utility functions noted above; second, show that the tax structures are scale invariant.

### Proportional Taxation

Consider the utility function $U^h = a \ln Y_h + b$. Equal absolute sacrifice implies $a \ln Y_h = a \ln (Y_h - T_h) = C$, for $h = 1, \ldots, H$. The left-hand side is constant at any income if (and only if) $T_h = tY_h$, that is, under a flat-rate, proportional tax:

$$a \ln Y_h - a(1 - t)Y_h = a \ln (1 - t) = C \quad (11.7)$$

A proportional tax clearly satisfies the homogeneity principle; the ratio $\theta T_h/\theta Y_h$ is independent of $\theta$. Young’s six principles of taxation have resurrected Aristotle’s call for proportional taxation, assuming loglinear utility.

### Progressive Taxation

Now consider the utility function $U^h = aY_h^p + b$. Equal absolute sacrifice implies

$$a Y_h^p - a(Y_h - T_h)^p = C \quad h = 1, \ldots, H \quad (11.8)$$

Rearranging terms and solving for $T_h$ yields

$$a Y_h^p - (Y_h - T_h)^p = C/a = -\lambda, \quad \text{with } a < 0 \quad (11.9)$$

$$\left(Y_h - T_h\right)^p = \left(Y_h^p + \lambda\right) \quad (11.10)$$

$$T_h = Y_h - (Y_h^p + \lambda)^{1/p} \quad (11.11)$$

Under this tax, individual tax collections can be multiplied by a scalar as needed for total revenues. Therefore, the tax is a flat-rate tax applied to a tax base in which taxpayers exempt an amount $(Y_h^p + \lambda)^{1/p}$ from their Haig–Simons income $(Y_h)$ in determining their taxable income. The tax has a number of interesting properties.

First, $T_h/Y_h$ is independent of $\theta$. This follows from dividing Eqn (11.11) by $Y_h$ and noting from Eqn (11.9) that scaling $T_h$ and $Y_h$ by $\theta$ scales $\lambda$ by $\theta^p$.

Second, the tax is progressive in terms of the standard average tax burden measure of progressivity. The average tax burden increases as $Y_h$ increases (divide Eqn (11.11) by $Y_h$ and recall that $P < 0$ and $\lambda > 0$).

Third, and most unusual, the exemption from the income in computing taxable income, $(Y_h^p + \lambda)^{1/p}$ increases as income increases. In all actual taxes with exemptions, the exemption either remains constant or decreases as income increases. Even so, the increasing exemption does not prevent the tax from being progressive.

Fourth, the homogeneity principle rules out graduated tax rates (although not progressive taxes).

In conclusion, Young has provided a rationale for either proportional or progressive broad-based taxes using the methods of cooperative game theory. In doing so, he has brought the old equal-sacrifice principle of taxation into the realm of modern economic theory. Whether he has done so successfully depends on a society’s willingness to accept his six principles of fair taxation. Perhaps some other set of sacrifice principles would be viewed as more attractive and imply quite different tax structures.

### Vertical Equity in the United States

The five major broad-based taxes in the United States give a mixed reading on how unequally the United States is willing to treat unequals. As we will see in Chapter 17, some of the taxes are progressive and others regressive. One can argue that the federal personal income tax gives the clearest signal of the U.S. view of vertical equity since it is designed on ability-to-pay principles. Unfortunately, it gives mixed signals as well.

The federal personal income tax appears to be fairly progressive on paper, with a graduated rate structure ranging from 10% to 39.6% and a large exemption of the first dollars of income to protect the poor from taxation. It
turns out to be much less progressive in practice, however, because capital gains and some other forms of income from capital receive highly favorable tax treatment, in some cases no tax at all. Capital income is highly concentrated among the richer taxpayers.

The recent history of the federal personal income tax has not clarified matters. TRA86 reduced the graduated rate schedule from 14 brackets ranging from 11% to 50% to five brackets ranging from 11% to 38.5% in 1987 and then to two brackets in 1988, 15% and 28%. The reduction of the top rate to 28% was done in large part to improve the dynamic efficiency of the tax. At the same time, TRA86 sharply increased the personal exemption to protect the poor from taxation. The history of the tax rates since 1988 has been marked by frequent changes: 1991, three brackets ranging from 15% to 31%; 1993, five brackets ranging from 15% to 39.6%; 2002, six brackets ranging from 10% to 38.5%; 2003, six brackets ranging from 10% to 35%; and 2013, seven brackets ranging from 10% to 39.6%.14

The message from all these reforms is unclear, except for a desire to protect the poor from taxation. The earned income tax credit, which grew rapidly during the 1990s, also greatly reduces the federal tax burden on the poor. At the same time, however, three of the other major U.S. taxes—the federal payroll tax, the state sales taxes, and the local property taxes—do not protect the poor from taxation.

A widely cited study of the U.S. tax system by Joseph Pechman and Benjamin Okner, last updated in 1984, estimated that the overall U.S. tax structure is mildly progressive at the lowest incomes and then roughly proportional over all remaining income levels. The low-end progressivity is due largely to the exemptions under the federal and state personal income taxes—the federal payroll tax, the state sales taxes, and the local property taxes—do not protect the poor from taxation.

A widely cited study of the U.S. tax system by Joseph Pechman and Benjamin Okner, last updated in 1984, estimated that the overall U.S. tax structure is mildly progressive at the lowest incomes and then roughly proportional over all remaining income levels. The low-end progressivity is due largely to the exemptions under the federal and state personal income taxes. (Other studies since 1984 have reached approximately the same conclusion.) The U.S. tax system does not redistribute much purchasing power in and of itself.15

### REFLECTIONS ON THE HAIG—SIMONS CRITERION IN PRACTICE: THE FEDERAL PERSONAL INCOME TAX

Despite its appeal to many public sector economists, the Haig–Simons income measure has not fared well in the United States. Only the federal and state personal income taxes pay so much as lip service to the Haig–Simons criterion. State governments rely heavily on sales taxes and local governments rely primarily on the property taxes, neither of which is valid according to the Haig–Simons criterion.

State sales taxes may appear to be consistent with the view that consumption is the ideal tax base. In practice, however, sales taxes are far removed from an ideal consumption tax. Sales taxes often exclude broad classes of expenditures from taxation, they usually tax all included items at one fixed rate, and they are levied on businesses. What expenditures tax proponents have in mind is a tax levied on individuals exactly as the federal income tax is, except that it would exclude saving from the tax base. A graduated rate schedule could easily be applied to individual expenditures, removing the stigma from sales taxes that they may be regressive.16

The federal personal income tax is the single largest tax in the United States. Of all the broad-based taxes, it comes closest to the Haig–Simons income measure as its tax base, but not really all that close. Recall that, according to the Haig–Simons criterion, the federal income tax base should include personal income and capital gains on assets held from the beginning of the tax year, without regard to the sources or uses of income. The only permitted deductions from Haig–Simons income are legitimate business expenses, i.e., expenses required to earn the income. The actual tax base falls far short of the Haig–Simons ideal, both the personal income and capital gains components.

### Personal Income

Taxable income is only about half of personal income. The main discrepancies between taxable and personal incomes fall into three categories: exemptions, exclusions, and deductions.

An exemption is an income that is recognized as taxable income by the Internal Revenue Service (IRS) but is simply not taxed. The main example is the personal exemption given to the taxpayer and all the taxpayer’s dependents. The exemption was $3950 per person in 2014, and it is adjusted each year for increases in the consumer price index (CPI).

Exclusions are sources of income that are counted as personal income by the U.S. Department of Commerce in the National Income and Product. Accounts are not counted as taxable income by the IRS. The principal exclusions are employee fringe benefits (primarily employer contributions to pension plans (along with the accrued interest on the investments under these plans), health insurance, and life insurance); interest income on IRAs and Roth IRAs, which are earmarked for retirement income; many federal, state, and local transfer

---


15. Okner and Pechman (1974). We will take a closer look at tax-incidence studies in Chapter 17.

16. This stigma may be more myth than reality. See the discussion in Chapter 17 on the incidence (burden) of the sales tax.
payments; imputed rental income on owner-occupied homes and imputed income on farm produce consumed on the farm; and interest received on state and local bonds, commonly referred to as “municipals.”

Deductions are not sources of income at all, but rather certain expenditures that can be deducted from personal income in computing taxable income. The most important itemized deductions are extraordinary medical payments and other uninsured losses, state and local income and property taxes (taxpayers can elect to deduct state sales taxes instead of state income taxes), interest payments on mortgages for the principal residence, contributions to charities and other nonprofit organizations, and business expenses. Taxpayers can elect to take a “standard deduction” (equal to $12,400 in 2014 for married taxpayers filing jointly) instead of itemizing deductions.

The various exemptions, exclusions, and deductions exist because Congress and the administration cannot avoid the temptation to use the income tax to pursue other social ends, such as protecting low-income families and individuals from taxation, promoting homeownership, helping people save for their retirement years, subsidizing state and local governments, encouraging charitable giving, and so forth. These may all be worthy goals, but they come at a cost. The exemptions, exclusions, and deductions can undermine the horizontal and vertical equities of the tax. They also introduce inefficient distortions into the income tax. (We will return to these points below.)

Capital Gains

The capital gains portion of the tax base is also far from the Haig–Simons ideal. Recall that capital gains should be taxed as they accrue, and at the same rate applied to personal income. Also, capital losses should be fully offset against other income because they represent equal dollar decreases in purchasing power. Finally, the tax base should reflect increases in real purchasing power only. Increases in income arising solely from inflation should not be taxed. If nominal income is used as the tax base, then at least all sources of income should be treated equally with respect to the effects of inflation on purchasing power.

Capital gains taxation is deficient on all counts. Capital gains are taxed on a realized basis (that is, only when an asset is sold) and then at a separate rate from the rates applied to personal income (for assets held for more than 1 year), a rate that is often far below the rates applied to the higher-income brackets (20% in 2014); in effect, part of the realized gains is excluded from the tax base. The ability to offset losses against income is mostly limited to offsets against capital gains.

Finally, the tax is levied on nominal capital gains, with no adjustment for the effects of inflation on purchasing power. As a result, capital and wage income are treated very differently in times of inflation.

Equity judgments about the income tax would be easiest if Haig–Simons income were the tax base (except for business expenses, which we will ignore from now on). Since Haig–Simons income is assumed to be the surrogate measure of utility, the tax payments themselves would be the appropriate basis for judgment. Horizontal equity would be satisfied if two persons with the same Haig–Simons income paid the same tax. Similarly, vertical equity would be appropriately measured by the difference in taxes paid by people with different amounts of Haig–Simons income.

Given the exemptions, exclusions, and deductions, however, the tax payments are no longer accurate measures of either horizontal or vertical equity. The problem is that markets react to any differences from the ideal tax base, and the market reactions have to be factored into any assessment of horizontal and vertical equities. They are sources of gains and losses to the taxpayers that matter every bit as much as the tax payments themselves. Tax burdens, not tax payments, determine the equity of the tax, and Feldstein’s versions of horizontal and vertical equities defined in terms of utility are the only appropriate basis of judgment. For example, the proper statement of horizontal equity is that two persons with the same utility before tax should have the same utility after tax. That is, they should bear the same tax burden, the same loss of utility.

Unfortunately, judgments based on before and after utility comparisons can be problematic and are likely to confuse the general public. People tend to see the individual incomes and the tax payments, not the additional market-induced gains and losses. As a result, the exemptions, exclusions, and deductions are branded pejoratively as “tax loopholes.” People see two taxpayers with the same Haig–Simons income who do not pay the same tax. Even worse, they see higher-income people paying less tax than lower-income people, an apparent equity reversal. The public’s sense of horizontal and vertical equities is offended.

The Taxation of Personal Income: The Tax Loopholes

Not all tax loopholes in the personal income portion of the tax base are equally distasteful. The personal exemptions, for example, do not generate much complaint from the public. Protecting low-income people from taxation is an accepted goal in a nation that has declared a war on poverty. One of the often-stated criticisms of sales taxes, property taxes, and the payroll tax is that they do not offer such protection. Even proponents of a single, flat-rate
income tax favor including a personal exemption for the taxpayers and their dependents.

Exemptions are a simple way to ensure that a tax is progressive, if progressivity is desired. Fig. 11.2 illustrates the case of a taxpayer with three other dependents. It assumes a flat-rate tax of 18% on all income beyond a personal exemption of $5,000, or $20,000 for a family of four ($20,000 was approximately equal to the poverty line for a family of four in 2010). The vertical axis pictures the marginal and average tax rates. The marginal rates are 0 up to $20,000 and 18% thereafter. The average rates are also 0 up to $20,000 but then rise steadily beyond $20,000 as tax payments begin, approaching 18% asymptotically. (For example, at an income of $120,000, the tax is $18,000, and the average tax rate is $18/120 \approx 15\%$). The tax is progressive by the usual average tax rate measure.

The exclusions and deductions are far more contentious “loopholes,” as perhaps they should be. They violate the pattern of vertical equity implicit in the tax structure, and they generate market and other forms of inefficiency. They may not be a source of horizontal inequity, however, despite the common perception that they are. The relationship between tax loopholes and horizontal equity is a particularly subtle issue that illustrates the importance of the market’s reaction to the loopholes.

**Tax Loopholes, Tax Capitalization, and Horizontal Equity**

Consider the three large tax breaks to homeowners relative to those who rent an apartment: the exclusion of imputed rent on the home, the deduction for the interest payments on the mortgage, and the deduction for the local property taxes on the home. Fig. 11.3 illustrates the market’s reaction to the tax break. The left-hand panel depicts the market for owner-occupied homes purchased by people within a certain income range (housing markets segment by income.) The right-hand panel depicts the market for rental apartments purchased by people within the same income range. The apartments are assumed to provide the same housing services as the owner-occupied homes, and the people in these markets are assumed to have identical tastes.

The equilibrium before these three tax breaks were introduced into the federal tax is given by the intersection of $D^0$ and $S$ in each market. Since the housing services are identical, the prices are the same in each market, $P_H^0$ and $P_A^0$ ($P_H^0$ is the annualized price of a home, the implicit rental value). The people are indifferent to owning or renting.

The introduction of the three tax breaks makes the owner-occupied homes more attractive. Demand shifts up
in the owner-occupied market and down in the apartment market, driving the (annualized) price of the homes up and the rentals on the apartments down. The new equilibrium occurs at the intersection of $D^1$ and $S$ in each market, with the new equilibrium prices, $P_H^1$ and $P_A^1$.

At the new equilibrium, the difference in prices $P_H^1 - P_A^1$ must equal the annualized value of the three tax breaks to the homeowner. The market is said to capitalize the tax breaks into the relative prices of the two forms of housing. The implication of the market capitalization is that once the new equilibrium has been reached, the people in this income range are once again indifferent to owning or renting. If they choose to buy a house, the higher price minus the value of the tax breaks equals the rent they would have to pay for the apartment, $P_A^1$. This has to be the case, since the housing services are the same for the homes and the apartments, the people have identical tastes, and they are free to purchase a home or rent. Indifference to owning or renting is the only possible long-run equilibrium, regardless of the tax system.

This example illustrates the principle that any two persons in these markets, who had equal utility before the tax breaks were introduced, must have equal utility once the market returns to equilibrium in response to the tax breaks. The tax breaks do not violate horizontal equity in the new equilibrium. The homeowners get the tax breaks but no gain in utility relative to the renters. The same analysis applies to any tax loophole and for the same reason: The value of the loophole is eventually fully capitalized by the market system.

Feldstein summarized this fundamental principle of tax design as follows:

"Once the market system establishes a long-run equilibrium in response to a given tax system, the tax system per se cannot be a source of horizontal inequity, where horizontal equity is defined in terms of burden or utility."

A corollary to this fundamental principle of tax design is an equally fundamental principle of tax reform:

"Any reform of an existing tax code will create horizontal inequities through unanticipated gains and losses, and will continue to do so until a new long-run equilibrium obtains in the market place."

Continuing with the housing example, suppose the three loopholes favoring homeownership were suddenly removed for promoting horizontal equity. Assuming the long-run equilibrium had been achieved, current homeowners surely lose, but not necessarily because they lose a tax advantage that had been unfairly given them, as the reformers intend. Rather, some of them will lose because they never received any gain in the first place at the higher prices they paid for their homes. These pure losses are an unavoidable consequence of any tax reform that removes the “loopholes.”

A final point is that determining who gains from the three tax breaks is difficult once the market has reached its new equilibrium and the homes have changed hands a few times. Tax loopholes can even be capitalized in anticipation of the loopholes, before they become part of the law. In conclusion, simply looking at tax payments gives a very misleading picture of horizontal equity when the tax contains various exclusions and deductions from Haig–Simons income.

### Tax Loopholes, Vertical Equity, and Inefficiency

Although tax loopholes may not be a source of horizontal inequity, tax reformers can still make a good case for removing them. They are likely to give rise to vertical inequities, and they lead to various kinds of inefficiencies. Therefore, the gains to vertical equity and efficiency from removing the loopholes may exceed any temporary horizontal inequities plus the lost benefits associated with whatever social goals the loopholes are trying to promote. This is especially so if there are more effective ways of promoting the social goals.

The housing example above illustrates the possibility of vertical inequity. Both homeowners and apartment renters gain equally from the three tax breaks. Their

---


18. For further analysis of the owner-occupied tax breaks that include supply adjustments, see White and White (1977). Boris Bittker tells an amusing anecdote illustrating the principle of capitalization. It concerns an eager, young law student who searches in vain for the benefits of a tax-sheltered apartment building in his hometown known as Rainbow Gardens. The tax shelter had been in existence since the inception of federal income taxation under the Revenue Act of 1913. Rainbow Gardens was for sale, but the law student quickly surmised that at the asking price he would only realize a normal return on his investment. He also learned that the current owners were selling because they, too, were only able to earn a normal return despite the existence of the tax shelter. The same had been true of the previous owners, and the ones before them, and so on. Alas, they all paid too much to realize an economic profit from the tax shelter. The persistent student was able to trace the line of ownership all the way back to R. E. Greison, who had purchased Rainbow Gardens in 1896. Greison possessed a remarkable foresight. In 1896, he was clerking for a U.S. Supreme Court justice when the Court ruled that a federal income tax was unconstitutional. Greison nonetheless correctly predicted that the Court’s decision would eventually be overturned by a constitutional amendment (the 16th), and further that the income tax law, when drafted, would tax shelter apartment buildings. Based on these predictions, Greison bought Rainbow Gardens. Sad to say, the capitalization of the tax shelter predated Greison. His epitaph read: “Sacred to the memory of R. E. Greison, who learned that before every early bird, there is an earlier bird.” See Bittker (1975).
annual costs fall from \( P_0^S ( = P_0^D) \) to \( P_A^S \) because of the tax breaks, whether they own or rent. But, as noted above, housing markets segment by income. It is possible, therefore, that the decrease in their housing costs is greater than the decrease in the housing costs of other lower-income people. If so, the larger break to the higher-income people is likely to offend people’s sense of vertical equity.

Exclusions and deductions always generate this kind of vertical inequity under an income tax with graduated rates. Since the exclusion or deduction is taken off the tax base, it reduces the taxpayer’s liability by \( t \) cents per dollar of exclusion or deduction, where \( t \) is the taxpayer’s marginal tax rate. Under a graduated tax, the value of the tax savings rises with income. For example, every dollar given to charity that can be deducted from taxable income costs the taxpayer only \( $ (1 - t) \). The richer the taxpayers, the lower their costs of contributing to their favorite charity, church, or school. This is why economists tend to favor tax credits over exclusions or deductions if the tax system is to be used to encourage certain activities. A 10% tax credit is taken directly against the tax liability, after the tax has been computed, so that it is 10% for all taxpayers.

The housing example also illustrates the market inefficiencies of tax loopholes. As we will learn in Chapter 13, anything that drives a market away from its normal equilibrium is to be used to encourage certain activities. A broader tax base means that the same tax revenue can be collected with lower tax rates, which sharply reduces the size of the deadweight loss in the marketplace. We will see in Chapter 13 that the deadweight efficiency loss from a tax varies directly with the square of the tax rate.

Would the gains to vertical equity and efficiency from removing the exclusions and deductions more than offset any temporary horizontal inequities that may arise and the social benefits of the loopholes? This remains an open question, but many economists favor removing most of the exclusions and deductions. The economists in the Treasury Department during the Reagan administration put forth such a plan in their proposal for TRA86. The administration’s proposal called for almost a textbook version of an income tax based on Haig—Simons income, with little more than the personal exemptions and legitimate business expenses as reductions to the tax base. The administration’s proposal could not stand up to the special interest groups favoring the loopholes, however, and all the major exclusions and deductions were retained. No major tax base reform proposals have received a serious hearing in Congress since 1986.

The problem is that only investors who can save more than $20 in taxes will purchase the municipals. For example, at the assumed interest rates, a person in the 28% tax bracket can earn interest of $80 net of tax on the municipal at 8% but only $72 net of tax on the taxable bond at 10%. Therefore, the U.S. Treasury loses more than $1 in tax revenue for every $1 of interest subsidy received by a state or locality, $28 of lost revenue for a $20 subsidy in this example. In contrast, a direct federal subsidy (grant-in-aid) for capital expenditures would give $1 of subsidy for every $1 of tax revenue collected, a more efficient subsidy from the federal government’s viewpoint. Also, the direct subsidy avoids the deadweight loss inefficiencies in the bond market as the tax break to the municipals is capitalized into a lower interest rate relative to the taxable bonds.

Another source of efficiency gain from removing the loopholes is that the tax becomes much simpler, which saves on administrative and compliance costs. In addition, a broader tax base means that the same tax revenues can be collected with lower tax rates, which sharply reduces the size of the deadweight loss in the market for houses. Tax loopholes inevitably lead markets away from their natural equilibrium as the markets capitalize the loopholes. Thus, they necessarily generate efficiency losses (unless supply or demand is perfectly inelastic, which is rare in the long run).

Tax loopholes lead to other inefficiencies as well. Consider the exclusion for interest received on state and local bonds, the municipals. The exclusion acts as a subsidy to the lower-level governments, equal to the reduction in debt service made possible by the municipals’ tax-free status, but this is a particularly inefficient form of subsidy from the federal government’s point of view. Suppose a state government can offer an interest rate of 8% rather than 10% because of the exclusion, a savings of $20 of interest income on each $1000 bond.

---

19. We will see in Chapter 13 that efficiency losses arise only in markets that are distorted by features such as taxes or subsidies and market power. In this example, the competitive apartment market is not distorted and thus not a source of efficiency loss.
saving and investment. Taxing capital gains on a realized rather than an accrued basis generates further sources of inefficiencies and vertical inequities. The next two sections consider these issues.

THE INFLATIONARY BIAS AGAINST INCOME FROM CAPITAL

The U.S. Tax Codes were written for a noninflationary economy. This, in itself, generates an extra tax burden on income from capital relative to wage income, simply because inflation causes nominal asset income to grow more rapidly than nominal wage income. Consequently, equal growth in nominal income from these two sources reflects unequal growth in real purchasing power, so that equal taxation implies unequal treatment, in violation of horizontal equity. The inequity is compounded by the tax base of asset income may be taxed at higher rates. Graduated rate schedule, because the artificially expanded tax base of asset income may be taxed at higher rates.

To see how the differential inflation effect arises, assume that an economy has been experiencing inflation since time \( t = 0 \). Define the accumulated inflation to time \( t \) as

\[
I(t) = \exp \int_0^t i(s)ds \quad (11.12)
\]

where \( i(t) \) is the instantaneous rate of inflation at time \( t \). Assume further that inflation is fully anticipated so that \( i(t) \) represents both the actual and expected rates of inflation. If \( W(t) \) represents wage income at time \( t \) without inflation, then

\[
W'(t) = W(t) \cdot I(t) \quad (11.13)
\]

is wage income with inflation.

Let \( Y(t) \) represent income from capital in the absence of inflation:

\[
Y(t) = r(t) \cdot V(t) \quad (11.14)
\]

where \( r(t) \) is the real rate of return and \( V(t) \) is the value of an asset without inflation. The basis arises because expected inflations affects income from capital in two ways. It increases both the value of assets and the rate of return on assets. Let

\[
V'(t) = V(t) \cdot I(t) \quad (11.15)
\]

represent the value of assets with inflation, and

\[
n(t) = r(t) + i(t) \quad (11.16)
\]

represent the nominal rate of return. Hence

\[
Y'(t) = n(t) \cdot V'(t) = [r(t) + i(t)]V(t) \cdot I(t) \quad (11.17)
\]

where \( Y'(t) \) is income from capital with inflation. Dividing Eqn (11.17) by Eqn (11.13), using Eqns (11.14) and (11.15), and rearranging terms yields

\[
\frac{Y'(t)}{W'(t)} = \frac{(r(t) + i(t)) \cdot V(t) \cdot I(t)}{W(t) \cdot I(t)} = \frac{Y(t)}{W(t)} + \frac{i(t) \cdot V'(t)}{W'(t)} > \frac{Y(t)}{W(t)} \quad (11.18)
\]

Therefore, capital income grows more rapidly than wage income simply because of the inflation factor. If the tax base is nominal income, capital income is overly taxed. By inspection of the right-hand side (RHS) of Eqn (11.18), the inflationary bias can be removed by subtracting from nominal asset income the expected rate of inflation times the value of assets with inflation, \( i(t) \cdot V'(t) \), before applying the tax rates.

The inflation adjustment should be applied to all sources of capital income, but the nature of the adjustment varies depending upon the particular form of the asset. For example, if the income derives from an interest-bearing asset, the taxable income should include only the proportion of the interest resulting from the real rate of return. That is, if \( Y'(t) = (r(t) + i(t)) \cdot V'(t) \), then

\[
Y'(t) - i(t) \cdot V'(t) = r(t) \cdot V'(t) \quad (11.19)
\]

But,

\[
r(t) \cdot V'(t) = \frac{r(t)}{n(t)} \cdot n(t) \cdot V'(t) = \frac{r(t)}{n(t)} \cdot Y'(t) \quad (11.20)
\]

Thus, the taxpayer would report actual interest payments times the ratio of the real to the nominal rate of return.

For a straight capital gain without interest payments, the taxpayer would increase the purchase price by the accumulated inflation factor before subtracting it from the current value to compute the capital gain. For these assets,

\[
Y'(t) = CV - PV \quad (11.21)
\]

where

- CV = current value, inclusive of inflation.
- PV = original purchase value.

Adjusting \( Y'(t) \) yields

\[
Y'_{\text{adjusted}} = (CV - PV) - (CV - PV)_{\text{inflation}} \quad (11.22)
\]

\[
Y'_{\text{adjusted}} = (CV - PV) - (PV \cdot I(t) - PV) \quad (11.23)
\]

---


21. Vertical equity is also necessarily violated by the inflation bias, presumably in an anti-rich, pro-poor direction. Horizontal equity may not be violated if investments in physical and human capital are perfect substitutes. In that case, the inflation bias would drive investment toward human capital, lowering wages and raising the return to physical capital until the difference just equaled the value of the inflation bias against physical capital.
Finally, money holdings would receive a credit equal to \( t(t) \cdot V(t) \), since there is no nominal return from which to subtract this adjustment factor. Diamond recommends ignoring this adjustment on the grounds that the bookkeeping for cash assets would be especially difficult and that the liquidity services from holding cash are untaxed anyway.\(^{22}\)

These inflation adjustments are correct as given only if inflation is always fully anticipated and all income inflates at the same rate over time. In actuality, of course, neither of these is true, and it is not clear what should be done to correct the discrepancies. As a practical matter, governments would surely have to use actual rather than expected rates of inflation for any adjustment. Economic research has not even been able to determine how inflationary expectations are formed. Yet nominal interest rates and capital values almost certainly adjust to some degree for anticipated inflation. Thus, it is probably more accurate to adjust capital income by some long-run smoothed inflation index than to make no adjustment at all. People who anticipate inflation incorrectly will either make windfall gains or losses relative to the theoretical ideal, but this is unavoidable.

The practical question remains as to which long-run series to use, since assets and other sources of income inflate at different rates. A broad series such as the CPI is probably a reasonable choice for practical purposes, although again some people will receive (suffer) windfall gains (losses) in purchasing power relative to the ideal.

**TAXING REALIZED GAINS: AUERBACH’S RETROSPECTIVE TAXATION PROPOSAL**

Economists have long understood the equity and efficiency problems caused by taxing capital gains on a realized basis rather than an accrued basis. The equity problem is that deferring taxes on the capital gains until they are realized places the government in the position of offering interest-free loans each year to the asset holders. The loans equal the amount of the tax liability that would have been paid on an accrued basis. Since assets that generate capital gains, such as common stocks, are disproportionately held by the rich, the pattern of loans is likely to violate the public’s sense of vertical equity. The efficiency problem is that taxing capital gains on a realized basis alters the pattern of buying and selling of assets that would occur if the gains were taxed properly on an accrued basis. By taxing on a realized basis, investors have an incentive to “lock-in” the gains on successful assets (choose not to sell) to defer the tax payment and to sell unsuccessful assets early to deduct the losses against other sources of income. A related inefficiency is the incentive to take income as capital gains to defer the tax, an example being executives who take stock options in lieu of salary.

The inequities and inefficiencies notwithstanding, no one has seriously proposed taxing capital gains on an accrued basis. The difficulty comes with assessing the accrued gains on real assets that are infrequently traded. How much capital gain accrued last year on the house that has not been on the market since 1980 or the painting that has been hanging in the den since 1985? Tax authorities have no good way of estimating the gains (or losses) for these assets. Even if they could evaluate the accrued gains, people whose wealth consisted primarily of real assets may have to sell some of their assets to pay the accrued tax liability. The public would tend to view this as unfair. (An analogous situation is the elderly couple that is forced to sell their house they have lived in for 50 years because they can no longer pay the local property taxes out of their retirement pension.) For all these reasons, capital gains will almost certainly continue to be taxed on a realized basis.

In 1991, Alan Auerbach achieved a substantial breakthrough in solving the problems of taxing gains on a realized basis. He proposed a tax reform that avoids the lock-in and early sales effects by leaving investors always indifferent between: (1) holding an asset for one more period and (2) selling the asset and investing the after-tax proceeds in a risk-free asset for one period. His proposal also protects the government from making interest-free loans, at least on an expected value basis. The beauty of the Auerbach proposal is its practicality. It continues to tax capital gains on a realized basis and makes use of data that are readily available at the time the asset is sold.

**A Two-Period Example**

The following simple two-period example provides the intuition for the nature of the realization problem and how Auerbach proposes to overcome it. Consider two options for investing $1 at the start of the first period.

**Option 1:** Hold the asset for one period, realize the gain at the end of the period, and invest the after-tax proceeds in a risk-free asset during the second period.

**Option 2:** Hold the asset for two periods and then realize the capital gain over the two periods.

Assume:

\[ g = \text{the capital gain during the first period} \]
\[ i = \text{the one-period return on the risk-free asset} \]
\[ r = \text{the (uncertain) capital gain during the second period} \]
\[ t = \text{the income tax rate} \]

\[ Y'(t)_{\text{adjusted}} = CV - PV \cdot I(t) \quad (11.24) \]
Option 1: Sell and Invest Risk-Free Asset
The value of the asset at the end of the first period is $(1 + g)$. The realized gain $g$ is taxed at rate $t$, leaving net of tax proceeds of $[1 + g(1 - t)]$ to be invested in the risk-free asset during the second period. The proceeds grow at rate $i$ and the interest is taxed at rate $t$. Therefore, the net-of-tax value of the asset at the end of period 2 is 

$$[1 + g(1 - t)][1 + i(1 - t)]$$

For comparison with option 2, rewrite the net-of-tax value as 

$$[(1 + g) - tg][(1 + i) - it] = [(1 + g)(1 + i) - t[(1 + g)i + g(1 + i(1 - t))]]$$

The first bracketed term is the gross-of-tax value and the second the tax liability.

Option 2: Hold for Two Periods
The value of the asset at the end of period 2 is $(1 + g)(1 + r)$, and a tax is paid on the capital gain, leaving a net-of-tax value at the end of period 2 equal to:

$$[(1 + g)(1 + r) - t[(1 + g)(1 + r) - 1]]$$

$$= [(1 + g)(1 + r) - t[(1 + g)r + g]]$$

On comparing the two outcomes, note that $r$ is an uncertain return at the end of period 1. Assume that the certainty equivalent of $r$ is $i$. That is, investors are indifferent between investing at the uncertain return $r$ or the certain return $i$. Under this assumption, the certainty equivalent net-of-tax value of option 2 is 

$$(1 + g)(1 + r) - t[(1 + g)i + g]$$

Thus, option 2 is more valuable by the amount $(tg)(1 - t)$, equal to the after-tax interest on the portion of the accrued tax liability that is avoided by taxing the capital gain on a realized basis. The tax savings can be thought of as an interest-free loan by the government $(tg)$ made at the end of period 1. The taxpayer invests the risk-free loan at rate $i$ during the second period, pays a tax on the interest at rate $t$, and pays back the principal on the loan, for a net gain of $(tg)(1 - t)$, the after-tax interest on the loan.

The value to the asset holder of taxing on a realized basis follows the same pattern for any number of periods. The value equals the net-of-tax value on the current value of the taxes that would have been collected each year if capital gains were taxed on an accrued basis. (Note that the tax is paid once, when the asset is sold. The deferred tax liabilities, the “loans,” accumulate at untaxed interest until the sales date.)

The Vickrey Proposal
In 1939, William Vickrey proposed the following tax on capital gains to remove the interest-free loan advantage from taxing on a realized basis: Tax the gain in the final period on a realized basis, and add to the tax the interest on the current value of accrued tax liabilities to date, with the interest being tax deductible (Vickrey, 1939). The combined tax plus interest payment would make asset holders indifferent at any given time between holding the asset for one more period or selling the asset and investing the after-tax proceeds in a risk-free asset.

Under the Vickrey scheme, the instantaneous increase in the tax at time $s$ if the asset is held one more period is, in general,

$$\dot{T}_s = i(1 - t)T_s + trA_s$$

(11.25)

where

$T_s$ = the current value of the accumulated deferred tax liabilities to date at time $s$.

$r_s$ = the gain in period $s + 1$.

$A_s$ = the current value of the asset at time $s$.

In terms of the two-period example above, $s$ is the end of period 1, $T_0 = g$, and $trA_0 = tr(1 + g)$.24

The problem with Vickrey’s scheme is that it is as impractical as taxing on an accrued basis. It requires knowing the entire pattern of accrued tax liabilities to the time of sale, which is the same as knowing the entire pattern of gains. The current value of the total accrued taxes due on an asset held for 10 years is quite different if all the gains came in the first year, or in the last year, or evenly over time. In other words, the data requirements are the same as they would be under an accrued tax, data that would be unavailable for infrequently sold real assets.

The Auerbach Proposal
Auerbach (1991) proposed a variation of the Vickrey scheme that is practicable for all assets. It is based on the certainty equivalence operator, $V(\cdot)$, which gives the value that an investor would require, with certainty, to be indifferent to an uncertain return that is the argument of the function $V$. The idea is that investors make their portfolio choices prospectively. They are indifferent to holding an uncertain asset for one more period if the certainty equivalence of the after-tax return on the asset is equal to the risk-free after-tax return. In terms of the operator $V$, indifference requires that at time $s$

$$V(A_s - \dot{T}_s)/(A_s - T_s) = i(1 - t)$$

(11.26)

23. $i = E[r]$ under risk neutrality.

24. The part of the realized tax liability, $gt$ in the two-period example, is the tax on the first-period gain. It is not part of the tax increase.
Applying First-Best Principles of Taxation—What to Tax and How  

Chapter | 11  195

where \( \hat{A}_s \) is the uncertain next period return on the asset, and the other terms are as defined above. Multiplying both sides of Eqn (11.26) by \((A_s - T_s)\) yields

\[
V(\hat{A}_s - \hat{T}_s) = i(1 - t)[A_s - T_s]
\]  
(11.27)

\( V(\ ) \) is a linear operator. Therefore,

\[
V(\hat{A}_s) - V(\hat{T}_s) = iA_s - i(1 - t)T_s - itA_s
\]  
(11.28)

But \( V(\hat{A}_s) = iA_s \), the certainty equivalent next period return on the asset. Thus, indifference to holding or selling requires that

\[
V(\hat{T}_s) = i(1 - t)T_s + itA_s
\]  
(11.29)

Auerbach’s proposal is Vickrey’s proposal from an ex ante rather than an ex post perspective. \( V(\hat{T}_s) \), the ex ante certainty equivalence of the increase in the taxes, is the net-of-tax risk-free interest on the deferred tax liabilities plus the tax on the certainty equivalent return for the next period.

Auerbach proves that the required \( V(\hat{T}_s) \) is achieved if and only if the accumulated tax liability upon realization, \( T_s \), is

\[
T_s = (1 - e^{-is})A_s
\]  
(11.30)

Note that \( T_s \) depends only on current data at the time the asset is sold: the risk-free market interest rate, \( i \); the number of periods that the asset has been held, \( s \); the marginal tax rate, \( t \); and the current value of the asset, \( A_s \). The taxpayer could easily determine the tax liability by looking it up in a table. Note, also, that \( T_s = 0 \) when \( s = 0 \) (an asset bought and sold immediately yields no income and incurs no tax); and \( T_s \to A_s \) as \( s \to \infty \) (the accumulated deferred tax approaches the value of the asset as the holding period extends into the future without limit).

We will demonstrate the sufficient conditions and leave the necessary conditions to the interested reader.\(^{25}\) Suppose \( T_s = (1 - e^{-is})A_s \). Then, the instantaneous increase in taxes from holding one more period is

\[
\dot{T}_s = (1 - e^{-is})\dot{A}_s + it e^{-is}A_s
\]  
(11.31)

Add and subtract \( A_s \) to the RHS and multiply and divide the first term by \( A_s \):

\[
\dot{T}_s = (1 - e^{-is})\dot{A}_s/A_sA_s - (1 - e^{-is})it A_s + it A_s
\]  
(11.32)

\[
\dot{T}_s = (1 - e^{-is})A_s[(\dot{A}_s/A_s) - it] + it A_s
\]  
(11.33)

But, \( \dot{A}_s/A_s = i + e \), where \( e \) is a random variable with mean zero. Thus,

\[
\dot{T}_s = (1 - e^{-is})A_s[(i + e - it)] + it A_s
\]  
(11.34)

\[
\dot{T}_s = (1 - e^{-is})A_s[e + it(1 - t)] + it A_s
\]  
(11.35)

But \( V(\epsilon) = 0 \), by definition of the certainty equivalence operator. Therefore,

\[
V\dot{T}_s = (1 - e^{-is})A_s[i(1 - t)] + it A_s
\]  
(11.36)

or

\[
V\dot{T}_s = T_s[i(1 - t)] + it A_s
\]  
(11.37)

as required for investor indifference for holding the asset one more period or realizing and investing in the risk-free asset.\(^{26}\)

The only caveat to Auerbach’s proposal is that the tax is essentially a prospective tax because it is based on the certainty equivalence of the next period return rather than the actual return. Many proponents of income taxation tend to believe that the fair way to tax is on the ex post actual returns and not the ex ante expected returns. From the ex post perspective, exceptionally good assets are undertaxed and exceptionally poor assets are overtaxed under Auerbach’s proposal. Nonetheless, investors do base their decisions on prospective returns, so that Auerbach’s proposal does avoid the lock-in effect. Whether it is entirely fair or not depends on the ex post versus ex ante point of view, and this is largely a matter of taste.

Economists who favor expenditures taxes based on lifetime utility arguments tend to be indifferent between taxing on an ex post or ex ante basis. For example, they are indifferent between taxing the value of a house when it is purchased or the stream of housing services as they accrue, because the purchase price of the house equals the expected present value of the stream of housing services. The IRS would tax the value of the house when purchased under an expenditures tax because it is the only practical alternative. Whatever one’s view of its equity implications, Auerbach’s proposal for taxing-realized capital gains must be considered a landmark in the theory of tax design for having solved the capital gains lock-in problem in a practical manner.

### Capital Gains Taxation: A Postscript

Congress has never protected income from capital from inflation nor even remotely considered adopting Auerbach’s tax scheme. Instead, it has favored either excluding a portion of “long-term” capital gains, the gains on assets held for more than 1 year, thereby effectively taxing the

\(^{25}\) The necessary conditions establish that Eqn (11.30) is the only possible \( T_s \) that is a function of only \( i, s, t, \) and \( A_s \). See Auerbach (1991), pp. 172–173.

\(^{26}\) Auerbach also presents more complicated cases, such as the appropriate tax for indifference when there are both capital gains and dividends.
THE TAXATION OF HUMAN CAPITAL

Louis Kaplow (1996) has taken a provocative position regarding the appropriate taxation of wage income under an ideal income tax if one views wages as the returns to a person’s stock of human capital. His point is simply that physical and human capital should be treated identically under an ideal income tax. If this were done, however, it would lead to a sharp increase in the share of taxes collected from income received by labor.

Wages are not treated as returns to human capital in the standard Haig–Simons version of the ideal income tax presented earlier in this chapter. Instead wages are viewed as arising completely and concurrently with the supply of labor and are therefore taxed in full as they are realized each year, exactly as they would be taxed under an ideal wage or payroll tax. Viewing wages as returns to human capital would lead to very different tax treatment.

Compare, for example, the decision to save and the decision to invest in human capital. Under an ideal income tax, the saving is taxed immediately (not deducted from taxable income), and the returns to the saving are taxed as they accrue, whatever form they may take (e.g., interest, capital gains, a stream of returns from a depreciable asset). In the case of a real asset, the taxable returns are the gross returns less the depreciation on the asset each year, with the depreciation equal to the decline in the value of the asset. An investment in human capital is most directly equivalent to an investment in a real, depreciable asset. The initial investment costs should be taxed, that is, not deducted from taxable income. Also, the net returns from the investment—equal to the increase in the wages less the annual depreciation of the stock of human capital—should be taxed each year. Neither requirement is met under the standard income tax. Investment in the human capital may well be expensive, that is, deducted in full from taxable income, if it takes the form of lower wages received while participating in an on-the-job training program. Also, no deduction is allowed for the depreciation of a person’s human capital.

The wages are taxed in their entirety each year as they are realized. Notice that, from the human capital perspective, the full taxation of wages each year is completely wrong in the person’s last year of work because the stock of human capital necessarily depreciates to zero in the last year, and always by an amount equal to the wages earned in that year. The final-year tax liability should be zero under an ideal accrued income tax.

Kaplow takes the taxation of human capital one step further by assuming that all wages can be thought of as a return to human capital. Under this view, the stock of human capital is a gift received at birth that should be subjected to two forms of taxation if treated symmetrically to physical capital.

First, the receipt of a gift of physical capital, or of any financial asset that is ultimately a claim against the earnings of physical capital, is treated as income under an ideal income tax and subject to full taxation. Therefore, the initial gift of human capital should be treated as income and subject to full taxation at birth. The value of the gift is the present value of a person’s lifetime stream of wages less any expenses/investments incurred to generate the wage stream. In a world of perfect certainty, all future expenses/investments associated with the maintaining and increasing the stock of human capital would be known at birth, as would the entire stream of future wages arising from the human capital. The cash flow from the human capital would be lower in years in which future investments were made and higher in the noninvestment years. In other words, the initial gift of human capital at birth is its capacity to engage in certain kinds of investments in human capital throughout one’s lifetime, along with the lifetime wages that result from the investments.

In addition, any accrued income (net of depreciation) earned by the physical capital gift in subsequent years is subject to taxation each year. Similarly, the stream of wages each year net of depreciation resulting from the gift of human capital should also be taxed under an ideal income tax. Given the usual pattern of depreciation of human capital, the present value of the depreciation is likely to be less than the present value of the wage stream because wages will far exceed depreciation except in the last working years. Thus, the annual stream of wages and depreciation represents a second source of taxable income.

To summarize, the appropriate tax base for human capital under an ideal, accrued income tax consists of: (1) the initial gift of human capital at birth, equal to the present value of lifetime wages less any lifetime expenses/investments and (2) the annual stream of wages less the depreciation of the stock of human capital. This tax treatment is equivalent to the ideal tax treatment of a gift of physical capital.

In fact, gifts of physical capital (or financial assets) are stepped up in basis when passed on to heirs, so that the initial value of the capital escapes taxation. This is not supposed to happen under an ideal income tax, but because it does happen one could argue for exempting the initial gift of human capital from taxation. If so, then the tax base for
human capital is just the annual stream of wages less depreciation of the human capital stock. Since the standard "ideal" income tax calls for full taxation of wages, it actually overtaxes wage income when it is viewed as a return to human capital.

Suppose the income tax were reformed to include all gifts of capital as income, as called for by an ideal income tax. Then, if human capital escapes taxation at birth, the Auerbach/Vickrey method of retrospective taxation could be employed to capture the escaped tax liability when the wages (returns to human capital) are realized. The taxes due on the wages each year would include tax-deductible in-wages (returns to human capital) are realized. The taxes due on them at birth have been receiving implicit interest since birth on the taxes that should have been collected on those wages at birth. The later in life that the taxes occur, the higher the tax due on them, because the taxes due on them at birth have been receiving implicit interest tax free since then.

For example, the present value at birth of $1 of wages received at time $i$ equals $\frac{1}{(1 + ri)^i}$, where $r$ is the annual gross-of-tax interest rate, assumed constant over time. Had a tax been collected on that wage at birth, the value of the human capital at time $i$ would have been $(1 - t) \left[ \frac{(1 + ri)^i}{(1 + r)} \right]$, where $t$ is the tax rate and $r_i$ is the after-tax rate of interest, both assumed constant over time. Therefore, the current value of the taxes that should have been collected at birth, increased by the after-tax interest rate since birth, is

$$\left[ 1 - (1 - t) \frac{(1 + ri)^i}{(1 + r)} \right].$$

If the escaped taxes are to be collected retrospectively at rate $t$ in period $i$, then the $1$ of wages has to be scaled by

$$\left[ \frac{1}{(1 - t) \frac{(1 + ri)^i}{(1 + r)}} \right]$$

to collect this portion of the tax due under the ideal income tax.

Kaplow presents some calculations to show that wages received in the last few working years would have to be increased by a factor of 2–3 to capture retrospectively the escaped taxes since birth. Such scaling of the wages would be equivalent to scaling the returns on tax-deferred pension instruments such as IRAs if the taxes were collected retrospectively when the returns were realized. Under an ideal accrued income tax, savings for retirement should not be deducted from income as IRAs are.

Taxing human capital at birth, or scaling wages later on in life to account for taxes that should have been paid, would undoubtedly lead to horrendous problems of evaluation and liquidity—people might not trust how the tax liabilities were calculated or be able to pay the taxes when they are due. The fact remains, however, that Kaplow’s suggested treatment of human capital is the proper one under an ideal accrued income tax if wages are viewed as the returns to human capital.

The discussion so far has assumed perfect certainty. The taxation of uncertain income streams would be resolved as the uncertainty is resolved: Unexpected favorable (unfavorable) returns to human capital would increase (decrease) its value and the taxes due.

In conclusion, the only three ways that taxable income from human capital can arise are at birth, over time (the stream of wages less depreciation), and as uncertainties about future income streams are resolved. Proponents of the ideal taxation of physical capital should favor similar taxation of human capital if they view wages as a return to human capital. This point takes on special force given the widely cited estimate by James Davies and John Whalley that the stock of human capital in the United States is on the order of three times the stock of physical capital. If gifts of human and physical capital were counted as income as they should be, then the share of tax revenues collected from labor income would rise substantially.

**SUMMARY**

This chapter has emphasized that the problem of designing equitable broad-based taxes is one of the more vexing in all of public sector economics. First-best theory offers two guidelines for tax design: the interpersonal equity conditions of social welfare maximization and the ability-to-pay principle. The interpersonal equity conditions are preferred by the mainstream theory, yet the ability-to-pay principle has won the day in terms of informing tax policy. Even so, ability-to-pay principles are subject to various interpretations. Furthermore, even if ability-to-pay principles can be agreed upon, it is extremely difficult to determine who has actually gained or lost from a given tax system and who will gain or lose from particular tax reforms.

A brief review of the U.S. federal personal income tax served to highlight these problems. The tax pays lip service to the ability-to-pay principle on paper, but there are many slips in application. The chapter considered a number of reforms that would make the tax conform more closely to traditional ability-to-pay principles, such as removing certain exclusions and deductions. But we were forced to admit that these reforms would not necessarily make the tax more equitable under a proper utility-based interpretation of these same principles, since reforms themselves generate inequities. Equity in taxation is as difficult to achieve as equity in any other context.

27. Davies and Whalley (1991). Kaplan is definitely not proposing that the personal income tax be reformed to treat wages as returns to human capital. To the contrary, he does not believe that the ability-to-pay perspective is a useful addition to tax theory. He prefers the modern social welfare function perspective on taxes, transfers, and distributive equity generally which, as discussed earlier in the chapter, is concerned much more with issues of the tax structure (vertical equity) than with precisely defining the tax base.
REFERENCES