1. Introduction

There are many ways in which a history of ideas may be written. One is a strict chronological accounting, moving from year to year and encompassing the entire subject. Another is to proceed on an author by author basis, focusing on what the main contributors had to say. Then there is the option to select major themes and to see how they evolved, thus taking a number of runs down (or up) the time-path of doctrinal development. I have chosen the latter option as most appropriate to accommodate the wide range of issues to be covered in a history of the fiscal theory. They are here arranged in five essays, dealing with the theory of public expenditures (Section 2), equity in taxation (Section 3), efficiency in taxation (Section 4), shifting and incidence (Section 5), and macro aspects of fiscal policy (Section 6). Hopefully, this arrangement will set forth both the internal structure of fiscal doctrine and its historical development. But beyond this, there is no single or correct view of intellectual history. Such a history, in both selection and arrangement, is bound to reflect the perspective of the time in which it was written, as well as that of the author in whose prism past events are collected.

The history of fiscal doctrine, perhaps more than that of any other aspect of economics, carries a particular fascination. On the one hand, it reflects the advance of analytical economics, an enrichment of the tool box, to use Joan Robinson’s terms, which may then be applied to the solution of fiscal problems. As we shall see, the key fiscal tools were in fact forged by a line of great general theorists, not fiscal specialists. This line ranges from Smith over Ricardo, Mill, Depuit, Edgeworth, Wicksell, Pigou, and Keynes to Samuelson. The close linkages between general and fiscal theory is most evident for the analysis of tax incidence, which at each stage reflects the prevailing theory of price and distribution. The analysis of tax equity was affected profoundly by the growth of utility theory. The rise of Paretian welfare economics permitted the modern analysis of social goods. The advent of Keynesian economics placed the role of budget policy in a new perspective, and so forth.
But there is more to a history of fiscal doctrine than the tools of economic analysis. This history also responds to changing economic and social institutions. With the decline of feudalism, the property income of the Crown had to be replaced by taxation; and as the rise of modern legal and financial institutions came to be reflected in the tax structure, the complexity of tax structure analysis increased vastly. The growth of popular democracy, in turn, altered what is viewed as the appropriate range of governmental functions, with budget policy replacing the barricades as the area of struggle among group and class interests.

Changing social philosophies and values, finally, also have their bearing on the development of fiscal doctrine. The displacement of Lockean rules of entitlement by the utilitarian model of Bentham greatly altered the premises of tax justice; and the rise of egalitarian philosophy proved a major factor in the growth and significance of transfer payments. Single-minded economists may find these cross-currents disturbing, but to this observer they add sparkle to what even without them would be an intriguing story.

For reasons of space, but also substance, our analysis begins with Adam Smith and *The Wealth of Nations*. This, to be sure, is not the beginning of fiscal doctrine. The Physiocrats had their theory of taxation and the Cameralists had written explicitly on the administration of public finances. However, Smith offers a convenient point of departure to trace the emergence of modern thought. The major issues are already present and neatly arranged, from the duties of the Prince to provide public services to appropriate ways of raising the necessary revenue. What follows over the next two centuries are variations, if dramatic ones, on his essential theme. At the other end of the scale, we shall carry the story up to the 1960s, leaving off where other chapters in these volumes take over to present the current doctrine. Given our space constraint, our treatment has to be selective, both as to issues and authors. It is hoped, however, that the reader will be encouraged to pursue matters further on his own account.¹

### 2. Public goods

The core of fiscal theory addresses the question of what public services should be provided by the public sector, and how much. Our first task, therefore, is to examine how this question has been addressed as fiscal theory developed. Not surprisingly, it is here and in matters of tax equity that approaches have differed most widely, in contrast to the theory of incidence which does, and the newer role of fiscal policy which should, offer a more uniform body of thought.

¹Among histories of fiscal doctrine, see Seligman (1908, 1909), Myrdal (1929), Mann (1937), and Groves (1974). For selected readings, see Musgrave and Peacock, eds. (1958), and Musgrave and Shoup, eds. (1959).
2.1. The duties of the sovereign

The operation of the public sector, as developed by the classical economists, is seen in the context of a natural order which calls for reliance on and non-interference with the market. Public provision and taxation for its finance is called for only where exceptional circumstances demand it. A definition of these circumstances is attempted, but the tools for precise analysis were still lacking. An essential feature of the classical approach, still widely followed, is that the economics of expenditures and taxation are pursued as separate issues: while benefit taxation was viewed as the ideal, the bulk of tax revenue and hence tax analysis had to be examined in a context of ability to pay, with the required total set from the expenditure side.

2.1.1. Adam Smith

Adam Smith, in concluding his critique of mercantilist policy, sets out the “obvious and simple system of natural liberty that will establish itself of its own accord” once governmental restraints are withdrawn [Smith (1776, vol. II, p. 184)]. The prince, therefore, is “discharged from any duty of superintending the industry of private people”. But the system of natural liberty still needs the prince. It requires him to perform three duties, “duties of great importance indeed, but plain and intelligible to common understandings”. These duties are, first, to protect society against invasion from abroad; second, to protect every member of society from the injustice of every other member; and third, to provide certain institutions and public works. Having staked out these three functions, Smith examines each in detail. Interesting and cogent observations are offered on how these functions have grown and how they should be conducted, but the core issue of why they must be undertaken by the prince (read public sector) remains unsolved.

The discussion begins with defense. As the art of defense becomes more complex, self-protection becomes impossible and even reliance on a militia inadequate. Efficient defense, based on the division of labor, calls for a professional and standing army. But Smith does not go beyond this to question why this army must be maintained by the prince rather than by private providers. Next, Smith shows how the administration of justice is needed to safeguard life, limb, and property against internal offense. The more property there is, so he argues, the more inequality there will be, and the more costly will the protection of property become. Yet, its provision is essential to the functioning of the market system.

Smith then examines how justice should be administered to assure impartiality, and how its finance may be secured without burdening general revenue. Once
more he does not address why the administration of justice must be public rather than private. Perhaps no special explanation seemed needed to justify the public conduct of defense and justice. Though a proponent of the market, Smith was not an unbounded libertarian. He did not believe that civil society could be based on market forces only. Natural liberty requires a framework of security and legal rules, and government is needed to provide it. Smith’s view of the world may be read between the lines of the Wealth of Nations, but is developed in detail in his earlier work, The Theory of Moral Sentiments (1759). It is in this earlier work that Smith the moral philosopher presents an extremely complex and subtle structure of human interaction. Comprising a multiplicity of forces and motivations, individuals guided by the invisible hand are led to interact so as to produce a socially desirable outcome. In this interaction, benevolence as well as self-interest has an important role to play.

Turning to the expense for education, both university and elementary education are considered. The finance of university education is examined in its bearing on the quality of teaching, with scathing comments on how public financing and endowment lead to laziness and abuse once faculty is paid without relation to services rendered to their students [Smith (1776, vol. II, p. 249)]. Major concern, however, is with elementary education. Here public support is needed because division of labor, as it progresses, leads to monotonous and simple tasks, a specialized dexterity which is “acquired at the expense of the intellectual, social, and martial virtues” of the laboring population [Smith (1776, vol. II, p. 267)]. For moral as well as economic and military reasons, government must take some pain to prevent this process, and this calls for public provision of elementary education. Implicit in the argument is that education involves externalities, but once more the criteria of publicness is not addressed in explicit form.

The need to do so arises only when Smith turns to the provision of public works. Here a criteria had to be established to determine just which public works government should undertake. Such a criteria is offered in a preamble to the section on public works [Smith (1776, vol. II, p. 539)]:

“The third and last duty of the commonwealth is that of erecting and maintaining those public institutions and those public works which, though they may be in the highest degree advantageous to a great society, are, however, of such a nature that the profit could never repay the expense to any individual or small number of individuals, and which it therefore cannot be expected that any individual or small number of individuals should erect or maintain.”

As is evident from this passage, Smith recognized that market failure occurs in the provision of certain goods, goods which it does not repay the individual to provide. But we are not told just why this is the case. Nor do we receive an explanation in the more detailed discussion that follows, much of which focuses
on how finance can be arranged by fees, so as not to burden general revenue. We thus have to rely largely on the preamble. Much can be read into it, including the key concepts of joint consumption, externalities, and free-rider behavior which enter the modern view of social goods. But none of these are made explicit, so that it would be unduly generous to attribute them to Adam Smith. Nevertheless, his passage is not hostile and indeed amenable to these later developments. It contains the important premise that there exist certain functions which for objective (not ideological) reasons need be provided for by the public sector.

Unfortunately, however, Smith appears to have overlooked a strikingly insightful passage in Hume’s *Treatise of Human Nature* (1739), a treatise which had appeared over 30 years prior to the *Wealth of Nations*. Hume describes how two neighbors might agree to drain a meadow but how a thousand persons cannot reach such agreement, as each will try to lay the whole burden on others. Political society overcomes this difficulty by reliance on magistrates whose interest it is to reflect the interests of [Hume (1739, p. 539)]:

> “any considerable part of their subjects ... Thus bridges are built ... by the care of the government which tho' compos'd of men subject to all human infirmities, becomes, by one of the finest and most subtle inventions possible, a composition, which is, in some measure, exempted from all these infirmities.”

Here, as in other instances, an idea running ahead of its time finds no response until the situation is ripe, and then springs up at once in many places.

### 2.1.2. David Ricardo

Leaving Adam Smith, it is disappointing to find that nothing is to be said in this context about Ricardo. Although his concern with the affairs of the government was paramount, he dealt with the effects of taxation on the private sector only. There is nothing to be found on public expenditures in his treatise, except his approving quotation of the “golden maxim” of J.B. Say, “that the very best of all plans of finance is to spend little, and the best of all taxes is that which is least in amount” [Ricardo (1817, p. 159)]. The same view appears in his scathing rejection of Owen’s proposal for workhouse reform [Ricardo (Collected Works, vol. VIII, p. 46)].

### 2.1.3. John Stuart Mill

John Stuart Mill, our next author, viewed society through a quite different window. Concerned as he was with the works of the early Socialists such as
Fourier, Owen, and Sismondi, he addressed the proper scope of government in detail. Like Smith, he held that “laissez-faire should be the general practice: every departure from it, unless required by some great good, is a certain evil” [Mill (1848, p. 952)]. But also like Smith, he found important instances where a departure is called for. These instances are divided into “ordinary” and “optional” functions. The ordinary functions include above all provision for a legal system which secures life, limb, and property, a security which is prerequisite for the system of laissez-faire. But beyond this, “there is a multitude of cases in which governments, with general approbation, assume powers and execute functions for which no reason can be assigned except the simple one, that they conduce to general convenience”. As examples he cites not only coinage and the setting of standard weights, but also paving and lighting of streets, erection of harbors, lighthouses, and dykes. The range of functions seems to be left wide open, limited only by the rule that “interference of government should never be admitted but when the case of expediency is strong” [Mill (1848, p. 800)].

Though sceptical whether a specific principle of demarcation can be developed, he nevertheless tries to set forth the conditions under which the principle of laissez-faire and reliance on individual choice may be interfered with. Three such situations are noted: (1) Individuals may be unable to evaluate the utility of certain products. Thus children may be required to undergo elementary education. (2) Individuals lacking foresight may undertake irrevocable contracts and need to be restrained. (3) Regulations may be needed where individuals delegate decisions to managers, whose interests differ. Regulation of stock companies is thus called for, especially of monopolies in whose profits governments should share.

Coming closer to the economic content of publicness, he addresses “matters in which the interference of law is required, not to overrule the judgment of individuals – but to give effect to that judgment: they being unable to give effect to it except by concert, which concert again cannot be effectual unless it receives validity and sanction from the law” [Mill (1848, p. 963)]. Suppose, so he argues, that it is advantageous to reduce the working day from 10 to 9 hours. The individual worker cannot do so, nor can it be accomplished by agreement among them, as particular individuals will find it convenient to break the agreement. Hence legislation may be required. As a further illustration, he points to colonization. Individual colonies will wish to appropriate as much land as possible, while it would be to the advantage of the colonists to require newcomers initially to work as hired hands (the Wakefield system) so as to permit more intensive cultivation of specific parcels. Mill thus recognizes, implicitly at least, the existence of a prisoner’s dilemma and free-rider problem, conditions which require public intervention.

He also notes cases in which “acts done by individuals, though intended solely for their own benefit, involve consequences extending indefinitely beyond them,
Ch. 1: A Brief History of Fiscal Doctrine

to interests of the nation or posterity" [Mill (1848, p. 970)]. A case in point is again colonization, which requires regulation in line with the permanent welfare of the nation. The same principle, so he continues, "extends to a variety of cases wherein which important public services are to be performed, while yet there is no individual especially interested in performing them, nor would any adequate remuneration naturally or spontaneously attend their performance". A voyage of exploration might produce benefits of great public value, yet "there is no mode of intercepting the benefit on its way to those who profit by it, in order to levy a toll for the remuneration of its authors". Thus, "no one would build lighthouses from motives of personal interest, unless indemnified and rewarded from a compulsory levy by the state" [Mill (1848, p. 975)]. Mill's explanation why certain goods require public provision thus moves beyond Adam Smith's generalization, but still falls short of precise formulation. Emphasis is on the difficulty of collecting tolls, an argument also advanced subsequently by H. Sidgwick. Indeed, it was not until a hundred years later that the lighthouse problem [Coase (1974)] was placed in its proper perspective, i.e., that fee finance of social goods would be inefficient even if fees could be collected. Nevertheless, both Smith and Mill were aware that the nature of certain goods requires public provision, even though they assigned primacy to the market and flayed the inefficiency of governmental action.

2.2. The public economy

The traditions of British authors, from Adam Smith on, viewed the market as the rule and the public sector as the exception, needed to step in if and where a specific market failure occurs. The tradition of continental and in particular of German authors was to view the economic system in dual terms, with the public sector (Staatswirtschaft) equal in birthright to the private sector (Privatwirtschaft). This difference in emphasis had various roots. British fiscal theory emerged from the background of the Lockean model, a society based on individual entitlements and free exchange, guided by the beneficent rule of an invisible hand. The continental approach emerged from the cameralist teaching which had developed rules for the conduct of public affairs in the enlightened state. Kant's view of the state as limited to its productive function had been superceded by the Hegelian vision of the state as "immaterial capital"; and the Historical School's approach to economics, dominant in the closing decades of the 19th century, invited a view of growing state activity as a natural outcome of the historical process. Moreover, a sympathetic view of the public sector was supported in Germany by the rising concern of academic economists (the so-called pulpit socialists) with matters of social welfare.

Among major contributors to this view of public finances we may note Dietzel, Schäffle, and Wagner. Dietzel (1855) addressed the role of the state as a producer
of capital, both fixed and "immaterial", with public credit an important instrument of economic growth. Schäffle (1867) advanced his rule of "proportionate satisfaction" of public and private wants, anticipating Pigou's formulation of the mid-1920s. Wagner (1883), the leading figure of his time, formulated his law of expanding state activity, based on technical factors such as increased density and urbanization, as well as a growing acceptance of social-policy objectives in fiscal affairs. This line of thinking, moreover, exerted considerable influence on American scholars, who at that time tended to do graduate work in continental and particularly German universities. This influence is apparent as one compares the basic American text of Adams (1899) with its British counterpart of Bastable (1892).

2.3. Subjective value and public provision

But though the cameralist tradition of German authors had provided a more open-minded view of the public sector, it did not furnish an economic theory of public goods. Such a theory emerged only after the basis had been laid in the 1880s, when the analysis of subjective utility had grounded value theory on the demand side. This new approach, as developed by Menger and Jevons, was soon to find its application to the budget. Thereby the analysis of public provision was placed into an entirely new perspective. Focus was no longer on the duty of the sovereign, but on the demands of the individual consumer. The public sector appeared no longer as an awkward, albeit necessary, exception to the laws of economics. The same principles of efficient resource use were now to be applied to both the public and the private sphere. Integrated into the general theory of value, public sector economics was legitimized.

2.3.1. Marginal utility

The breakthrough emerged in the late 1880s from the contributions of Austrian and Italian writers, among whom Sax (1883), Panteleoni (1883), Mazzola (1890), and de Viti de Marco (1888) may be noted. While nuances differed, the essence of the new doctrine was this: Given the preferences of individuals, welfare is maximized by having each equate marginal utility with price. This basic efficiency rule applies to both public and private goods. To be sure, there is a difference: In the private good case, goods are sold at a uniform price, with individual consumers equating price and marginal utility by quantity adjustment. In the case

---

2 For references to these authors and excerpts translated from their major works, see Musgrave and Peacock, eds. (1958).
of public goods, the critical feature of indivisibility (already vaguely noted by Mill) requires the same quantity to be available to all consumers. Since the marginal utility of the same quantity differs among them, the equating process calls for differential prices to be charged. Thus benefit taxation—greatly broadened from its Hobbesian origin as payment for protection—becomes the “supreme law of the fiscal economy” [Musgrave and Peacock (1958, p. 81)].

With public expenditures linked to consumer evaluation, the basis for the modern theory of public goods was laid. But, not surprisingly, this early formulation had its shortcomings. By framing the efficiency rule in terms of benefit taxation, attention was diverted from specifying just how indivisibility affects efficiency conditions, conditions which may be met with or without benefit finance. It would be over half a century for these conditions to be worked out. Moreover, by focusing on the benefit rule as an analogue to market pricing, attention was diverted from the political, not market, process needed to reach an efficient solution. Not that the Italians were unaware of this problem. Mazzola noted that budgetary decisions are made by agencies, but held that agencies must act so as to satisfy voters, lest political equilibrium be disturbed [Musgrave and Peacock (1958, p. 44)]. De Marco, viewing the income tax as a subscription price, thought it to secure a fair solution. While the analogy to a market solution was central, the concept of a competitive political process was also present.

2.3.2. Knut Wicksell

This was the aspect of the problem which drew Wicksell’s attention and solicited his seminal contribution to the theory of public finance [Wicksell (1896), also see Musgrave and Peacock (1958, pp. 72–118)]. Wicksell accepted the new doctrine that provision of public goods should be designed to maximize individual satisfaction, and that the benefit rule would accomplish this. But then two concerns arise.

A first reservation relates to the equity implications of benefit taxation. Justice may be said to be served if consumers pay in line with their marginal evaluation but, so he ads, “it is clear that justice in taxation presupposes justices in the existing distribution of property and income” [Wicksell (1896, p. 143), Musgrave and Peacock (1958, p. 108)]. Wicksell thus views distributive justice as primary but then separates it from justice in taxation as payment for the cost of public services, a separation to which we will return later on.

Notwithstanding this qualified endorsement of the benefit rule, Wicksell did not consider it a realistic option. Indeed, he rejected as “really meaningless” [Musgrave and Peacock (1958, p. 81)] Mazzola’s new model. Analogy to the market, so he argued, was inapplicable since individuals would not reveal their preferences without injection of a political process. With large numbers the offer
made by any one individual has no significant effect on total supply so that consumers “will pay nothing whatsoever” [Musgrave and Peacock (1958, p. 81)]. Wicksell, as did Hume one hundred and fifty years earlier, thus clearly recognized what later came to be known as the free-rider principle. Impatient with the assumptions of altruistic and omnipotent behavior [Wicksell (1896, p. 90)], the real problem as he saw it was to design a practicable process which will approximate an optimal outcome. Ideally, consumers would be asked to vote on bundles of options combining a complete set of budgets and tax shares, with provision based on that bundle which carries unanimous support [Musgrave and Peacock (1958, p. 92)]. Since this ideal situation is impracticable, Wicksell settles for a rule of approximate unanimity, but stresses the need to protect minority rights. He thus laid the basis for a normative approach to voting models, a problem which has again become a subject of lively discussion in recent years.

2.3.3. The Lindahl price

The story continues with the appearance of Eric Lindahl’s doctoral dissertation [Lindahl (1918) written under the auspices of Wicksell].

Lindahl visualizes two consumers who must share in the cost of a public good. The more A pays, the less B will have to pay. Given the cost schedule for the product, A’s offer curve may thus from B’s point of view be translated into a supply curve, and vice versa. The two curves are plotted and their intersection determines the quantity to be supplied. At this solution each pays a tax price (the famous Lindahl price) equal to the value of the marginal utility which he derives, with the sum of the two tax prices adding up to the cost of the product [Musgrave and Peacock (1958, p. 89)].

One wonders how pleased Wicksell was with this formulation of his student. To be sure, the Lindahl price (i.e., benefit taxation) was efficient, but the market analogy inherent in the Lindahl diagram was precisely what Wicksell had objected to most in Mazzola’s presentation. While a bargaining solution might be reached in the small number case, it hardly need be the efficient one. More important, the analogy does not extend to the large number case, where preferences are not revealed, reducing the Lindahl schedules to “pseudo demand curves”, as referred to later on by Samuelson (1954). Lindahl, of course, was aware of these limitations. He notes that the intersection solution is reached only on the assumption of “equal bargaining power” and, in his later writings, expanded upon the political setting in which the budget determination occurs [Lindahl (1928), Musgrave and Peacock (1958, pp. 214–232)]. Nevertheless, the concept of the Lindahl price and its initial demand–supply presentation has remained his key contribution.

For excerpts, see Musgrave and Peacock (1958).
2.4. Pigouvian externalities

The Austrian and Italian model of fiscal analysis and its Wicksellian interpretation in voting behavior did not enter the purview of English language authors for over half a century. Continuing the classical tradition, an ad hoc approach to the delineation of expenditure functions prevailed from Mill to the 1920s. Marshall had little to say on the subject, nor did Jevons apply his marginal analysis to the public sector. Bastable's (1892) *Public Finance*, the major English language treatise, offered little advance over Mill, even in its later edition (1903). Various categories of public expenditures are discussed but little attention is paid to the nature of publicness and its bearing on an efficient solution.

Thus it was not until Pigou's (1920) *Economics of Welfare* that a new perspective was introduced. This perspective emerged from the concept of externalities, central to the Pigouvian distinction between social and private net product [Pigou (1920, ch. 9)]. The private net product measures the internalized costs and benefits which are recorded in market price. The social net product will be larger when further benefits accrue to persons other than those engaged in the sales transaction; and it will be smaller if costs are imposed on third parties, costs which need not be compensated for and hence are not reflected in price. Where social benefits are in excess of private, a bounty need be paid to allow for the addition (external) benefits which are not reflected in market demand. Where social costs exceed private cost, a tax is in order. Thus fiscal instruments become a mechanism of adjusting for externalities, be they of the benefit or cost type.

Where a bounty is appropriate, its magnitude will depend on the spread between private and social product. Thus Pigou notes that a moderate bounty to farming may be suitable if farming yields the "indirect service of developing citizens suitable for military training ... A more extreme form of bounty, in which a governmental authority provides all the funds required, is given upon such services as the planning of towns, police administration, and, sometimes, the clearing of slum areas." Having advanced this close to providing a criteria for public provision, it is frustrating to find that Pigou does not proceed to do so. Yet the logic of this argument suggests that public provision (i.e., a full bounty based on tax finance) is needed where the private net product is zero and the social net product absorbs the entire value.

Nor did Pigou (1928) offer an explicit linkage between externalities and social goods in his subsequent *Study in Public Finances*. This volume, as we shall see presently, made major contributions on the theory of taxation, but gave only brief attention to the expenditure side of the budget. A distinction is drawn between transfers and exhaustive expenditures leading to a principle of expenditure allocation, similar to Schäffle's (1880) much earlier rule of proportional satisfaction. Within a given budget size, the program mix is to be adjusted so as to balance the marginal benefits derived from various projects; and in determining
the size of the budget, the marginal benefits from public outlays are to be equated
with those from private outlays. Such would be the simple formula if the
community was a unitary being [Pigou (1928, p. 52)]. Since it is not, and since the
desire of any one taxpayer to contribute depends on the contribution of others,
government, as the agent of its citizens collectively, must exercise coercion upon
them individually. This coercion creates indirect costs which must be allowed for.
But though the need for coercion is noted, there is no consideration of the
mechanism by which it will inform government how individuals value social
goods. Pigou, evidently, was still unaware of what had been contributed by the
Austrians and Swedes three decades earlier. His contribution, however, was to
place the extreme case of "full bounty" into the broader spectrum of externalities,
which may differ in degrees and kind, i.e., may range from "mixed" social goods
to the polar case of full external benefit.

2.5. Pareto efficiency with public goods

It was not until 1939 that the continental discussion of the 1880s and 1890s was
brought to the attention of the English (and only English) reading part of the
profession [Musgrave (1938)]. Howard Bowen’s (1948) vertical addition of de-
mand curves for social goods reinvented Lindahl's earlier formulation, and the
breakthrough came with Samuelson’s (1954,1955) two three-page papers on the
subject. Carrying a benefit-page ratio without rival, at least in the literature of
fiscal theory, these papers met the long-delayed need for a rigorous integration of
social goods into the conditions of Pareto efficiency. Thirty years later, his
solution may seem evident to the well trained senior, but at the time it offered a
giant step forward.

2.5.1. The Samuelson model

Addressing the implications of indivisibility and joint consumption for Pareto
efficiency, Samuelson assumed that there exists an omniscient referee to whom
individual preferences are known. Based on this information, given resources and
technology, the referee then determines a set of optimal solutions, each involving
a mix of output between private and social goods and a division of the former
among consumers. Each solution reflects different positions of welfare for particu-
lar consumers with the optimal solution or "bliss point", chosen by application of
a social welfare function. The set of efficient solutions meets the condition of
equality between the sum of differing marginal rates of substitution in consump-
tion and the rate of transformation in production. It differs from the case of
private goods where the marginal rate of substitution is the same for all consumers and equal to the rate of transformation. Social goods are thus amenable to the same (if somewhat altered) principles of efficiency as private goods, to be dealt with as a subcase of the Paretian rule.

How did this formulation relate to the Wicksellian system and its use of benefit taxation as an efficiency rule? The two approaches are similar in that both yield efficient solutions. The Lindahl price, approximated by the Wicksellian voting system, is also among the solutions which meet Samuelson’s efficiency conditions. But the approaches differ in dealing with distribution.

For Wicksell, a distribution of money income and charging of tax prices is essential. Preferences must be determined by voting, in line with a pattern of effective demand based on a given distribution of money income. In Samuelson’s (1969) model, money income and taxes may be inserted, but they are not needed and only clutter up the problem. Given an omniscient referee whose preferences are known, the referee may proceed directly to the solution. Having determined the set of efficient solutions, he then resolves the distribution problem by applying a social welfare function. Optimal distribution is determined in terms of welfare positions. The Wicksellian model begins as just noted with a distribution of money income; and this distribution must be just to begin with if the voting process is to arrive at tax prices which are “just” as well as efficient. But how can a just distribution of money income be taken as predetermined when the distribution of welfare (which is what matters) depends on relative prices (including those of public goods) as well as on money incomes? The appearance of circular reasoning is resolved, however, once the voting rule is allowed for in determining the distribution of money income, and determination of the voting rule is added as a further equation in the system [Musgrave (1969; 1984, p. 67)]. Multiple policy objectives may then be resolved in an interdependent fashion, including a benefit-tax based allocation branch which provides for social goods, and a tax-transfer based distribution branch which secures the desired distribution of money income [Musgrave (1959, ch. 1)].

2.5.2. Extensions and additions

The modern discussion of budget determination thus involves two traditions, both valid and compatible with each other but addressing different aspects of the problem. The analytical neatness and abstract formulation of Samuelson’s model meets the pure spirit of Paretian welfare economics and as such has invited the attention of economic theorists. The greater realism of the Wicksellian approach has offered a more workable stepping stone to the problems of budget policy.
Subsequent work has drawn on both traditions. Among the most important extensions, the following are noted briefly, leaving further discussion to subsequent chapters of this volume where the current state of the art is examined:

1. Extension of the Samuelson model has focused on examination of non-polar cases, such as benefit spillover and congestion [Oakland (1971, 1972)], as well as on the conditions under which private supply of public goods is feasible [Demsetz (1970)].

2. The Pigouvian analysis of external effects, and in particular external costs, has become a central aspect of growing concern with pollution and environmental economics [Baumol and Oates (1975)].

3. Much attention has been given to the analysis of local public goods and the implications of spatial benefit limitation [Tiebout (1956)]. Based on Tiebout's hypothesis of voting by feet, the feasibility of Tiebout equilibria has been examined and tested in empirical work [Oates (1972)]. A theory of fiscal federalism based on a theory of clubs [Buchanan (1965)] was developed.

4. Empirical research has been directed at estimating expenditure functions for government, based on the hypothesis that the political process approximates the preferences of consumers as expressed by their voting behavior. Utilizing median voter models [Downs (1966)] and cross-section data for state and local government, such research has become a rich area of fiscal analysis [Inman (1978)].

5. Wicksell's primary concern with the political mechanisms by which fiscal decisions are reached was extended in the context of voting theory [Black (1948)], and a model of democracy viewed in analogy to a political market [Downs (1955)]. Early optimism was dampened by Arrow's demonstration that an unambiguous social welfare function cannot readily be determined [Arrow (1951)], and fiscal issues played a central role in the rapid development of public choice theory [Mueller (1979)].

6. As distinct from normative considerations, a new direction of fiscal analysis emerged as a positive theory of government behavior [Buchanan and Tullock (1962)]. Critical of public sector growth, the new model centered on the propositions (a) that the voting process is biased towards overexpansion, and (b) that this bias is accentuated by the desire of bureaucrats and politicians to maximize their budgets [Buchanan (1975), Niskanen (1971), Borcherding (1977)]. The role of government is seen not as a servant of majority preferences but as a self-interested actor in its own right. Attention shifted to a theory of "government failure" and resulting need for governmental restraint counteracting in Hegelian fashion the earlier concern (underlying Pigouvian externalities and public goods theory) with market failure and the need for remedy by governmental action. This analysis, though very different in content, might be viewed as a resumption of interest in fiscal sociology which, fifty years earlier, had been pursued in the context of Marxist thought [Goldscheid (1917), see also Musgrave and Peacock (1958)].
2.6. Cost–benefit analysis

It remains to note a further and more operational approach to the economics of public expenditures. Moving to the forefront of fiscal work in the 1960s cost–benefit analysis was designed to provide a practical basis for evaluating “public works and development projects of a sort for which measures of value can be established empirically” [Eckstein (1961, p. 440)]. This episode thus matched the partial approach of tax theorists who, as we shall see presently, dealt with the tax side of the budget only, while disregarding the expenditure side [Musgrave (1969, p. 103)].

While cost–benefit analysis became the vogue in the 1960s, its beginnings, in important part, date back over a century to the work of Jules Dupuit. Drawn to the problem as an engineer in the Corps des Ponts et Chaussées, he was the first to pursue cost–benefit analysis on a rigorous basis and indeed anticipated the essence of much that was to come later [Dupuit (1844), Vickrey (1968)]. In particular, he developed the concept that benefits are measured by the area under the demand curve, not by what is actually paid. The next major contribution appeared nearly one hundred years later when Hotelling (1938, p. 158) formulated the case for marginal cost pricing: “The efficient way to operate a bridge”, so he argued, “is to make it free to the public, so long at least as the use of it does not increase to a state of overcrowding.” The common assumption “that every tub must stand on its own bottom” is rejected and the rationale for subsidies to increasing cost industries is given. Beginning in the late thirties, the U.S. government adopted standards for cost–benefit analysis of water projects, and studies of cost effectiveness blossomed in the Department of Defense during World War II.

Beginning with the late fifties, an extensive literature on cost–benefits analysis emerged and became the vogue of the 1960s [Layard (1972)]. Theoretical interest centered on problems such as the choice of an appropriate discount rate, measuring the opportunity cost of capital withdrawal from the private sector and the introduction of shadow prices [Margin (1963), Harberger (1969), Feldstein (1973)]. By the end of the sixties or early seventies, the major analytical issues had been resolved and cost–benefit analysis had become an important tool of applied fiscal analysis.

3. Equity in taxation

In the preceding section, our focus has been on the development of expenditure theory, a development which in large part proceeded independent of the taxation side of the fiscal process. The classics first examined the obligations of the prince
and then turned to tax analysis. With benefit taxation applicable to only a small part of the revenue total, criteria of "good taxation" for most of the revenue had to be developed independent of expenditures. Pigou once more dealt with the expenditure side in terms of externalities, then took up taxation theory as a separate issue. Samuelson's model, to be sure, covered both the uses and sources side of the fiscal picture, but at a level of abstraction which did not involve tax institutions. Thus only the Austrian approach and its Wicksellian extension opted for a simultaneous solution to both sides of the budget equation. Taxation theory similarly developed largely in isolation from the expenditure side. Such was the case with the classics, with the Schanz–Simon approach to income taxation and now replays in the context of optimal taxation theory.

Criteria for "good taxation" found an early statement in Smith's (1776, vol. II, p. 310) famous "maxims". Among them Smith includes equality, certainty, convenience of payment, and economy in collection as most important. Equality or equity in turn was interpreted along two lines, i.e., that contributions should match benefits received, and should also reflect ability to pay. In both camps, there emerged a long debate over whether burden distribution should be proportional or progressive, with the ability doctrine more open to egalitarian interpretation. Moreover, it was necessary to specify an index by which benefit and ability to pay is to be measured.

3.1. The benefit doctrine

The benefit approach to tax equity was congenial to the political philosophers of the enlightenment, such as Hobbes, Grotius, and Locke. With legitimacy vested in the hand of the governed, the contractarian model would call upon them to pay the state for protection received. Under the Lockean concept of entitlement, each person had property in the fruits of his labor [Locke (1690, p. 327)], an entitlement compatible with taxation as payment for services rendered but not with state-taking on other than a quid-pro-quo basis. Smith's (1759) grand design for the human condition, as painted in his earlier work, squarely fitted this pattern.

The first of his famous maxims of taxation states the rule of tax equity as follows [Smith (1776, vol. II, p. 310)]:

"The subject of every state ought to contribute towards the support of the government, as nearly as possible, in proportion to their respective abilities; that is, in proportion to the revenue which they respectively enjoy under the protection of the state. The expense of government to the individuals of a great nation, is like the expense of management to the joint tenants of a great estate, who are all obliged to contribute in proportion to their respective..."
interests in the estate. In the observation or neglect of this maxim consists, what is called the equality or inequality of taxation."

While the maxim begins with ability to pay, its thrust is in the direction of a benefit rule. As stated in the bottom line, equality in taxation calls for payment in proportion to one's interest in the public estate. Placing Smith in the benefit camp, while somewhat controversial, also matches his repeated call for fee-finance in the expenditure chapters, including a timely admonition that professors be paid in line with student attendance [Smith (1776, vol. II, p. 249)].

While contribution is to be in proportion to revenue received, Smith then qualifies this in various ways. Most important, he excludes wage income needed for subsistence. A tax on subsistence wages (as we shall see below) would have to be absorbed by higher-income consumers or by landlords. Its imposition would be "absurd and destructive" [Smith (1776, vol. I, p. 350)]. He thus shares the view, frequently held by early advocates of proportional taxation, that a subsistence minimum (then necessarily in the form of wages) should be exempt. A further exemption arises in connection with a tax on house rents. Smith supports such a tax, even though it would in general fall upon the rich, adding "that in this sort of inequality there would perhaps not be anything very unreasonable. It is not very unreasonable that the rich should contribute in proportion to their revenue but something more than in that proportion" [Smith (1776, vol. II, p. 337)]. Smith once more is too wise a man to permit neat classification as an all-out advocate of proportional taxation.

Benefit theorists, however, were far from unanimous in support of any particular pattern of rates. Most of the early contractarians, such as Hobbes and Grotius, supported the proportional view. So did Bentham, subject to the exemption of subsistence wages. Sismondi, Rousseau, and Condorcet, among others, favored progression. Robespierre rejected progression as insulting to the poor, while John Stuart Mill (though rejecting the benefit rule) thought it to call for regression: the poor, so he noted, are most in need of protection and thus would have to pay most [Mill (1921, p. 805)]. The question, who benefits most, it appears, was not easily resolved. [See Seligman (1908, part II, ch. 2).]

In past no less than current controversy, disagreement frequently reflects a difference in the question that is asked; but, unhappily, the question is not readily defined until the problem is resolved. The issue, in the interpretation of the benefit rule, is whether focus is on the cost of the service rendered to a particular person, or whether it is on what a person (given his or her income and preferences) would be willing to pay. In the latter case, the benefit tax as we have seen becomes a Lindahl price; and the issue of progressive taxation then hinges on the income and price elasticities of demand, factors which depend on the particular services in question and cannot be generalized upon for the budget as a whole [Buchanan (1964)]. Others have questioned whether Lindahl pricing meets
the spirit of the benefit rule. Thus Myrdal has suggested that benefit taxation should be in line with total (not just marginal) benefit received. Interpreted in this fashion, benefit taxation would still be “efficient” but would yield surplus revenue which would then have to be disposed of in other ways [Myrdal (1953)].

3.2. The ability to pay doctrine

The ability to pay doctrine also claims a long history. As with the benefit doctrine, views on the resulting burden distribution varied. Montesqieu and Say found in favor of progression, while early observers such as Bodin deduced a proportional rate [Seligman (1908, part I, ch. 3)].

The modern doctrine may be said to begin with John Stuart Mill’s formulation. Writing in the 1840s, Mill responded to a wholly different philosophical and political setting than had Adam Smith. Representative government had progressed and the accepted functions of the state had broadened. Bentham’s utilitarian framework had replaced natural law and the Lockean view of entitlement had given way to a new concept of distributive justice. The case had been made as early as 1802, that welfare from a given income total would be maximized by equal distribution [Bentham (1830)]; and the concept of justice, to quote Mill (1848, p. 805), came to “consist not in limiting but in redressing the inequalities and wrongs of nature”. Moreover, “government must be regarded as so preeminently the concern for all, that to determine who is most interested in it is of no real importance”. The just pattern of taxation, therefore, is not to be derived from the expenditure side of the budget, but is to be based on a general rule of social justice.

Mill (1948, p. 804) defines equal treatment as follows:

“For what reason ought equality to be the rule in matters of taxation? For the reason that it ought to be so in all affairs of government. As a government ought to make no distinction of persons or classes in the strength of their claim on it, whatever sacrifices or claims it requires from them should be made to bear as nearly as possible with the same pressure upon all, which it must be observed, is the mode by which least sacrifice is occasioned on the whole ... means equality of sacrifice.”

Mill then interpreted equal sacrifice to call for a proportional tax on income above subsistence, in line with Pitt’s income tax and its 3 percent rate. Following Bentham, he feared the disincentives of progressive income tax rates but (unlike Bentham) he favored sharp progression in inheritance taxation. While mistaken in equating “equal” with “least total” sacrifice, his concern with the latter anticipated the subsequent shift from equity to efficiency aspects of sacrifice doctrine.
His main contribution was to have posed the problem of equity in terms of equal sacrifice, thus setting the framework for half a century of subsequent discussion. Nearly fifty years later, the argument was resumed by F.Y. Edgeworth. Edgeworth (1897, p. 101) begins with "the purest, as being the most deductive form of utilitarianism, from which Bentham reasoned down to equality". He then deduces equal marginal or least total sacrifice as the optimal solution, not as a principle of distributive justice but on a game theoretical basis. Edgeworth (1897, p. 103) considers two self-interested parties, contracting in the absence of competition:

"In this setting, neither party in the long run can expect to obtain the larger share of the total welfare... Of all principles of distribution which would afford him now a greater now a smaller proportion of the sum-total utility obtainable on each occasion, the principle that the collective utility should be on each occasion a maximum is most likely to afford the greatest utility in the long run to him individually... The solution to this problem in the abstract is that the richer should be taxed for the benefit of the poorer up to the point at which a complete equality of fortunes is attained."

He does not explain why the individual player should agree to this solution, but intuition might have suggested an argument in line with Hasyani's (1953) maximization of expected utility under the veil of ignorance.

Edgeworth thus not only applied the equal marginal sacrifice rule to distributing the cost of public services but he extended it into a system of transfers resulting in an equal distribution of income. But having thus sighted the "acme of socialism", he finds it "immediately clouded over by doubts and reservations" [Edgeworth (1897, p. 104)]. The detrimental effects of the extreme solution are noted, including reduction in output and, here quoting Mill, threats to personal liberty. Thus "minimum sacrifice, the direct emanation of pure utilitarianism, is the sovereign principle of taxation; [but] it requires no doubt to be limited in practice" [Edgeworth (1897, p. 106)].

In addition to minimum sacrifice, two other species of the "hedonisti theory of taxation" enter. These are the rules of equal absolute and equal proportional sacrifice, as first explored by the Dutch economist Cohen Stuart [Stuart (1889), Musgrave and Peacock (1958)]. Given Bernoulli's assumption of a unit-elastic marginal utility of income curve, the two rules call for proportional and progressive taxation, respectively. More or less elastic schedules respectively call for progressive and regressive rates for equal absolute sacrifice, with proportional sacrifice not amenable to a simple solution of this sort. Suspecting the decline of utility to be more rapid than Bernoulli suggests, Edgeworth concluded that both rules require progression, but once more we are warned that "other utilities" must be allowed for as well, and that this prohibits rigid application of either sacrifice rule.
Writing as a contemporary of Edgeworth, a generally similar position was taken by Henry Sidgwick. Benefit charges should be applied where possible, but the principle of payment in line with services received is of limited applicability. Where it cannot be applied, the "obviously equitable principle – assuming that the existing distribution of wealth is accepted as just or not unjust – is that equal sacrifice should be imposed on all" [Sidgwick (1883, p. 562)]. He takes this to call for exemption of a minimum income, but hesitates to favor progressive rates as these may become excessive and unduly depress capital formation. But given that concern with capital formation is the reason for limiting progression, a case is made for the exemption of saving and the taxation of luxury goods.

Along with this debate among British economists, the case for progressive taxation was argued by Adolf Wagner, the German fiscal economist noted earlier for his "law of expanding state activity". Also writing toward the close of the nineteenth century, Wagner distinguished between (1) a purely financial and (2) a social welfare principle (sozialpolitisches Prinzip) of taxation. The former calls for proportional taxation in the finance of public services, while leaving the distribution of income unchanged. It is then supplemented by the latter, which calls for progression to reduce income inequality. With the development of modern society, Wagner expected this development to expand and viewed it with less concern than did his English contemporaries [Wagner (1883)].

The discussion of ability to pay then lapsed, but was resumed in Pigou’s (1928) *Studies in Public Finance*. Pigou accepted least sacrifice as an absolute principle of taxation, viewing it as counterpart to the general rule that public policy should be directed at maximizing welfare. He explored tax formulae applicable to various equity rules and arrived at conclusions similar to those of Edgeworth. Like Edgeworth, he found no convincing basis on which to choose between equal absolute and equal proportional sacrifice, with preference again given to the equal marginal sacrifice rule.

Conclusions regarding progressivity, from Edgeworth to Pigou, had been based on the assumptions (1) that utility is comparable across individuals and measurable in cardinal terms, (2) that there exists a known marginal income utility schedule, (3) that this schedule shows marginal utility to decline with income, and (4) is identical for all people. Doubts regarding (3) and (4) had been raised by Edgeworth and Pigou, but the fundamental break with (1) did not come until the 1930s. At that time, the feasibility of inter-personal utility comparison was rejected [Robbins (1932, 1938)], and modern welfare economics was restated in terms of Pareto optimality. A welfare gain could be recorded only if there was an improvement in A’s position without worsening that of B or, less demanding, if A’s gain was sufficient to permit potential compensation of B. Having thus advanced (or retreated) to the impeccable shores of Pareto efficiency, the rug was pulled out from under the older sacrifice doctrine. The distribution of the tax burden, henceforth, would be a matter of social ethics or politics, but no longer of economics.
But the issue of distribution proved too basic to be excluded from economics forever, and the distribution of the tax burden proved too central a part of Public Finance to be permanently expunged from its books. Distributional considerations soon reentered via a social welfare function [Bergson (1938)]. Assuming the shape of such a function to be agreed upon by society, it was then applied to determine the "bliss point" on the utility frontier [Samuelson (1954)]. It soon reappeared in the use of distributional weights in cost–benefit analysis [Weisbrod (1968)] and more recently in optimal income tax rates. By postulating a welfare function which reflects social value judgment (arrived at by a political process as based on the social preference of individual voters), the disputed premise of cardinal measurement and comparability is avoided. But the heart of the matter, we venture to suggest, was not changed as much as is commonly thought.

In concluding this review of equity doctrines, we note that while the ability to pay view of tax equity does not allow for benefit considerations, ability to pay considerations may enter into the benefit doctrine. They will not do so as long as benefit differentials are viewed in terms of differences in service levels provided to the rich and the poor. Thus, as Adam Smith saw it, the rich should pay more because they have more carriages to protect than the poor. But ability to pay enters once it is seen that the rich will value the benefit per carriage more highly than do the poor, as reflected in differential Lindahl prices. The later interpretation, and with it a linkage between benefit and ability to pay doctrines, also enters in the Wicksellian context where tax prices are needed for purposes of preference revelation.

3.3. The index of equality

So far we have traced the debate over how taxpayers with differing levels of capacity should be taxed. This leaves open the question of how this capacity should be measured. The answer has to be viewed in historical terms, as it depends on the prevailing economic institutions and objects of taxation which may be taken as representative of ability to pay. The answer also depends on the availability of "tax handles", i.e., objects of taxation which can be reached by fiscal administration. From the Middle Ages to the Elizabethan poor laws, "faculty" had been interpreted as property and this was still the case in early colonial taxation. Specific forms of property, such as cattle, windows, or carriages in turn served as indicies for property at large. There then occurred a gradual shift to a broader view of property; and beyond it, faculty came to be interpreted as income, viewing the tax base in terms of flows rather than stocks. The development of tax bases may thus be seen to reflect the institutional changes which accompanied the rise of modern industrial and financial society. Schumpeter (1918) in particular viewed the rise of the income tax as a corollary to the growth of capitalism and a pecuniary economy. The increasing complexity of economic
institutions in turn was reflected in the emerging technical discussion over how the specifics of taxation income should be defined.

3.3.1. The income base

Income, as measure of taxable capacity, dates back to Adam Smith and before. But the concept of the comprehensive income base emerged only slowly. Certain sources of income, it appeared, could not be taxed, even if an attempt were made to do so. Thus the Physiocrats viewed the rent of land as the only feasible tax base, simply because land was considered the only genuine source of income [Quesnay (1760)]. The classics, from Smith to Mill, took a broader view of the income base, but still thought it useless to tax subsistence wages or necessities, as such taxes would have to be passed on to rent or profits. Thus only rent, profits, and income flowing into luxury consumption remained as eligible sources of taxation. With taxation of profits involving undue meddling, causing capital flight [Smith (1776, vol. II, p. 333)], and detrimental to growth [Ricardo (1817, p. 94)], the viable tax sources were reduced to rent and income used for luxury consumption.

The modern idea of a comprehensive and global income tax, as the best index of equality and taxable capacity, only emerged towards the close of the nineteenth century. Intensive discussion of the income concept, especially by German authors, then led to the concept of accretion, first proposed by Georg Schanz (1896) and subsequently introduced into the American literature by Haig (1921). With a person's income defined as the money value of net accretion to his/her economic power, the measurement of taxable income was made tax-specific and distinguished from the concept of income shares in the context of national income accounting. Pioneered by Neumark (1947) in Germany and Simons (1938, 1950) in the United States, the development of a broad-based income tax came to occupy much of the tax literature over the following decades. With accretion as the guiding concept, specific issues of tax base measurement could be dealt with in coherent form, covering such items as the treatment of capital gains independent of realization, the integration of corporate source income into the income tax base, the economic measurement of depreciation, and so forth [Seligman (1914), Vickrey (1947), Pechman (1959), Musgrave (1967), Shoup (1969), Goode (1977)]. The comprehensive income tax base thus became the banner of tax reform in the United States, designed to secure equal treatment of taxpayers with equal income (horizontal equity) as well as to provide a global base on which progressive rates could be assessed in a meaningful fashion (vertical equity). How much impact this movement has had on the actual income tax is a different matter, but it clearly provided the focus of analysis and delight for a generation of tax economists in the United States.
In contrast, British public finance literature was reluctant to embrace the accretion concept. The schedular approach to the income tax was deep-rooted. Ursula Hicks (1947), following J.R. Hicks’ analysis in *Value and Capital*, still defined income so as to exclude windfall gains. Similarly, a Royal Commission of the mid-fifties rejected the usefulness of establishing criteria for an income concept, with a minority view in favor of the accretion approach joined by Kaldor (1955). More recently, however, the accretion concept and comprehensive base approach has gained acceptance [Prest (1975), Kay and King (1978)], but with it there emerged an alternative approach to broad-based taxation in the form of consumption.

3.3.2. The consumption base

Though emphasis among tax economists had traditionally been on the income base, the case for consumption as index of equality also claims a long ancestry. Thus, over three centuries ago, Hobbes (1651, p. 386) stated the equity case for the consumption base as follows:

"Equality of imposition, consisteth rather in the equality of that which is consumed, than of the riches of the person who consumes the same. For what reason is there, that he which laboureth much, and sparing the fruits of his labour, consumeth little should be more changed, than he that living idly, getteth little and spendeth all he gets; seeing the one has no more protection from the Common-wealth, than the other."

Adam Smith, while featuring income in his first maxim, subsequently retreated to rent and luxury consumption as the appropriate bases. Ricardo, bypassing the issue of equity, prevailed against the taxation of income which would be returned to capital. It was thus left to John Stuart Mill to lay the modern basis for the consumption-base doctrine. Beginning with an income-based view of sacrifice, Mill rejected preferential treatment of temporary as against permanent income, but not all income was to be treated alike. Suppose there are two people with the same income, so he argued. One is a wage earner, has no capital and must save for old age. The other has interest income and need not save as his capital will provide for retirement. The wage earner, so Mill (1848, p. 813) argued, has less left for consumption, and to treat him equally, his savings should be omitted from his tax base:

"If, indeed, reliance could be placed on the conscience of the contributors, or sufficient security taken for the correctness of their statements by collateral precautions, the proper mode of assessing an income tax would be to tax only that part of income devoted to expenditure, exempting that which is
saved. For when saved and invested (and all savings, speaking generally, are
invested) it thenceforth pays income tax on the interest and profits which it
brings, notwithstanding that it has already been taxed on the principal.
Unless therefore savings are exempted from income tax, the contributors are
twice taxed on what they save, and only once on what they spend...The
difference thus created to the disadvantage of prudence and economy is not
only impolitic but unjust.”

Mill concluded that savings should be exempt if this can be administered
without abuse, but was sceptical that this can be done. As a more practical
solution, he suggested that “life incomes” be taxed at only three-quarters of the
rate applicable to “incomes of inheritance”.

While Edgeworth conducted his analysis of sacrifice doctrines in income terms,
the case for the consumption base – in theory, if not in practice – was made by
distinguished theorists such as Marshall (1927), Fisher (1909), Einaudi (1912),
and Pigou (1928). With the exception of Fisher, these authors argued the case in
principle, but, like Mill, did not think a personalized tax on consumption to be
practicable. The first detailed proposal for practical implementation was made by
Fisher (1942) and a second by Kaldor (1955). As the readers of this essay are well
aware, the expenditure tax recently moved to the center of the academic tax
discussion [Pechman, ed. (1980)]. While its fate in the arena of actual tax policy
remains to be seen, the idea of personalized and progressive expenditure tax freed
consumption taxation from its previous association with regressive burden distri-
bution, an association which had prevailed as long as consumption taxes were
viewed as in rem taxes on retail sales.

3.4. Unjust enrichment

Before leaving the topic of tax equity, we briefly return to the income base. This is
to note views that certain types of earnings should be singled out for taxation, not
because they reflect a higher ability to pay, but because the recipient is not
entitled to them. Thus, Aristotle and St. Thomas, while defenders of property,
questioned the legitimacy of interest income [Schumpeter (1954, p. 82)]; and
though this scruple disappeared later on, there remained a presumption that
“earned” income might be given preference over “unearned” income, as indeed
has been the case until recently with the earned income exemption under the U.S.
income tax.

The main instance of differentiation, however, was with regard to land. This
thought goes back to John Locke who, in quoting scripture, distinguished between
the fruits of labor to which a worker is entitled and the fruits of land which God
gave to man to be held in common [Locke (1698, book 11, ch. 5)]. The same
theme was taken up with John Stuart Mill who noted that there were certain exceptions to uniform taxation, "consistent with that equal justice which is the groundwork of the rule" [Mill (1948, p. 817)]. Such is the case with regard to rent from land. The ordinary progress of society increases the income of landlords who have no claim, based on the principle of desert to this accession of riches. Introduction of a penalty tax on prevailing land values, so argued Mill, would do injustice to the present owners; but an extra tax on increments is appropriate, provided they reflect the progress of society rather than the industry of the owner. This discussion was continued by Edgeworth and Marshall who, while agreeing stressed the difficulty of isolating external effects [Edgeworth (1887, p. 216)].

The same theme was continued by Henry George, whose single tax doctrine swept the United States in the 80s and 90s of the last century [George (1880)]. While a staunch defender of private property, George viewed the entitlement to land as held in common. Impressed with rapid gains in land values at his time, he viewed such gains as the source of inequality and social injustice. Following Herbert Spencer, his case for a 100 percent tax on the rent of land was more drastic than Mill's as it was to apply to entire land values and not only increments therein. George thus became the founder of the single tax movement, a movement which was subsequently supported by Brown (1918) and still continues in existence. As will be noted below, land as the prime base of taxation was to receive further endorsement on efficiency (as distinct from equity) grounds. As shown in Chapter 8 of this volume, the taxation of natural resources has remained a problem of great interest.

Nor is this the only instance in which selective taxation of certain sources of income has been argued in the name of tax justice. The taxation of wartime profits under an excess profits tax, for instance, has been common practice. While the concepts of accretion and global base have been central to the equity rule, the underlying premise of general entitlement has been subject to certain qualifications. Changing views on tax equity must indeed be understood in the context of philosophical views regarding the nature of property, individual entitlement thereto, and the relationship of the individual to the state. Thus, there is a vast gap between the neo-Lockean view of taxation as "forced labor" [Nozick (1974, p. 169)] and, say, Justice Holmes' view of taxation as the cost of civilization.

4. Efficiency in taxation

But equity is not all there is to the construction of a good tax system: efficiency also matters, and here economic analysis takes over. Adam Smith, in his fourth maxim, counsels that "every tax ought to be so contrived as to take out and to keep out of the pockets of the people as little as possible, over and above what it brings into the public treasury of the state" [Smith (1776, vol. II, p. 311)].
Reference is to the cost of tax administration, obstruction to industry, the burden of penalties, and odious examinations. Taxes should be "as little burdensome to the people" as possible. John Stuart Mill quotes Smith with approval but finds that no elaboration is needed. At one point, Mill (1948, p. 803) almost recognizes that payment of similar amounts under different taxes may impose differential burdens, but then backs away from this conclusion. Edgeworth, as noted before, accepted least total sacrifice as the utilitarian solution, but also warned that the "productional" consequences of taxation may outweigh "distributional" requirements. At some point "the utilitarian must sadly acquiesce in inequality of taxation" [Edgeworth (1897)]. Similar views were expressed by most early contributors to the equal sacrifice debate, although concern with potential loss of output and reduced growth was more serious in some cases [Bastable (1892, p. 311)] than in others [Adams (1899, p. 351)]. However, reference was to loss of output rather than to dead-weight loss.

The modern formulation of efficiency in taxation was anticipated once more by Dupuit (1844). As noted earlier, Dupuit anticipated modern utility theory by exploring the conditions under which a public works project should be undertaken. In the process he developed (or came close to developing) the concept of a demand curve, and measured the net loss from a tax diagrammatically by the triangle which after Marshall became the standard picture of excess burden. Indeed, Dupuit already recognized that the net loss is proportional to the square of the tax base. But Dupuit's insight was far ahead of its time, as was that of Gossen's (1854) early vision of marginal utility analysis.

The concept of consumer surplus reemerged forty years later in the works of the marginal utility theorists such as Wieser, Menger, and Jevons. Jenkin (1871), as noted below, was the first to use demand and supply curves in incidence analysis, and to show how the burden of a tax exceeds the amount of revenue collected. Marshall (1890, book III, ch. 6), during the same period, developed the concept of consumer surplus on his own terms and warned of the underlying assumption of constant marginal utility of income. He then applied the concept to tax analysis. Inquiring whether competitive equilibrium produces maximum welfare, he suggested that welfare may be raised by giving a bounty to decreasing cost industries while taxing those with increasing cost. In a footnote, he added that the "net loss" (now referred to as excess burden, dead-weight loss, or efficiency cost) of a product tax would be larger for the case of a luxury than for a necessity since demand tends to be more elastic [Marshall (1890, p. 467)]. Ever since, the concept of consumer surplus has played a key role in tax economics – first in the evaluation of particular taxes and most recently in the theory of optimal taxation.

Marshall's application of consumer surplus to taxation was by way of illustrating general principles of price theory, as was the case for most of his tax analysis.
The efficiency implications of taxation were given central focus only thirty years later. Pigou (1928, part II, ch. 5), after discussing sacrifice rules as had Edgeworth, proceeded to an explicit analysis of "announcement effects". As a tax is introduced, a taxpayer finds his options changed and adjusts his behavior accordingly. This results in "announcement burdens" or loss of consumer and producer surplus. The announcement burden of the income tax will be larger if a given revenue is drawn from a particular taxpayer under a progressive rate schedule, than if a proportional or regressive schedule applies. But for a given total revenue to be obtained, the use of less progressive rates requires that the taxpayer with lower income be taxed more. It thus remains questionable which schedule has the greater announcement burden for the group. In all, Pigou expects labor supply to be relatively inelastic and concludes that distributional considerations should be given the major weight.

But the principle of least sacrifice calls for the exclusion of saving from the income tax base. Inclusion, so he argued correctly, taxes future consumption at a higher rate than present consumption. This offends the principle of least sacrifice. It does so, so his argument continues somewhat strangely, because saving is the more elastic use of income and should, if anything, be taxed at a lower rate [Pigou (1928, p. 138)]. Efficiency considerations thus call for an expenditure tax. But Pigou holds a progressive expenditure tax unworkable and thus opts for exclusion of investment income as an equivalent solution. To avoid unjust windfalls, this exclusion is limited to earnings from future investment income only, a proviso which anticipates the transition problems in the expenditure tax debate to follow fifty years later [Pechman, ed. (1980)]. Other situations may also arise where announcement considerations call for discrimination against certain sources of income. Thus the unimproved value of land is a prime source of taxation, as there are no announcement effects. And so are unanticipated windfalls. However, once more, care need be taken lest sudden introduction of a high land tax discriminates against old holders.

While allowing for bearing of demand and supply elasticities on announcement effects of product taxes, Pigou (1928, p. 128) was aware of the complexity of the problem and concluded that a "more powerful engine of analysis is needed" to construct an optimal system. This analysis was provided by F.P. Ramsey who, in response to Pigou's inquiry, laid the basis for what has now come to be known as the theory of optimal taxation. Ramsey (1927) demonstrated that "the optimal system of proportionate tax yielding a given revenue will cut down the production of all commodities in equal proportions. Assuming labor supply to be fixed, this will be achieved by uniform proportional taxes; but with labor supply variable, differential ad valorem rates will be called for, depending on the elasticities of demand and supply". But differentiation between products causes distributional inequities among taxpayers with equal income but different tastes. Given this
further conflict between equity (now within the income group) and efficiency, Pigou (1928, p. 132) suggests that progression be applied via the income tax, supported perhaps by some luxury taxation. He thus stresses the importance of taste differentials, a factor destined to be largely neglected by the optimal taxation to follow fifty years later.

While the contributions of Pigou and Ramsey laid the basis for the modern theory of optimal taxation, its current formulation was slow in coming. The first major contribution, following Pigou (but evidently unaware of Pigou’s and Ramsey’s writing) was that of Hotelling (1938). Departing from Dupuit’s work, Hotelling derived the superiority of a lump sum over an excise tax in general equilibrium terms, based on ordinal analysis and without using the concept of consumer surplus. More questionably, he then extended this conclusion to claiming superiority of an income over an excise tax. In subsequent writings, this superiority became accepted doctrine. It remained so until Little (1951) showed that the earlier conclusion had depended crucially on the assumption of fixed labor supply and that, with labor supply variable, no such a priori judgment could be drawn. As recognized later (although not noted by Little), allowance for a variable goods–leisure choice invalidates the a priori case for ranking a general consumption ahead of a general income tax. Focus on the importance of the goods–leisure choice pointed to product complementarity with leisure as key factor in the selection of an efficient tax base [Corlett and Hague (1953)].

With these foundations laid, Harberger (1964) carried the argument beyond theorizing into the empirical measurement of dead-weight losses for particular taxes. Thus the analysis of excess burden was moved to an applied base and has been actively pursued since then. At the theoretical level, it was not until 1971 that the model of optimal taxation, visualized fifty years earlier by Ramsey, was resumed and expanded [Diamond and Merlees (1971)]. As treated in the following chapter, optimal taxation then became the center of tax theoretical work in the 1970s.

5. Shifting and incidence

Economists have for long been aware that there exists a difference between the point at which taxes are imposed (their “statutory” incidence) and the “final” point at which burdens come to rest. The transition or shifting process has been at the center of tax economics from the Physiocrats on. Indeed, the development of incidence theory closely reflects the development of economic theory at large. Developments in the theory of tax incidence have mirrored the advances of price
and distribution theory, including both their general and partial equilibrium settings.

5.1. The precursors

The first general equilibrium model was that of the Physiocrats and it also spawned the first well defined incidence theory. But the discussion of incidence dates further back. Hobbes' proposal for a tax on expenses, made in 1651, was based on the premise that such a tax falls on consumers. The growth of excises that was to follow focussed the early debate on such taxes. Leading up to the furor created by Walpole's excise reform of 1733, a wide variety of views emerged. Thomas Mun (1664) argued that taxation of necessities would not only raise their price, but also cause wages to rise accordingly. The final burden would thus fall on the rich. Sir William Petty (1667), in what was the first English treatise on tax theory, held that all excises, even those on necessities, will be borne by consumers. This was the case even for consumers of necessities. For some this was reason to reject such taxes as inequitable; but others (including Petty) thought it to be a virtue. A tax on necessities, so they argued, would reduce laziness, add to output, and in line with the contemporary doctrine favoring low wages, would thus be to the advantage of the British economy. Anticipating Physiocratic doctrine, John Locke (1692) held that all taxes, including excises on necessities, would be borne by the landlord. The landowner cannot shift a tax on land since such a tax does not change the "tenant's bargain and profit". A tax on necessities raises wages and thereby the cost to the farmer who, in turn, is able to pay less rent to the landlord.

There is thus a wide range of early opinion based on diverging views regarding the shape of labor supply, including vertical and backward bending. Moreover, many points featured in later discussion already appear in one or another part of this early debate. These include allowance for how tax revenue is spent, the concept of capitalization, the idea that old taxes are good taxes, and a warning that excessive rates will reduce revenue. However, the views are advanced mostly in ad hoc terms and will not be pursued here. The interested reader is referred to Seligman's (1899, book I) scholarly account.

As noted before, rigorous incidence theory begins with the Physiocratic model of income generation, and its first vision of an equilibrating economic system [Schumpeter (1954, part II, ch. 4)]. According to this model, only land was able to produce a net product [Quesnay (1758)]. Labor could merely produce an output needed to maintain itself; and the capitalist's return, net of compensation for risk, was similarly limited. With land the only factor capable of producing a surplus, it followed that land could be the only lasting source of taxation. Taxation of wages
or of products could only lead to economic decline without any lasting revenue gain to the Crown. The sensible way to tax, therefore, was to proceed directly to a tax on land. The Physiocratic model thus led to a warped view of incidence, less realistic indeed than the ad hoc theorizing which had gone before. It is not surprising, therefore, that Turgot, while an ardent proponent of the doctrine, made no attempt to introduce the single tax on land during his tenure in office.

With one extreme leading to its opposite, we may note here the theory of Canard (1801), celebrated at its time, that the search for rents of surplus leads to a diffusion of tax burdens which continues until the burden is shared equally by all participants in the exchange. The burden of taxation, so he held, results from the disturbance caused by this adjustment process and vanishes as the tax comes to rest. Hence, the conclusion that “every old tax is good; every new tax is bad” [Seligman (1899, p. 162)].

5.2. The classics

The system of the classical economists, like that of the Physiocrats, centered on the division of output among factor shares. But the essentially two-factor model of the Physiocrats was now extended to include capital, reflecting the change in perspective from an agricultural to a manufacturing economy. Focus on the return to the three factors not only served as a central analytical tool to explore the laws of value and production, but also dealt with the division of output among the major classes—landlords, capitalists, and workers—which defined the social and economic structure of the times. A view of incidence theory as distribution of the tax burden among these factor shares, therefore, not only fitted the analytical scheme but also provided a political economy of taxation.

In addition to adding the third factor, capital and manufacture, the classical model also broadened the framework of tax analysis by tracing taxation effects through the price adjustment of the market and by drawing a distinction between short- and long-run responses. In this broadened setting, the classics remained true to the Physiocratic tradition of viewing incidence in the context of a truly general equilibrium system. Moreover, the assumption of infinitely elastic labor supply was largely retained, at least in the longer-run context, so that the expanded model still yielded a set of relatively simply solutions.

5.2.1. Adam Smith

The heart of classical incidence analysis is to be found in the work of David Ricardo, but his analysis responded to the pattern developed by Adam Smith. It is thus well to begin with that version as developed in *The Wealth of Nations*
After presenting his maxims, Smith in a series of chapters offers a detailed discussion of the major taxes, including their incidence.

Not surprisingly, the story begins with a tax on the rent of land. If imposed directly on the landlord, so Smith asserts without further explanation, the tax will be absorbed in rent. The same result obtains if the tax is levied on the tenant. The tenant is charged a rent equal to the amount by which the value of his output exceeds what he needs to maintain himself [Smith (1776, vol. I, p. 145)]. Thus he cannot absorb the tax and deducts it from his rental payment. "The landlord is in all cases the real contributor" [Smith (1776, vol. II, p. 313)]. While arriving at a valid conclusion, the reasoning still rests on a Physiocratic notion of net product, rather than on a view of rent as an intra-marginal return. This is evident when Smith arrives at a faulty result for the tax on agricultural produce. Such taxes, Smith (1776, vol. II, p. 321) asserted, "are in reality taxes upon rent", and like taxes on rent they are eventually paid by the landlord. The essential distinction between a tax on rent and a tax on agricultural produce was not as yet recognized.

A direct tax on the wages on labor, so Smith continues, cannot be borne by the worker. The wage is set by the cost of subsistence and therefore cannot be reduced. If the tax is on the wages of agricultural labor, the farmer (as employer of such labor) must pass it back to the landlord through reduced rent. The outcome is similar to that of a tax on agricultural produce. If the tax is on the wages of manufacturing labor, the manufacturer will add it to price. What happens next depends on whether the taxed labor is engaged in the production of luxury goods or necessities. In the former case, the tax is borne by the consumer. In the latter case, the consumer, already living on a subsistence wage, cannot absorb the tax. Wages must rise and the tax, once more, is passed back to the landlord in the form of reduced rent. Smith (1776, vol. II, p. 357) thus concludes as follows:

"Taxes upon necessities, so far as they affect the laboring poor, are finally paid, partly by landlords in the diminished rent of their lands, and partly by rich consumers, whether landlords or others, in the advanced price of manufactured goods. [Therefore] the middling and superior ranks of people, if they understand their own interest, ought always to oppose all taxes upon the necessities of life, as well as direct taxes upon the wages of labor."

*Smith throughout argues that a tax on wages or products would raise prices "in a higher proportion" than the rate of tax [Smith (1776, vol. II, p. 349)], and thus impose an additional burden on the landlord or rich consumer. One reason is that a tax of 10 percent imposed on the gross wage must raise the gross wage by 11 percent to keep the net wage from falling. True enough, but hardly a reason to conclude that the real burden of the tax is increased. Another reason is that the producer will charge a profit on the funds needed to advance the tax, thus resulting in what is later referred to as "tax spiralling".*
Smith notes that the "middling and superior ranks of people" should not only be indifferent between taxes on luxuries and rent, which they pay directly, and taxes upon wages and necessities, which they must absorb indirectly; they should indeed prefer the former. The reason, it appears, is that the latter are taken to raise prices by more than the tax, thereby imposing an additional burden.

The conclusion that wages and necessities cannot be taxed conflicts with Smith's (1776, vol. I, p. 71; vol. II, p. 384) recognition that subsistence may be more or less liberal, depending on whether the demand for labor is increasing, stationary, or declining. Given this range, he might have noted that circumstances may allow for a reduction in the market wage net of tax, on at least a temporary basis. His focus, however, is on the longer run, where the wage returns to its subsistence level. If reduced below that level, population would fall, economic advance would be retarded, and the revenue base would be lost.

Next consider Smith's view of a general tax on profits. Profits or the "return from stock" are divided into compensation for trouble or risk of employing the stock and into interest which belongs to the owner. The former cannot be taxed, as entrepreneurs also seem to have their subsistence wage. The part which reflects interest, however, is likened to rent [Smith (1776, vol. II, p. 331)]. "With the quantity of stock or money in the country, like the quantity of land, being supposed to remain constant, the same after tax as before", Smith (1776, vol. II, p. 352) concludes that interest, like rent, can absorb taxation. The assumption that stock remains constant, however, is qualified by subsequent counsel against its excessive taxation as causing undue inquisition and capital flight.

Taxes which are imposed on profits of particular industries, finally, are passed to the consumer, as capital will be withdrawn until the tax is recovered in higher prices [Smith (1776, vol. II, p. 310)]. Smith thus recognized that returns will be equalized across industries, but he mistakenly interpreted this as burdening the consumer rather than as spreading the tax among all uses of capital.

5.2.2. David Ricardo

We now turn to David Ricardo, the main architect of the classical system of incidence theory.\(^5\) Ricardo's central concern with taxation is evidenced by the very title of his major work, the *Principles of Political Economy and Taxation* (1819). The market, so Ricardo held, does best without interference; but, unhappily, public expenditures are made and taxes are needed to finance them. Thus interference is inevitable: "It is here then that the most perfect knowledge of the science is required." Indeed, "political economy, when the simple principles of it are once understood, is only useful as it directs governments to right measures of taxation" [Ricardo (*Collected Works*, vol. VIII, p. 132)].

\(^5\)For a penetrating discussion of the Ricardian incidence analysis, see Shoup (1960).
All taxes, so Ricardo notes at the outset, are either paid from income or from stock. But government expenditures are "unproductive consumption". They add neither to capital nor provide for advances to labor. Therefore if such expenditures are financed by taxes which fall on revenue, i.e., reduced private consumption, the national capital remains unimpaired; but if such consumption is not reduced, taxes must fall on capital, and eventually distress and ruin follows [Ricardo (Collected Works, vol. I, p. 151)]. No neater formulation of the supply side view of the budget could be desired.

Following the pattern laid out by Adam Smith, Ricardo then turns his attention to particular taxes, taking a critical view of earlier doctrine. A tax on rent, or a land tax levied in proportion to rent, does not apply to marginal land which yields no rent. Since this is the land on which the price of produce is determined, a tax on rent cannot be reflected in that price and must be borne by the landlord. The conclusion is similar to that of Adam Smith, but the reasoning differs. The Physiocratic view of land as the basic source of income is now replaced by rent as an intra-marginal return which does not affect price. With this clarification, Ricardo proceeds to correct Adam Smith's conclusion that a tax on raw produce is borne by the landlord. By raising the cost of produce at the margin of cultivation, such a tax also raises the price of produce. Hence the tax is not paid by the landlord but by the consumer. Such at least is the case until further adjustments, similar to those of a tax on manufactured products, are allowed for. As noted previously, such a tax, if on necessities, cannot be borne by the consumer.

Ricardo's most intriguing argument applies to a tax on wages. As a wages tax is imposed on the worker, nominal wages must rise. This must be the case since labor supply is fixed in the short run and the wage rate is at subsistence. As wages are raised, profits are reduced. Suppose now that as nominal wages rise, the employer comes to recoup his profits by raising prices. This would call for a further rise in wages so as to maintain the real wage at subsistence, generating a further increase in prices, and so forth. Ricardo (Collected Works, vol. I, p. 225) rejects this reasoning as "indefensible", as it suggests that the tax is paid by no one. To determine where the tax falls, he views the problem in terms of resource use. If total output is fixed and part thereof is transferred to government, some other use of resources must be cut. But these cuts cannot be in the wages fund. Since government engages in unproductive consumption (consumption which does not add to necessary advances to labor), since labor supply is fixed and since wages are at subsistence, the wages fund (circulating capital) must remain intact. Therefore, the only resource uses that can be cut are consumption by capitalists.

6 Ricardo further notes that the overall demand for labor remains unchanged with the introduction of the tax. While demand based on capitalist consumption and investment falls, increased government demand takes its place. Thus the wage bill net of tax remains constant.
and the stock of fixed capital. Since both are paid for out of profits, this is where
the tax must fall. Any attempt to recoup the increase in wages by raising price
only requires further wages increases and will not help.

Much the same reasoning applies to a tax on profits. Such a tax cannot be
recouped in higher prices because this would require an increase in wages, nor can
wages be reduced since the wages fund must be kept intact to compensate the
fixed labor supply at its subsistence wage. Since "a tax on wages is in fact a tax
on profits", so Ricardo (Collected Works, vol. I, p. 226) concludes, "I should
think it of little importance whether the profits of stock or the wages of labor were
taxed." Turning finally to a tax on manufactured products, much depends on
whether the product is in the form of necessities or luxuries. A tax on the former
must again fall on profits, whereas a tax on the latter can be absorbed in reduced
consumption of the well-to-do.

Given the short-run context in which labor supply and hence the wage bill in
real terms is held fixed, Ricardo thus arrives at these two conclusions: (1) taxes on
rent, profits, and luxury products are absorbed by the payee, whereas taxes on
wages and necessities are passed on to and borne by profits; and (2) the resource
release from the private sector must be either in reduced consumption of
landlords and capitalists, or in their reduced contribution to the maintenance or
expansion of the fixed capital stock. But the story does not end here. In the
longer-run, reduced accumulation will result in a decline "in society's demand for
labor" [Ricardo (Collected Works, vol. I, p. 222)]. As a result, population
decrees – Ricardo (Collected Works, vol. I, p. 218) quotes Malthus with
approval – until the wages fund is distributed among fewer workers and the
market wage has been returned to its natural level of subsistence. Thus, a new
equilibrium is established at a lower level of population. The net real wage rate is
restored, rent is reduced, and there is a lower capital stock. Taking the very long
view, profit taxes hasten the arrival of the stationary state, as the net (after tax)
return to capital reaches zero at an earlier point and (returning to the Physiocratic
outcome) only rent remains as a taxable income.

Ricardo's argument might be interpreted thusly: Suppose that before tax wages equal $80 and
profits equal $20. Expenditures on necessities equal $80 and investment plus capitalists' consumption
equal $20. After a tax of $10 is introduced, gross wages rise to $90, net wages remain at $80, and
profits fall to $10. Expenditures on necessities, equal to net wages, remain at $80, outlays on
capitalists' consumptions and investment fall to $10, and government outlays rise to $10. The total
remains at $100 and prices are unchanged.

One wonders what would happen to Ricardo's argument if government outlays were made for
"productive consumption". In that case, wages could remain constant, permitting a decline in net
wages. There would no longer be a need for the tax to fall on profits! Due to the peculiarity of the
Ricardian model (holding the net wage bill fixed in real terms), incidence thus depends directly on
how the revenue is used.
This, to be sure, is a simplified version of a highly complex system in which many additional factors are involved. Thus, Ricardo considers how adjustments to a partial tax will differ depending on whether the industry is intensive in fixed or circulating capital, how adjustments to profit and commodity taxes will differ depending on whether the output of precious metals (the monetary standard) is included in the tax base or not, on how the role of trade is affected, and on how government spends the funds. Due to these complications and abundant quarrels with other authors, it is difficult to draw out the core of his argument. Our summary therefore cannot but involve interpolation and interpretation [see also Shoup (1960)].

From the perspective of later analysis, the system is biased by conducting most of the arguments under the assumption of fixed labor supply and subsistence wage, supplemented in the longer run by a Malthusian labor-supply response. Nevertheless, Ricardo offers an impressive structure of micro and macro analysis. Schumpeter (1954, p. 473) may have been uncharitable, therefore, in disposing of the Ricardian model as “an excellent theory that can never be refuted and lacks nothing save sense”. Indeed, as we shall see below, it has only been in recent years that incidence analysis on a Ricardian scale has been resumed in the context of neo-classical growth models.

5.3. The marginalists

Adam Smith and Ricardo, of course, were not the only classical economists who wrote on the incidence of taxation. Others to be noted in a more detailed accounting include McCulloch (1845) who stressed the “reproductive effect of taxation”, and Mill (1849) who restated the Ricardian position in a more flexible fashion and extended its application to international trade. However, the essential theme has been given with Adam Smith and Ricardo, so that we may proceed directly to the next stage, i.e., the rise of marginalism and the modern view of factor pricing.8

The revolution in economic analysis which occurred in the closing decades of the 19th century began with the recognition of utility as a determinant of value. Value was no longer derived from input of labor but from utility in use; and demand, based on relative utilities, was assigned a strategic role in setting relative product prices. This advance was followed by application of marginal analysis to factor pricing and the theory of distribution. The return to labor was no longer determined by a subsistence wage and the Malthusian mechanism of adjustment was dropped. The rule of capital as a factor of production, dealt with ambigu-

8For a review of this development, see Stigler (1941).
ously by the classics, was given specific content. The return to capital and saving was now seen as compensation for contribution to increased productivity via round-aboutness in production. The pricing of all factors in line with their marginal product thus became subject to one and the same principle of compensation. "The theories of the values of labor and of the things made by it", as Marshall (1890) put it in the introduction to his *Principles of Economics*, "cannot be separated: they are parts of one great whole; and what differences there are between them even in matters of details, turn out on inquiry to be, for the most part, differences of degree rather than of kind." The new model was bound to revolutionize incidence doctrine, just as the classical formulation had superceded that of the Physiocrats.

5.3.1. Fleming Jenkin

A first and striking contribution was made by Jenkin (1871); see also Musgrave and Shoup, eds. (1959). Drawing on Jevons' presentation of marginal utility curves to show gains from trade, Jenkin interpreted these as offer curves in relation to price. He was thus the first to have viewed incidence analysis in terms of supply and demand curves, with taxes resulting in shifts therein. He then uses this newly found apparatus to show how the burden of a unit tax is divided between buyers and sellers, and how the injury to each exceeds the tax paid. The total loss for each then depends on the slopes of the demand and supply curves. The terminology, to be sure, was not as yet in terms of consumer and producer "surplus"; and the concept of elasticity remained to be introduced. Nevertheless, the substance of Jenkin's analysis was essentially the same as may be found in textbooks of today.

5.3.2. Leon Walras

While Jenkin was the first to apply marginal analysis to incidence theory in a partial equilibrium setting, it is not surprising that Walras (1874), in his *Elements of Pure Economics*, was the first to apply it in the context of general equilibrium. Walras concluded his treatise with a chapter on taxation. Incidence is viewed in the context of an interdependent set of factor and product prices. Taxes on the three factors (land, labor, and capital) are examined, as are taxes on products. A distinction is drawn between partial and general taxes. While taxes are not formally entered into the set of Walrasian equations, the general argument and its conclusions are in line with modern doctrine.

The incidence of a tax on capital income (i.e., on interest, as there are no profits in competitive equilibrium) will depend on how saving responds. Since this
cannot be predicted, “we may as well assume that the incidence of the tax falls on the capitalist” [Walras (1874), p. 454]. A tax on wages, similarly, will depend on the response in labor supply, which once more cannot be foreseen. Special attention is given to taxes on capital which apply unequally to various uses. Two effects are distinguished. Suppose that the tax is imposed on rental income from housing. Capital in the housing industry will decline, rentals will rise and tenants will bear the burden. But this is not all and Walras (1874, p. 455) continues as follows:

> “Hence a tax on house rent would work out like a tax on consumption—or at least in part, for, if we look at the matter closely, we observe that a portion of the burden is borne by the capitalist. Since some of the capital goods previously employed in the construction of houses will be transferred to all sorts of other employments, a general decline in the rate of income (from capital goods) will result, and this decline will be to the detriment of all capitalists including home owners and to the advantage of all consumers, including tenants. One could, therefore, inquire into the extent to which the consumers thus recover, through the decline in the prices of other services and products, what they lose by the rise in house rents.”

Given a somewhat modern interpretation, Walras thus distinguished neatly between (1) how the depressing effects of the tax on net capital income are generalized among capital in all uses, and (2) how consumers are affected by more or less offsetting “excise effects”. The modern theory of property tax incidence [Mieszkowski (1972)] has its antecedent.

The incidence of product (or indirect) taxes, finally, will be borne partly by the consumers of the taxed product and partly by the owners of the productive services which are employed in their manufacture. The outcome, therefore, is extremely complex, depending on the conditions of demand and supply in the particular markets. Walras’ discussion, while held in fairly general terms, is unobjectionable, to be improved upon only in its specifics in later analysis.

5.3.3. Knut Wicksell

Wicksell’s primary contribution to fiscal theory, as noted earlier, is related to the voting process as a mechanism of preference determination. However, his fiscal treatise begins with an extensive analysis of incidence. Wicksell (1896, p. 5) opens with two methodological observations of importance. First, he rejects the term “shifting” as misleading, because it suggests that A, the initial payee, passes part

---

9 It may be noted that this version first appears in the third edition (1889) prior to which the tax was assigned entirely to consumers [Walras (1874, p. 609)].
of the tax to B and so forth until the entire tax (equal to revenue) is distributed among a chain of payees. This is misleading, because the burden at any one stage may exceed the amount of tax paid, so that the total burden may exceed the total tax. Dead-weight loss, in modern language, should be allowed for. Secondly, he notes the confusion encountered in earlier analysis, which combines the expenditure and tax sides of the budget. While the pattern of government demand matters, this should be separated from the analysis of tax incidence. To do so, government expenditures should be held constant and the incidence of alternative taxes (collecting the same revenue) should be compared. That is to say, incidence should be conducted in differential terms [see also Musgrave (1959, p. 212)].

Wicksell then turns to a tax on monopoly profits. Based on Cournot's (1838) much earlier work [see also Musgrave and Shoup, eds. (1959)], he shows that a tax on profits cannot be shifted. For the case of product taxes, he further shows that the increase in price needed to obtain a given revenue will be smaller under an ad valorem than under a unit tax. The ad valorem approach, therefore, imposes a lesser burden and is to be preferred.

Wicksell's major concern, however, was not with changes in product prices but with the classical problem of incidence among social classes and factor shares. A new formulation was needed, as the Ricardian model of a fixed wages fund, divided among a fixed labor supply, had been discarded by the advances of marginal analysis. Wicksell (1896, p. 35) attempted to fill the void by application of Böhm Bawerk's capital theory. Beginning with the simplest case, he assumes a two-factor model, including labor and capital only, engaged in the production of a single product. Moreover, labor supply and the capital stock are fixed. The only question is how the capital stock is to be used, i.e., how long the "average period of production", \( t \), should be. As \( t \) is lengthened, the productivity of labor is increased, and for any given wage the producer will choose that period for which the rate of interest is maximized. But the wage rate, equal to \( k/t \), where \( k \) is the capital stock, must fall as \( t \) is increased. In combination, these two relationships establish an equilibrium position, determining \( t \) as well as the wage and interest rate.

Wicksell then uses this model – with its peculiar mixture of marginal productivity and wage fund setting – to examine incidence. A tax on income, be it on wages or on interest, will have no effect on the optimal period. The wage and interest rate are unchanged and the tax is absorbed by the payee. But a product tax will affect the outcome, as it is equivalent to an increase in the cost of labor and hence leads to a lengthening of \( t \). Both interest and wages are reduced as a result, with the outcome depending on the shape of the production function or relation between \( t \) and labor productivity. By assuming a one-product model, Wicksell thus bypasses the issue of how incidence is affected by partial taxes and differing production functions. He does, however, amend his model by introducing land and allowing for variable capital and labor supply.
5.3.4. Alfred Marshall

Marshall frequently used the analysis of tax changes to “throw side-lights on the problem of value” [Marshall (1890, p. 412)]. In particular, tax illustrations are used to show how the nature of return to capital depends on the length of time under consideration. Returns obtainable from a given stock or machine are in the nature of quasi-rents, and taxes thereon (like taxes on the rent of land) cannot be shifted. The situation differs, however, in the longer run, when supply is variable. The return to capital is no longer a rent and the tax enters as a cost. Emphasis on the distinction between short- and longer-run adjustments may thus be considered one of his major contributions to incidence analysis. Marshall again shows how a tax on monopoly profits cannot be shifted, while a tax on the monopolist’s product leads to adjustments. Once more, general and selective taxes are distinguished.

Special attention is given to the incidence of local rates [Marshall (1901)]. A distinction is drawn between “onerous” rates, which leave the property without benefits and “beneficial” rates which are reflected in public improvements. He notes that capital movement in response to local differentials relates to the net of the two, but such movement is not considered substantial. Incidence is shown to differ for taxes on sight or building values and once more the adjustment process and the resulting incidence depend on the length of period allowed for.

5.3.5. F.Y. Edgeworth

Next in our parade of early neo-classical incidence theorists, Edgeworth’s (1897) contribution is to be noted. This contribution is distinguished by its systematic approach. Combining assumptions regarding fixed and variable supply, fixed and mobile uses of factors, and increasing and decreasing cost, the incidence of product taxes under various combinations is explored. Special attention is given to “peculiar cases” which arise under conditions of complementarity among products in consumption and production. He thus presents the famous “Edgeworth’s paradox” where it is shown that imposition of a tax on first-class fares may lead to a reduction in both first- and third-class fares [Edgeworth (1897, p. 93), Hotelling (1932)]. Enriched by lovely illustrations of changing slopes drawn from hiking trips in the French Alps, Edgeworth exhibits virtuosity in addressing fine points of incidence.

10Page references are to the 9th edition.
5.3.6. Enrico Barone

Finally, we note Barone's (1899) ingenious application of marginal utility analysis to determine the effects of a tax on work effort [see also Musgrave and Shoup, eds. (1959)]. Primus, producing for his own use, will work so as to equate the marginal utility of output $x$ with the marginal disutility of work $c$. Maximum utility is established at the maximum difference between $U(x) - C(x)$, i.e., where $u(x) = c(x)$. After a lump-sum tax is introduced, this becomes $u(x - a) = c(x)$. The marginal utility of $x$ curve shifts to the right and intersects the marginal utility of work curve at a higher level of output. If, however, the tax is proportional to output, the difference to be maximized equals $U(x - tx) - Cx$.

The optimal value of $x$ is given by $(1 - t)u(x - tx) = c(x)$ and $x$ may rise or fall. Barone thus anticipates later results arrived at by the distinction between income and substitution effects. Moreover, he shows that output will increase or decrease, depending on whether the elasticity of the marginal utility schedule at the pre-tax equilibrium exceeds or falls short of $1 - t$.

5.4. Later developments

These and other contributions to incidence theory during the closing decades of the last century had provided the major breakthrough. Subsequent developments offered improvements built on that base. We shall note very briefly some of the steps in this development, by no means complete but leading up to the current state of the art.

5.4.1. Imperfect competition

We have noted repeatedly how innovations in price and value theory came to be reflected in incidence analysis. A prize exhibit is provided by the work on product tax incidence which flourished in the late 1930s, following the birth of imperfect and monopolistic competition [Robinson (1933), Chamberlin (1938)]. Robinson, in developing the principles of imperfect competition, made extensive use of tax analysis, restating and expanding on Cournot's earlier work on monopoly taxation and even designing a tax device by which to correct monopolistic practice. This was followed by a spate of papers, exploring the relation between unit and ad valorem taxes under competition and monopoly, and given varying cost and demand conditions [Fagan and Jastram (1939), von Mering (1942)].
5.4.2. **Income and substitution effects**

In the earlier discussion, it had become evident that the incidence of factor taxes will differ, depending on how factor supplies respond to a reduction in the net rate of return. The tools for addressing this problem were refined with the distinction between income and substitution effects [Hicks (1938, p. 31)] and its application to tax analysis. Since the two effects work in opposite directions, it followed that no a priori conclusion can be drawn whether factor supplies will fall or rise, a conclusion which had already been reached by Barone. Going further, subsequent analysis pointed to a significant difference between proportional and progressive rates. Since the substitution effect depends on marginal and the income effect depends on average rates of tax, factor supply will tend to be lower under a progressive tax [Hicks (1938)], but not necessarily so since substitution of a progressive for a flat schedule not only raises marginal rates of tax for some but also lowers them for others [Musgrave (1959, ch. 11B)].

5.4.3. **Risk**

The taxation of capital income, by reducing the net rate of return, may reduce saving and the supply of capital, following reasoning similar to that for a tax on wages. But the return to investment is not certain. Rather, it is the expected value of probable gains and losses. The effect of a tax thus depends on how probable gains and losses are dealt with. If the tax law is such as to assure loss offset (be it by carry-backs, carry-overs, or refunds), government becomes a participant in both possible gains and losses and the outcome is not readily predicted [Simons (1938), Lerner (1943), Domar and Musgrave (1944), also Musgrave and Shoup, eds. (1959)]. Examination of taxation effects thus leads into portfolio analysis and investment choice [Tobin (1958), Feldstein (1969)], a topic examined in Chapter 5 of this volume.

5.4.4. **Depreciation**

Also relating to the definition of taxable income from capital, much attention has been directed at the treatment of depreciation. The effective rate of tax is shown to depend on the nominal rate and the time pattern at which depreciation is allowed [Brown (1948)]. Depreciation rules, unless carefully designed, may lead to differential effective rates of tax for industries with different assets lives. Thus the analysis of depreciation was directed at both the use of accelerated depreciation as tax incentive and at devising a depreciation rule which would be neutral among investments of differing asset lives. Economic depreciation as the neutral method
emerged [Samuelson (1964)] and the topic resurfaced fifteen years later when it was given new currency in the context of neutrality of tax incentives [Harberger (1980)] and of inflation [Auerbach and Jorgenson (1980)].

5.4.5. General equilibrium

Recent developments of incidence theory have moved towards a rigorously formulated general equilibrium approach. As we have seen, the classics had such an approach, but the underlying model was incomplete (a lack of capital theory) and unrealistic (the population response). The development of marginal productivity analysis at the close of the last century required a new model. It was recognized that taxes on any one factor may affect returns to other factors, as well as relative product prices; and that taxes on any more product may affect the prices of other products as well as factor returns. Returned to the general equilibrium perspective [Brown (1924)], incidence analysis came to be viewed in differential terms, with any particular tax substitution affecting both the uses and sources' sides of their accounts [Musgrave (1959)].

Mathematical models of general equilibrium incidence made their appearance [Shephard (1944), Meade (1955)] and took hold with Harberger's (1962) model of corporate tax incidence. This was the first model to offer a general yet workable approach. The burden of a profits tax on one industry is shown to be distributed among labor, capital, and consumers, depending on certain characteristics of the taxed industry relative to those of tax-free industries. Assuming the elasticity of substitution of capital for labor to be unity in both industries, factor shares are unaffected and the tax is absorbed by profits, including profits in both taxed and tax-free industries. The analysis is directed at an intermediate period with capital allowed to move, but the total capital stock is held fixed. Moreover, perfect capital mobility and competitive markets are stipulated. Based on this model, a wave of theoretical and empirical work developed and is still in process [Mieszkowski (1969)].

5.4.6. Growth models

Appearance of the neo-classical growth model [Solow (1956)] was followed by introduction of tax variables, thus opening a new dimension of incidence analysis [Krzyzanisk (1967), Feldstein (1974)]. Thus incidence theory closed the circle by returning to the long-term perspective of the classics. With focus directed at effects on factor shares under conditions of steady growth, earlier conclusions from comparative statics were changed. Incidence under steady growth is shown to depend on savings propensities as well as on elasticities of factor supply. Thus,
substitution of a tax on capital income for an equal yield tax on labor income will
(1) leave part of the burden on labor, even if both factor supplies are inelastic,
provided that the propensity to save out of capital income is higher; and will (2)
leave capital with the entire burden, even if labor supply is elastic, provided that
propensities to save are the same.

5.4.7. Empirical studies

Empirical studies of tax incidence began slowly but recently exploded under the
impact of newly available computer technology. Among the early work, mostly
directed at income tax, we may note the Colwyn Report with its contributions by
Coates (1927), as well as studies of income tax effects on labor supply [Break
(1957)] and on saving [Harberger and Bailey (1969), Break (1974)]. Current work
bearing on these key issues is examined in Chapters 4 and 5 of this volume and
effects on housing are taken up in Chapter 7.

At the same time, empirical studies tracing the actual incidence of particular
taxes from observed data remained relatively scarce. An early attempt at econo-
metric estimation of corporation tax incidence concluded that there may be
substantial shifting [Krzyaniak and Musgrave (1963)], but the outcome of this
and subsequent studies remained controversial [Harberger et al. (1967)].

While the concern of classical analysis had been with the distribution of the tax
burden by factor shares, this was no longer the relevant consideration for
purposes of policy analysis. With the change in social structure, attention had
shifted to the distribution of the burden among individuals or households
arranged by income brackets. A systematic attempt at assessing the distributional
effects of the entire tax system along these lines begins with a study by Colm and
Tarasov (1940), continued by Musgrave et al. (1951), and leading up to the
comprehensive work by Pechman and Okner (1974). The methodology underlying
this family of studies was to test the implications of various shifting hypotheses
by estimating the distributional effects which would result. Product taxes are
imputed to consumers, income taxes to the suppliers of factors, and alternative
assumptions are applied regarding the incidence of the corporation and property
tax. Proceeding along similar lines, attempts were made to allocate the distribu-
tion of expenditure benefits, thus aiming to arrive at a pattern of "net redistribu-
tion" through the budget [Musgrave et al. (1974)].

Among various shortcomings of this methodology, it has been noted that the
approach is based on shifting hypotheses rather than on empirical evidence as to
actual shifting. Also, the approach has been critiqued for being limited to a
partial equilibrium setting [Prest (1955)]. Taxes are assigned to either the uses or
sources' side of the taxpayer's account, while neglecting second-round effects.
Moreover, dead-weight losses are disregarded. In the defense of this methodology,
it has been argued that the distributional impact of first-round effects will dominate and that the implications of alternative shifting assumptions may be tested readily.

In recent years, a new approach to general equilibrium estimation has emerged, based on the tradition of the Harberger model and made possible by the advances of computer technology. Taxation effects are estimated in the context of a general equilibrium system, based on production functions and elasticities as deduced from prevailing output and price relationships [Shoven and Whalley (1984)]. Thus the complete chain of secondary effects is included and dead-weight losses are allowed for. However, the outcome still depends on the quality of the parameters which are built into the model. Moreover, it cannot be claimed that the model estimates the observed outcome of actual tax changes. Rather, it simulates the results which emerge on the assumption that adjustments proceed in a perfectly competitive and flexible economy.

6. Stabilization and debt

Up to the 1930s, fiscal economics, with few exceptions, dealt with the effects of budget policy on alternative uses of resources and the distribution of income. This analysis, as we have seen, was conducted in the context of a full-employment economy. Much attention was given to fiscal effects on the division of output between capital formation and consumption, and thereby on growth; but by the nature of the underlying macro model, effects on the level of employment were outside the confines of analysis. With the “Keynesian revolution” of the 1930s [Klein (1947)], aggregate demand became a major factor in determining the level of employment; and with it budget policy gained a new and strategic role. The stabilization function was added to the more traditional aspects of budget policy, and fiscal policy moved into the center of macro economics. Most of the fiscal literature of the 1930s and 1940s was directed at exploring this new dimension.

6.1. Fiscal policy and stabilization

While this phase of fiscal analysis gained dominance with the rise of Keynesian economics, it also had its antecedents.

6.1.1. Early Keynesians

Aggregate demand was of concern in mercantilist thought, and Sir James Steuart (1767, vol. II, pp. 642, 644), writing a decade before the Wealth of Nations, argued
that stagnant money “lent to government, is thrown into a new channel of circulation—thereby to augment the prominent income of the country”. Then there was that “brave army of heretics, Mandeville, Malthus, Gesell, and Hobson who, following their intuition, have preferred to see the truth obscurely and imperfectly rather than to maintain error, reached indeed with clearness and consistency and by easy logic, but on hypotheses inappropriate to the fact” [Keynes (1936, p. 371)].

That central hypothesis, of course, was Say’s (1821) law, i.e., the proposition that there could be no general glut since, commodities being exchanged against commodities, supply would create its own demand. Malthus (1820, p. 316), the chief heretic among the classics, was critical of “Mr. Say, Mr. Mill, and Mr. Ricardo, the principle authors of these new views”. An adequate level of “effectual demand” is needed, so Malthus (1820, p. 326) argued, to sustain output:

“But if the conversion of revenue into capital pushed beyond a certain point must, by diminishing the effectual demand for produce, throw the laboring classes out of employment, it is obvious that the adoption of parsimonious habits beyond a certain point, may be accompanied by the most distressing effects at first, and by a market depression of wealth and population afterwards.”

His primary problem with saving, it appears was not that funds will be withheld from the expenditure stream, as in the Keynesian model. Rather it was that too much is spent on capital formation, leaving consumer demand insufficient to absorb the increase in potential output that results. In consequence, profits will fall and accumulation will decline. Balance may be restored but only after distress has occurred. Malthus thus explains the depression which followed the Napoleonic War by under-consumption. Of particular interest in our context are certain implications which Malthus draws for budget policy. If the problem is one of deficient consumer demand, budget policy can be helpful. Consumer demand might be raised by redistributing income towards “those classes of unproductive consumers who are supported by taxes” [Malthus (1820, p. 410)]. Moreover, budget policy can be harmful if consumer demand is reduced by excessively rapid repayment of public debt. Having noted this much, Malthus (1820, p. 411) hastens to add that property rights must not be violated by redistribution, and assures the reader that he is not “insensible to the great evils of public debt”.

Malthus, along with Sismondi, was among the first to advance an underconsumption theory of crisis. Marx offered a related doctrine and underconsumption theories reached high fashion in the writings of the 1920s. [For a discussion, see Hansen (1927), McCord Wright (1942), Keynes (1936, ch. 23).] However, these contributions paid little attention to the role of the public budget, nor did monetary theories of business cycles of that period.
6.1.2. The Keynesian model

Analysis of budgetary effects on employment enter the public finance literature only in the mid-1930s, when the stage had been set by Keynes' (1936) *General Theory of Income and Employment*. The impact of Keynesian economics on fiscal theory profoundly changed its focus. Whereas the problem had been to observe resulting shifts in resource use, concern now was with effects on its overall level. With employment seen to depend on aggregate demand and with budget policy a direct contributor thereto, budget policy became a critical determinant of the level of employment. This new function of budget policy was the more important because departure from full employment was seen no longer as a temporary aberration, but a central tendency of the economy. A continuing tendency towards oversaving [Keynes (1936, p. 31)] and stagnation [Hansen (1938, 1941)] was expected to prevail, with expansionary fiscal policy called for on a sustained basis in order to maintain high employment in a mature economy. Moreover, fiscal policy was not only one but the policy instrument with which to remedy the problem of underemployment. Monetary policy at least in the earlier depression phase of Keynesian economics was rendered ineffective by the existence of a liquidity trap. Aggregate demand, like a string, could not be pushed up by monetary expansion, but it could be pulled up by government outlays.

The Keynesian case for deficit finance, from the beginning, was doubly controversial. Not only were the underlying analytics questioned, but the model carried political and ideological implications which contributed to the heat of the debate. If fiscal expansion had to be secured via increased government expenditures, it would also add to the size of the public budget. Moreover, the central proposition that private saving may be a public vice offended deeply held values based on Puritan tradition. The notion that different principles of prudence applied to the public and the private sector ran counter to the concept of a society based on the beneficial interplay of individuals in the market.

In the course of time, the extreme view of the early Keynesian model was moderated, and the supreme powers of fiscal policy were turned down. However, its role in stabilization remained a key factor. The field of “Public Finance”, traditionally a subject in micro economics, became part of macro teaching and macro issues claimed dominant attention. No attempt can be made here to trace the development of macro theory from the thirties to the sixties, not to speak of entering into the current debate. Rather, we limit ourselves to a brief look at certain macro issues which arose in the context of fiscal policy. Among the major contributors to the development of the fiscal policy concept we may note Alvin Hansen, whose Fiscal Policy Seminar, conducted at Harvard during the late 1930s, was the mainspring of the new approach in the United States [Hansen (1941)] and, in Great Britain, Beveridge's (1945) program for full employment in a free society was to be the blue-print for macro policy in the post-World War II
years. In the United States, the Employment Act of 1946 made it the President's responsibility to "promote maximum employment, production, and purchasing power" and, as added in 1953, "a dollar of stable value". It also established the Council of Economic Advisors to pursue these goals.

**Multiplier analysis.** Central to the role of fiscal variables in the Keynesian model was their treatment in the multiplier formula. Initially this was thought of as the multiplier effect of an increase in "government investment", observed while holding tax revenue constant. By adding government purchases, the level of autonomous expenditures would be raised, thus offsetting a higher level of private sector saving and permitting a higher level of income. The measurement of the fiscal multiplicand was explored [Clark (1935), Currie and Krost (1939), Villard (1940)], based on the assumption of a fixed level of investment. Samuelson's (1939) multiplier–accelerator model then expanded multiplier analysis to include investment effects and to examine the pattern of resulting income fluctuations. While initial emphases had been on increases in government purchases, subsequent analysis admitted tax reduction as a second device, leading to the analysis of alternative packages of tax and expenditure changes which would secure the same overall leverage effect [Beveridge (1945), Hansen (1945), Musgrave (1945), Samuelson (1948)].

In line with the early focus on excess savings as the villain, the creation of a public deficit (public dissaving) was first seen as an essential feature of fiscal leverage. It therefore came as somewhat of a shock when it was demonstrated in the early 1940s that even a balanced-budget increase could exert a leverage, albeit with a multiplier of unity only [Gelting (1941), Salant (1942), Havelmo (1945)]. Within a short time span the "balanced-budget theorem" had been advanced independently in a number of places, a nice illustration of how what was once unthinkable becomes commonplace when the time is ripe [Salant (1975)]. Examination of the multiplier time period [Metzler (1948)] and further exploration of various policy lags followed [Friedman (1948)], and the estimation of multiplier effects became a central part of the newly developed breed of econometric models [Klein and Goldberger (1955)].

**Fiscal structure.** The newly formed role of fiscal policy also placed the quality of various taxes in a new perspective. Taxes which fell heavily on consumption would carry a larger (negative) multiplicand than those falling on saving and thus do less damage to the leverage of the budget. Thus full-employment policy was linked to progressive taxation. A tax on undistributed profits was enacted briefly in 1937 [Colm (1940)] and the feasibility of a tax on hoarding, linking back to Gesell's idea of stamped money, was considered [McWright (1942)].

Given these new uses of tax and expenditure instruments in the context of macro policy, the question arose how they could be reconciled with the traditional
fiscal objectives, i.e., the provision for needed public services and the design of an efficient and equitable tax structure. Keynes (1936, p. 220) was pleased to shock his bourgeois reader by noting that even the construction of pyramids or the digging of holes would increase the national income, without adding that constructing a useful highway may be even better. Lerner (1943), in his “functional finance” doctrine, viewed taxation merely as a means to reduce the purchasing power of the public in contrast to its traditional role of transferring resources to the public sector. This author, concerned with avoiding distortions due to conflicting objectives of various policy functions, proposed a “three-branch model” in which various functions of the budget could be reconciled and be performed in a compatible fashion [Musgrave (1959, ch. 1)]. Changes in fiscal leverage, in that context, would be expedited primarily via changes in tax rates with government purchases set so as to meet the need for public services at a full-employment level of income. These considerations, we may add, are no less relevant in the current setting where macro considerations point in the direction of budgetary restriction.

**Built-in flexibility.** By the late 1940s, a distinction emerged between the stabilizing effects of the fiscal system which arise as the result of discretionary changes in fiscal parameters and those which arise automatically in response to changes in the level of economic activity. Measures of built-in flexibility were devised and the comparative flexibility of various taxes was explored. Critics of discretionary policy held that changes in fiscal leverage should be limited to those which result automatically, while setting the level of tax rates so as to balance the budget at full employment [Friedman (1948), Committee for Economic Development (1947)]. Others held that discretionary changes cannot be dispensed with.

Over time it appeared that built-in flexibility might not be an unmixed blessing. Whereas the automatic decline in revenue in the course of a recession would be helpful, the automatic increase in the upswing, or a secular increase in response to growth might generate a fiscal drag [Economic Report of the President (1962), Heller (1967)].

6.1.3. The neo-classical model

The economic experience of World War II, with its massive growth in the budget and rise in GNP, had demonstrated the powers of expansionary policy under wartime conditions, and during the forties thinking about postwar policy projected a continued need for expansionary fiscal measures. As it turned out, the postwar decades produced a much stronger economy, and with it reinstated monetary policy as an effective policy tool. The approximate mix of fiscal and monetary policy was examined in the context of a “neo-classical” policy model,
designed to accommodate both monetary and fiscal policy variables and to allow for the effects of unbalanced budgets (deficit or surplus) on the structure of claims [Samuelson (1951)]. Various degrees of monetary and fiscal tightness (or ease) could be combined to achieve the same level of aggregate demand but they would differ in the resulting mix of consumption and investment with a tight fiscal and easy money mix more favorable to growth. There would also be a difference in balance of payment effects, with the tight money and easy fiscal mix yielding the more favorable results. Adding selective instruments here needed, policy tools would be adequate, if wisely used, to achieve various macro policy goals.

The high point of optimism regarding the New Economics [Heller (1967)] and the powers of stabilization policy was reached in the first Economic Report of the Kennedy Administration (1962) and the recovery following the tax cut of 1964. Thereafter, the changing economic scene, with its shift in concern from unemployment to stagflation, produced a setting less favorable to the powers of fiscal policy. With it, changing perspectives on macro theory and the ensuing monetarist-fiscalist debate called for reconsideration of earlier tenets. Thus a new chapter was opened, but one which extends beyond the time span of this essay.

6.2. Public debt

The economics of public debt, the final topic to be considered here, has been among the most controversial parts of fiscal doctrine. In some measure, this reflects its strategic role in fiscal politics. Resort to debt finance is said to facilitate spending, remove public outlays from taxpayer control, and burden the future. The very proposition that the rules of prudence in private debt accumulation may not apply to the public sector offends, as noted before, the image of a natural order based on the rules of the market. But beyond this, subtle problems of economic analysis arise, problems which are still (or better, again) at a debatable stage.

6.2.1. Debt burden and future generations

Central to the debate is the issue whether the burden of the debt is paid for by future generations. The mercantilists thought not. Credit was viewed as a creation of wealth and outstanding debt was no burden. As stated by Melon (1734), an associate of John Law, public debts, if domestically owed, are debts which the “right hand owes to the left”. Pintus, Voltaire, and Condorcet took similar positions. The growth of French and British debts in the eighteenth century, however, produced a more critical view. Thus Montesquieu and Hume rejected Melon’s proposition as specious [Hume (1742)]. Smith (1776, vol. I, pp. 410,412)
similarly rejected the transfer argument as “sophistry of the mercantile system” as well as the mercantilist view that the national debt is an addition to the nation’s capital. Debt finance may be needed in wartime, or on other special occasions, but tax finance is to be the general rule. Tax finance will be drawn largely from funds otherwise used in the employment of unproductive labor, whereas loan finance will divert funds from the maintenance of productive labor, thus impairing the country’s capital stock [Smith (1776, vol. II, p. 410)]. Moreover, the burden of tax finance is felt at once, thus creating taxpayer resistance and providing protection against public waste.

Ricardo’s contribution to debt theory has regained recent attention as the “Ricardian equivalence theorem”. Hidden in a chapter on commodity taxes in the Principles [Ricardo (1817, p. 244)] and restated in his essay on the Funding System [Ricardo (1820, p. 187)], he offered this intriguing contribution to debt theory: Suppose, so he argues, that £20 million has to be raised to pay for the expenses of a year’s war. In the case of tax finance, let a particular individual be called upon to pay £2,000, or 1/10,000 thereof. In the case of loan finance, and with interest of 5 percent, £1 million per annum must be paid in interest to the lenders. Of this, our taxpayer is asked to pay £100. Under tax finance, he could have borrowed from the same lenders to finance his tax of £2,000, being left once more with an annual charge of £100. From the taxpayer’s point of view, the two methods are therefore equivalent.

Put in modern terms, Ricardo concludes that the taxpayer in the loan finance case discounts his future tax liability and finds his net worth reduced as it would be under tax finance. “In point of economy”, there is no real difference “between the two modes” [Ricardo (1820, p. 186)]. The burden of the war is paid for by the taxpayer during the year in which it is financed, be it in his role of paying £2,000 at once or as assuming a tax obligation of £100 per annum. Future interest payments, therefore, are only transfers among the future generation and impose no burden.

But having posed this argument as holding “in economy” – meaning, we take it, on the assumption of perfect foresight and rational behavior – Ricardo hastens to reject it as unrealistic. Asked to pay the full £2,000 in the case of tax finance, the taxpayer will endeavor at once to “save speedily” that amount from his income. Under loan finance, he has to pay only £100, and will thus “consider that he does enough” by saving this lesser sum “and then deludes himself with the belief that he is as rich as before”. In short, “loan finance is a system which tends to make us less thrifty – to blind us to our real situation” [Ricardo (1817, p. 247)].

Tax and loan finance both involve the diversion of resources to wasteful use, but they differ in their effect on capital formation and hence on the position of future generations. Saving and capital formation are reduced as tax finance is replaced by loan finance, and the future generation is burdened by having a lower income. Ricardo’s rejection of the equivalence theorem could not be more
explicit, and it is strange that the equivalence theorem could now be presented under his banner [Shoup (1960), Driscoll (1977)].

The economics of public debt also received lively attention among continental authors. German writers, tending towards a more favorable view of the public sector, saw the growth of public debt with less alarm. Dietzel (1855), impressed with the economic advance of Great Britain, attributed it to the rapid growth of public debt. Equating growth of public debt with public capital formation, he viewed the former as a sign of growing national wealth. Public capital formation, moreover, would include not only the real but also the "immaterial" capital of the state, such as the existence of legal institutions. Notwithstanding his overly enthusiastic view of state outlays, he anticipated future thought by calling for loan finance in the case of capital and tax finance in the case of current outlays. A similar, though more cautious support, of public debt was advanced by Wagner (1883, p. 184), whose views on the growth of public expenditures we previously encountered. The Italian literature of the 1890s, that decade of flourishing fiscal theory, accepted and elaborated upon the Ricardian equivalence theme [de Viti de Marco (1893), and, for a review of the Italian literature, Buchanan (1960)]. The leading French, British, and American texts [Leroy-Beaulieu (1906), Bastable (1892), Adams (1892)], however, adhered to the view that debt finance reduces private capital formation and thus places a burden on future generations by reducing the capital stock which is bequeathed to them.

This view was shaken by the impact of Keynesian economics. Not only would creation of debt be a necessary byproduct of fiscal expansion needed to secure high employment, but outstanding debt would pose no serious subsequent problem. The old doctrine that interest payments constitute a transfer from the right to the left hand now reappeared as "we owe it to ourselves". Public debt, so Lerner (1948) argued, differs from private debt because the latter is owed "to others" whereas the former is owed to citizens of the "same nation". Creation of national debt, therefore, is no subtraction from national wealth, nor do interest payments by members of a future generation reduce the national income of that generation.

To be sure, tax finance of interest payments might induce burdensome disincentives and dead-weight losses. But, so Lerner argued, tax finance of interest charges is not needed, since interest payments may in turn be loan financed. Tax finance becomes necessary only after interest payments become so large, relative to earnings, that further loan finance would become inflationary. At this point, a large national debt might become a serious problem, but he did not think this situation likely to arise. The wealth effect of growing debt reduces the propensity to save and thus terminates the need for further debt expansion, leading to an equilibrium level of public debt at full employment.

Essential to this view of debt and interest burdens is the assumption of an underemployment economy which calls for aggregate demand to be raised by
fiscal expansion. Government borrowing activates funds (recall our earlier reference to Sir John Steuart), but does not reduce private investment. Hence there are no depressing effects on future generations by leaving it with a reduced capital stock. Given an extreme Keynesian model with fixed investment and excess saving, this conclusion follows, just as the opposite conclusion (that loan finance burdens future generations) is appropriate for a full-employment model.

While the case for deficit finance (increase in debt) came to be accepted as an appropriate employment policy, concern with the burden of interest service and its effects on future generations continued. This fear was allayed, however, by the proposition that interest burden was a function of the ratio of interest bill to GNP, rather than its absolute level. Given a constant ratio of deficit to GNP and constant yields, the ratio of interest bill to GNP also approaches a constant level [Domar (1944)].

Subsequent discussion returned the argument to a full-employment setting. Attack on the “new orthodoxy” (i.e., the we-owe-it-to-ourselves proposition) was led by Buchanan’s (1958) subjective approach. There can be no initial burden, so he argued, since lenders are not called upon to contribute anything. Thus no burden is imposed in the initial period. Future taxpayers, however, are burdened by having to finance interest payments. Thus burden transfer occurs. Others continued to stress reduced capital formation and the burden which it imposes as the future generation as a whole is left with a reduced capital stock [Shoup (1962)]. With future tax payments needed to finance interest service equal to the loss of capital income, the two formulations yield essentially similar results.

The difference between loan and tax finance was thus left to depend on resulting differences in resource withdrawal from consumption and capital formation [for major contributions to this debate, see Ferguson (1964)]. The fact that loan finance may involve burden transfer, however, is not necessarily an argument against it. In the context of public capital formation, transfer of burden via loan finance may serve as an instrument of intergeneration equity and a rationale for a capital budget approach [Musgrave (1959, p. 562)]. A further equity-oriented case for loan finance may arise in the context of war finance, where the use of “refundable taxes” permits a postwar correction of inevitably heavy wartime burdens on low income groups [Keynes (1939)].

6.2.2. Public debt and liquidity structure

Apart from the differential implications of tax and loan finance, much attention was given during the 1940s and 1950s to how an outstanding debt should be managed. The two major issues were the choice between marketable and non-marketable bonds, and the maturity structure of the debt. With the former essentially a question of distributional outcome, the latter led into the linkage
between fiscal and monetary theory. With issuance of debt viewed as a purchase of liquidity, the Treasury should choose between short- and long-term issues so as to minimize the cost of securing a given reduction in liquidity [Rolph, (1957)]. Short-term debt, being closer in nature to money, would buy less illiquidity than long-term debt and hence be worth more to the Treasury.

The proposition that debt issue reduces liquidity was questioned, however, as debt policy came to be viewed in the context of general portfolio choice [Tobin (1963)]. Moreover, attention was given to the fact that the choice of maturities so as to minimize interest cost involves not only the prevailing term structure of rates, but also anticipation of future changes therein [Smith, (1970)]. Finally, there was the question of how maturity structures of different lengths would affect the stability of the market. Long-term debt would avoid the hassles of refinancing but would result in larger fluctuations in the market value of outstanding bonds, thereby increasing the risk of “disorderly conditions”, especially in the case of monetary restrictions. With the drastic shortening of the debt in the postwar decade, these issues which once were lively topics have largely disappeared from the discussion.

7. Conclusion

This closes our account of the evolution of fiscal theory. Over the two centuries here surveyed, the economics of public finance has grown enormously both in breadth and sophistication, moving with and benefitting from the growth of economic analysis at large, but also contributing thereto. This growth, however, has been far from linear, with insights cropping up, dropping out, and reappearing when their time had come. But great progress has been made. Yet, the basic problems have remained the same. The questions of what public services should be provided, how they should be financed, and what role government should play in the macro conduct of economic affairs were visible to Adam Smith, and they still pose the basic problem.

So does the fact that many issues in public finance remain inherently controversial. To establish the economic case for the public sector is to delimit the sphere that can be left to the invisible hand and the rules of the market. The scope of existing externalities, the acceptability of a market-determined income distribution, the shape of the social welfare function, maintaining full resource utilization, the issues of inflation and growth, all these have powerful bearings on the appropriate size and activities of the public sector. So does the capability of public policy to apply appropriate corrections, with the scope of public policy failure matched against that of market failure. Given this array of problems and their linkages, ideological and value issues are never far away. Moreover, the tools of fiscal policy changes with changing fiscal institutions. It is not surprising,
therefore, that the history of fiscal doctrine deals with more than the development of economic analysis per se. To this writer at least, that adds to rather than detracts from the fascination of our subject.

References

Arrow, K., 1951, Social choice and individual values (Wiley, New York).
Barone, E., 1899, About some fundamental theorems on the mathematical theory of taxation, Giornale degli Economisti, Serv. 2, 4.
Bergson, A., 1938, A reformulation of certain aspects of welfare economics, Quarterly Journal of Economics 52.
Buchanan, J., 1959, Public principles of public debt (Irwin, Homewood, IL).
Canard, F., 1801, Principes d'économie politique (Paris).
Also see Musgrave and Peacock, eds. (1958).
Colm, G., 1940, Full employment through through tax policy, Social Research 7.
Ch. 1: A Brief History of Fiscal Doctrine

Cournot, A., 1838, Récherches sur les principes mathématiques de la théorie des richesses (L. Hachette, Paris).


Diezal, C., 1885, Das System der Staatsanleihen im Zusammenhang der Volkswirtschaft betrachtet (Heidelberg).


Einaudi, L., 1912, Intorno al concetto di reddito imponibile e di un sistema d'imposte sul reddito consummati (Turin).


Feldstein, M., 1974b, Tax incidence in a growing economy with variable factor supply, Quarterly Journal of Economics 88.


Hansen, A.H., 1927, Business cycle theory (Ginn & Co., Boston, MA).


Hansen, A.H., 1945, Three modes of expansion through fiscal policy, American Economic Review 35.


Hicks, J.R., 1939, Value and capital (Oxford University Press, Oxford).
Hobbes, T., 1651, Leviathan (Penguin Classics, Baltimore, MD).
Lindahl, E., 1919, Die Gerechtigkeit der Besteuerung (Gleerupska Universitets Bokhandeln, Lund).
Niskanen, W.A., 1974, Bureaucracy and representative government (Aldine, Chicago, IL).
Panteleoni, M., 1883, Contributo alla teoria del riparto delle spese pubbliche, Rassegna Italiana.
Reprinted in: M. Panteleoni, Scritti vari di economia, Vol. I.
Quesnay, F., 1760, 1762, Le tableau économique, and Theorie de l'impôt. See August Oncken, ed., 1888, Oeuvres économiques et philosophiques de François Quesnay (J. Bear, Frankfurt).
Säffle, A., 1867, Das gesellschaftliche System der menschlichen Wirtschaft (Tübingen).
Shoup, C.S. and R.A. Musgrave, eds., 1959, Readings in the economics of taxation (Irwin, Homewood, IL).
Smith, Adam, 1776, An inquiry into the wealth of nations, 1904 ed. by E. Cunnan (Putnam's Sons, New York).
Viti de Marco, A. de, 1888. Il carattere teorico dell’economia finanziara (Rome).
Walras, L., 1874, Elements of pure economics, 1954 translation by W. Jaffe (Irwin, Homewood, IL).