From the static to the dynamic: Some problems in the theory of taxation

Nicholas Stern*

*London School of Economics, London WC2A 2HD, UK

Revised August 1991

This paper considers problems and issues that arise when we attempt to extend the static theory of taxation and production planning to a dynamic context and to combine that theory with theories of growth. The reasons dynamic theory is less developed include problems with the analytical tractability of even those dynamic models which attempt to retain the basic structure of the static model. Furthermore, difficult and interesting issues arise in a dynamic context which are either not present or are much less severe in the static theory. Special attention is paid to these new issues and the difficulties they pose.

1. Introduction

Planning is about influencing the future and is therefore intertemporal. Amongst the challenges of the analysis of taxation in a planning context is the explicit extension of tax theory, much of which is static to a dynamic economy. As argued in the Introduction to this issue, this was one of three of the themes of the conference, the second being taxation and the level and allocation of investment, and the third the problem of taxation when some prices are administered. This paper may therefore be seen as developing some of the issues raised in the short Introduction to this issue.

The normative theory of taxation and its relation with production planning for a static economy now constitute a substantial body of literature. There is also a voluminous literature on the theory of growth, primarily

Correspondence to: N. Stern, London School of Economics, London WC2A 2HD, U.K.

*This paper was prepared for the Delhi conference on Taxation and Planning organised by the International Seminar on Public Economics and the National Institute of Public Finance and Policy, 11–13 January 1990. I am grateful to Dr. Amaresh Bagchi, Director of NIPFP, for his kind invitation, and for his comments on the paper. I am also grateful for the comments of Paul Hare and David Newbery, who were the discussants at the conference. Support was provided from the Suntory-Toyota International Centre for Economics and Related Disciplines at the London School of Economics from the Bradley Foundation, and from the ESRC Programme at STICERD on Taxation, Incentives and the Distribution of Income. Robin Burgess, Stephen Howes and Jean Lanjouw provided valuable advice and assistance. Helpful comments were received from Tony Atkinson and Mervyn King. All errors are mine.
positive but also normative. The purpose of this paper is to consider problems and issues that arise when we attempt to extend the static theory of taxation and production planning to a dynamic context and to combine that theory with theories of growth. How, for example, can the insights from the theory of taxation complement those from the theory of investment and growth? Ramsey (1927, 1928) was, of course, the founding father of the theories of both optimal saving and optimal taxation. And it is striking that many others of those who figure in the theory of taxation have also been contributors to the theory of growth, including Wicksell, Hotelling and Samuelson, as well as more recently Atkinson, Dixit, Newbery, Sandmo, Sheshinski, Stern and Stiglitz, for example, and particularly Diamond and Mirrlees. Yet our dynamic theories of normative taxation have been much less well developed than the static. The reason, we shall suggest, lies in a combination of a number of related factors. These include problems with the analytical tractability of even those dynamic models which attempt to retain the basic structure of the static model. Furthermore, difficult and interesting issues arise in a dynamic context which are either not present or are much less severe in the static theory.

Within our discussion of the extension of static theory to dynamic models we shall pay special attention to these new issues and the difficulties they pose. Thus this paper will largely be a partial survey, a statement of problems, an assessment of where we stand, and a research programme, rather than a presentation of new results.

We provide, in the next section, a brief summary of the point of departure, the static theory of taxation. In section 3 we set out first some of the basics of the normative theory of growth with exogenous technical progress, and then look at some of the recent ‘endogenous growth’ models. In section 4 we investigate the problems of extending the static theory and set out and examine some of the new issues that arise. The final section contains a brief evaluation of where we stand on dynamic tax theory.

2. The static theory of taxation and public production

As concerns taxation, the basic theorems of welfare economics linking Pareto efficiency and competitive equilibrium point to lump-sum taxes and transfers as the only non-distorting tools and generally educate us that efficient targets for taxation are goods in inelastic demand or supply. Under the standard assumptions for the theorems (no externalities, convexity, existence of all markets and availability of lump-sum transfers) relative consumer prices, producer prices and shadow prices (defined as social opportunity costs in terms of the social welfare function) are all equal. Hence public production decisions can be taken using profit maximisation at market prices. The standard theory of externalities alerts us to the use of indirect
taxes to alleviate inefficiencies that would otherwise arise from market prices failing to fully reflect social costs.

Much of the theory of optimal taxation in its recent form can be seen as an extension of the first-best theory to the case where lump-sum taxes and transfers are not possible, but where the other assumptions of the first-best theory are retained.

A natural point of departure in the recent literature on optimal indirect taxation is Diamond and Mirrlees (1971), which followed in the tradition of Ramsey (1927), Samuelson (1951, 1986), Boiteux (1956, 1981) and others.

As well as deriving the Ramsey rules (suitably generalised) for their model they showed that, under fairly weak assumptions, the optimum should be efficient in the sense of being on the production possibility frontier. Thus (under standard convexity assumptions) one can think of this optimum being decentralised using producer prices \( p \). Indirect taxes, \( t \), are then the difference between consumer prices, \( q \), and the producer prices, \( p \); \( t = q - p \). Where there is private as well as public production one can ask whether production taken as a whole should be efficient (so that marginal rates of transformation in the public and private sectors should be the same). Generally efficiency will not be a feature of the optimum in this case, unless all goods can be taxed and there is no influence of profits on income distribution (or the influence can be optimally controlled) either because of constant returns and perfect competition (or unrestricted taxation of profits). If overall efficiency is desirable prices for private production and for decentralisation of public production should be the same so that, for example, there should be no taxation of intermediate goods.

Where there is a single consumer the Diamond–Mirrlees framework yields, as a condition for the optimality of taxes, the famous Ramsey rule [Ramsey (1927)], which may be interpreted as saying that, at the optimum, the compensated quantity change resulting from a small uniform intensification of all taxes should be the same for all goods. Another way of expressing the result, which is popular in the literature, is to say that, for small taxes, the proportional reduction in compensated demand associated with the taxes should be the same for all goods.

A number of authors have examined the structure of indirect taxation associated with the Ramsey rule in terms of complementarity and substitutability with leisure. For the case where there are two goods and labour, Corlett and Hague (1953) showed that one should tax more heavily the good that is more complementary with leisure. A number of authors, e.g. Sandmo (1976), Sadka (1977), Deaton (1981), Besley and Jewitt (1987), have examined the Corlett–Hague analysis with many goods. They have shown that equality of compensated cross-elasticities with leisure implies that all goods (apart from labour) should be taxed at the same rate and have further related this condition on cross-elasticities to assumptions on the shape of preferences. It
is remarkable that the earliest paper on the subject, Corlett and Hague (1953), focused immediately on this critical assumption. It should also be noted that leisure can be interpreted as non-market time so that the Corlett and Hague result points to the taxation of goods which are complementary to non-market activities, including home production.

Where there are many consumers the many-person Ramsey rule may be interpreted as saying [Diamond (1975)] that the proportional reduction in compensated demand should be lower the more the good is consumed by the 'deserving' (as measured by the social marginal utility of income). The rate of taxation on a good in this many-person Ramsey theory can be seen as arising from two considerations. The first, deriving from the one-consumer analysis, is related to tax efficiency, and concerns the cross-elasticity with leisure, with greater complementarities mitigating in favour of higher tax rates. The second concerns distribution with a greater predominance of the rich (assumed less 'deserving' in the formal sense of the model) in the consumption of a good implying a higher tax rate.

The many-person Ramsey analysis involves a model with indirect taxes only and we should ask how the conclusions might be modified if an income tax is also available. A linear income tax can be introduced simply by adding a poll tax or grant, since proportional income taxation can be represented by a proportional increase in all goods' prices. The models examined generally assume that labour is the single source of earned income and that wages vary across households. Deaton and Stern (1986) [building on results of Atkinson (1977) and Deaton (1979)] considered the case of linear parallel Engel curves for goods, and labour separable from goods, but where the intercepts of the Engel curves vary across households. These intercepts may be linearly related to household composition and have a random term which varies across households. If there is an optimal system of transfers (demogants) which are linearly related to household composition and if the random term is uncorrelated with the social marginal utility of income, it can be shown that indirect taxes should be uniform. Where a non-linear income tax is available, then weaker assumptions on preferences imply that indirect taxes should be uniform. Atkinson and Stiglitz (1976) and Deaton (1981) show that with identical preferences (households differ only in the wage) and separability of labour the optimal indirect taxes are uniform. Essentially this means, for the reason given earlier, that one can dispense with indirect taxes altogether (since raising all prices is the same as lowering the net wage) and operate only an income tax.

In the previous models the departures from the full optimality of the standard theorems of classical welfare economics arise only from the limitations of the tax system and not from market imperfections. It is remarkable, however, that there is a straightforward and considerable generalisation of the earlier theory to imperfect general equilibrium models.
where some markets have sticky (strictly, fixed) prices and thus do not clear. This is achieved by introducing the basic planning tool of shadow prices, defined as social opportunity costs.

It is straightforward [Dreze and Stern (1987)] to show that the standard theory of optimal taxes outlined above goes through if we replace producer prices by shadow prices. In the first-best world of standard welfare economics we have (relative) consumer, producer and shadow prices all equal. The Diamond–Mirrlees framework, where the only distortions arise from the tax system itself, has (relative) producer and shadow prices equal. More generally, planners are concerned with public finance problems where we have limited taxes and there are imperfections in markets and it is therefore reassuring that, through the use of shadow prices, the theoretical framework can handle these problems simultaneously, whilst retaining much of the structure of the optimal tax framework from the perfect market case. Broadly speaking we may say that, ceteris paribus, higher shadow prices (in relation to producer prices) indicate the desirability of higher indirect taxes on the relevant goods.

Some of the lessons of optimal tax theory are summarised below, concentrating on those results having particular relevance for the extension of theory from the static to the dynamic and beginning with the results on uniformity of commodity taxes. The issue of uniformity is important in the dynamic context when considering the question of income versus expenditure taxation since the question concerns whether goods consumed at different times should be taxed at different rates.

(i) The single-consumer optimal tax theory leads to a focus on departures from uniform rates of commodity taxation as resulting from differences in (compensated) cross-price elasticities of goods with leisure. Common cross-price elasticities imply uniformity.

(ii) The absence of lump-sum taxes implies that social marginal utilities of income will not generally be equal across non-identical households, thus providing support for distributional arguments for departures from uniformity. Only under rather special assumptions concerning the shape of preferences and the optimality of income taxation can the force of these distributional arguments be eliminated.

(iii) Externalities provide a reason for taxing some goods at different rates from others. Where feasible they should be personalised and in a second-best world account must be taken of the effects of actions of one individual on taxes paid by others. Some have argued that such issues are of importance in considering aspects of savings where some may benefit from the saving of others.

(iv) Efficiency considerations indicate that taxation of intermediate goods should be avoided.

(v) Distortions of prices resulting from market imperfections which cannot
be tackled directly can be ameliorated by appropriate adjustment of indirect taxes.

We note briefly two further themes that arise within tax theory and which are of relevance for dynamic extensions although they have not been discussed at length above.

(vi) Merit goods may provide a reason for government intervention. If the government thinks that an individual is not acting in her or his best interest or that some activities are of special intrinsic value (and not wholly acknowledged as such by an individual), then it may subsidise certain activities for this reason. Again it may be that such considerations have relevance for policy towards saving.

(vii) Where second-best taxation is in force, then there are general theoretical arguments for supplementation with rationing or regulation. The linear income tax provides a simple example. This causes a divergence between the marginal productivity of labour and the marginal rate of substitution of leisure for goods. The upshot is that a compulsory extra hour of work would provide production of a greater social value than the hour of leisure that is lost.

The relevance of these theoretical observations for dynamic theory is discussed in section 4.

3. The normative theory of growth

3.1. Orthodox growth theory

Before examining the extension of the microeconomic theory of taxation to dynamic models it is useful to have in front of us a brief summary of some of those elements of the basic macroeconomic models of growth which have something to say about policy. Where \( K \) is capital stock, \( Y \) output, and \( s \) the savings ratio, we have

\[
\dot{K} = sY, \tag{1}
\]

and therefore

\[
\frac{\dot{K}}{K} = s \frac{Y}{K} = s \frac{1}{v}, \tag{2}
\]

where \( v \) is the capital–output ratio. This is the basic growth equation identified by Harrod (1939), who called \( s/v \) the warranted rate of growth, i.e. that rate of growth warranted by the savings rate, the capital–output ratio
and output market equilibrium. If \( v \) is constant, then it will also be the rate of growth of output.

Solow (1956) embedded this analysis in a model with output determined through a production function \( F(K, L) \) showing constant returns to scale, and with full employment of labour, \( L \). He shows that an increase in the savings rate in this model will, in the short run, increase the rate of growth of capital, \( \dot{K}/K \) [see (2)], and therefore the rate of growth of output (since, on logarithmically differentiating the production function, this is seen to be a weighted average of \( K/K \) and \( n \), the rate of growth of labour, with weights given by the competitive factor shares). However, an increase in the savings rate cannot affect the long-run rate of growth since if \( s/v \) exceeds (falls below) \( n \), the capital–labour ratio and thus the capital–output ratio will rise (fall), a process which sees \( s/v \) asymptotically equal to \( n \). We shall see below, however, that the newer, so-called 'endogenous growth theories' move the warranted growth rate \( s/v \) back to centre stage in the long run as well as the short run, so that an increase in \( s \) can also increase the long-run rate of growth. These theories essentially involve increasing returns to scale so that, in a manner described below, the growth rate of the labour force is no longer the long-run constraint on output growth.

Disembodied technical progress of the labour-augmenting kind can be introduced into these models whilst retaining their basic structure. The technical progress is such as to make labour more effective: specifically 'effective units' of labour become \( A(t) = A_0 e^g \). These 'effective units' grow at \( g = a + n \), where \( a \) is exogenous. The analysis then goes through as before with \( s/v \) tending to \( g \) and with output per head and capital per head growing at rate \( a \). As before, increases in savings rates increase short-run growth rates, and output per head in the long and short run, but not the long-run growth rate.

An important advance on these models was the work by Arrow (1962) on learning-by-doing. He used a vintage approach (where technical progress is embodied in new machines and does not extend to old) with fixed coefficients but the productivity on new machines was taken to be influenced by experience as measured by the integral, \( G \), of the firm's past gross investments. The labour requirements on new machines were written as \( G - \phi \), where \( 0 < \phi < 1 \). Arrow shows that the model is capable of steady growth at rate \( n/(1-\mu) \). But notice that the long-run steady-state growth rate remains immune to policy. Of relevance for policy, however, is the externality associated with investment – the experience generated lowers labour requirements on future investments. In such models, therefore, we would expect the social rate of return on investment to be higher than the private, thus suggesting an investment subsidy. Sheshinski (1967) showed how the result concerning the growth rate could be obtained in a simplified non-vintage framework with the factor for labour augmentation, \( A(t) \), equal to \((K(t))^\gamma\) (see subsection 3.2 for further discussion).
The normative model which has been worked through most thoroughly has been that of optimal growth with one good. The constraints on the optimisation are: the initial capital stock, $K_0$; that consumption, $C_t$, and $K_t$ should be non-negative; the labour availability at time $t$, $L_0 e^{n t}$; and the production function, $Y = F(K, L)$. On rewriting (1) in per capita terms, where relevant lower-case letters indicate upper case letters divided by $n$, we have

$$\dot{k} = f(k) - nk - c. \quad (3)$$

The objective is the maximisation of

$$\int_0^\infty Lu(c) e^{-\rho t} \, dt, \quad (4)$$

where $\rho$ is a pure time discount rate. This is a utilitarian sum of utilities across periods. From the perspective of a planner with equal commitment to all generations the discount rate may be justified by interpreting $e^{-\rho t}$ as the probability that the world survives at least as long as $t$. The discount rate is important for the existence of an optimum since generally we require $\rho > n$ for a solution to exist. Otherwise, we run into problems of divergence of the infinite integral (which warns us that the meaning of 'maximisation' must be carefully defined).

The first-order condition for the optimisation of (4) subject to (3) is

$$- \frac{d}{dt} \left( u' e^{-\rho t} / u' e^{-\rho t} \right) = f'(k). \quad (5)$$

This is essentially the familiar tangency condition between the indifference curve and the production-possibility frontier in the standard diagram with consumption now on one axis and consumption next period on the other; it may be interpreted as the equality between the social discount rate (or rate of fall of marginal utility of consumption) and the marginal productivity of capital.

Eqs. (3) and (5) are a pair of differential equations in $c$ and $k$. Thus, given $c_0$ and $k_0$ they trace the development of the system which will be followed provided the non-negativity constraints are satisfied. The final problem is to select $c_0$, since we already know $k_0$. This initial condition is selected by looking at 'terminal' conditions. In a finite horizon economy (horizon, $T$, say) this condition would be $K_T = 0$. In an infinite horizon economy the condition becomes that the value of capital, $K u' e^{-\rho T}$, should tend to zero as time goes to infinity.
The main economic insights provided in the simple theory of optimal growth are: (i) the valuation of output, whether used as consumption or investment, as $u'c^{-a}$; (ii) the equality, (5), between the social discount rate and the marginal product of capital which follows essentially from (i); and (iii) the selection of initial consumption through the long-run destination. The models do not have much to say about taxation. As is clear from the positive structure of the model, nothing can be done to change the long-run growth rate. If change were possible, then the potential gains from policy could indeed be very large.

The simple one-good growth model has provided us with some basic grammar for understanding dynamic issues. The Harrod expression for the warranted rate, $s/v$, for example, is a tool for understanding growth rates and the overall structure of planning models, which is indispensable. It points straight to basic questions such as the determinants of savings and capital–output ratios, the relation between savings and investment, and so on. And the normative theory does lead to an understanding of simple rules and alerts us to incorrect argument. We see, for example, that the condition that the social discount rate equals the social rate of return does not by itself determine optimal savings and investment. There are many paths satisfying this condition and, generally, at most one of them is optimal.

Generally, growth theory has been used more for positive issues in the understanding of growth and for understanding savings and investment, where it has indeed been useful, rather than for tax policy. Where it has been applied to tax policy many of the issues have not turned on the detail of the dynamic structure. Furthermore, it has not, by and large, had much to say about the determination of the long-run growth rate.

3.2. Policy in models with endogenous technical progress

In the models described above the sole source of long-run growth in output per head was exogenous technical progress. From the points of view both of understanding the determinants of growth and of designing policy to influence it, this feature is highly unsatisfactory. Recently there has been a resurgence of interest in the theory of growth with a focus on endogenous (long-run) changes in productivity. The source of these changes is essentially some version of increasing returns to scale.

The models of endogenous technical progress which have been the focus of attention fall broadly into two types. In the first, technical advance comes from externalities in investment. The act of investment itself advances knowledge and this knowledge is available to others. From this perspective there is a basic externality and therefore not only is technical change endogenous but there is a role for policy in encouraging it via investment. This type of externality is the essential feature of the models by Arrow (1962)
and Sheshinski (1967) discussed above and also of Scott (1989) and Romer (1986, 1989). The second type has technical progress arising from the allocation of resources to the production of designs which may then be used in the production process [see Uzawa (1965), Shell (1973), Lucas (1988) and Romer (1990)].

We can illustrate the line of argument used in models of the first type with a simple example which Romer (1989) employs. To concentrate on the other sources of growth we suppose that labour is fixed. The representative firm-consumer is then assumed to maximise

$$\int_0^\infty u(c) e^{-\rho t} dt$$

subject to

$$\dot{k} = f(k, K) - c,$$  \hspace{1cm} (6)

where $c$ is consumption and $k$ the capital stock of the individual firm. The overall level of technology depends on the total acts or quantity of investment that have been made in the economy, i.e. the capital stock $K$ which is equal to $Nk$, where $N$ is the number of firms. Romer (1989) conducts some of his discussion through a simple example where $f(k, K) = k^aK^{1-a}$. The following gives the essentials of the argument.

The optimising social planner would recognise that $K$ is equal to $Nk$ and maximise (6) subject to

$$\dot{k} = N^{1-a}k - c.$$  \hspace{1cm} (7)

For an isoelastic utility function $u(c) = c^{1-\sigma}/(1-\sigma)$ (where $\sigma > 0$) the condition that the social discount rate is equal to the marginal product of capital (5) is just

$$\rho + \sigma \frac{\dot{c}}{c} = N^{1-a},$$  \hspace{1cm} (8)

so that the optimal rate of growth is $(N^{1-a} - \rho)/\sigma$. The representative firm-consumer does not, however, recognise the effect of $k$ on $K$ and we have for that problem, instead of (9), a different r.h.s. which is the marginal product of $k$ as seen by the firm:

$$\rho + \sigma \frac{\dot{c}}{c} = \alpha k^{\sigma-1}(Nk)^{1-a} = \alpha N^{1-a}.$$  \hspace{1cm} (10)

Hence the 'market' would give a lower rate of growth $(\alpha N^{1-a} - \rho)/\sigma$.

A policy that removes the externality would therefore raise the rate of growth. Following a similar line of analysis, a policy that reduces the rate of
return perceived by the firm, for example through an output tax at rate \( \tau \), would reduce the rate of growth to \( (s(1-\tau) + \rho)/\sigma \). Government policy now has a very clear and possibly substantial effect on the rate of growth.

The crucial point is that the externality takes a form (the level of technology being proportional to total capital) that, when combined with constant returns to scale, gives output proportional to capital. The Harrod–Domar expression \( (s/v) \) now applies to the long and short run. Hence policy that increases \( s \) or decreases \( v \) raises the long-run growth rate. An unsatisfactory feature about using this model to make the point is that if the level of technology increases with total capital less then proportionately [e.g. like \( K^\gamma \) with \( \gamma < 1 \) as in Arrow (1962) and Sheshinski (1967)], then we are back with the conclusion that the long-run growth rate is immune to policy [being \( n/(1-\gamma) \)]. That such an important conclusion turns on such a fine distinction (which is unlikely to be settled empirically) should make us uneasy about relying on the Romer (1989) model as a basis for explaining the role of policy in determining the rate of growth. On the more positive side, however, the analysis does suggest that the learning induced by investment may have substantial effects on growth rates which may endure for considerable lengths of time.

The second type of model has improvements in knowledge arising from resources being applied directly to the production of those improvements (see above references). The details are not provided here, but again we find that the socially optimal (long-run) rate of growth is higher than that of the private market. In the market economy, insufficient resources are allocated to research, and those ideas that are produced are not sufficiently disseminated owing to the monopolisation of designs. The difficulty is, of course, that if designs cannot be monopolised (or at least partially restricted to the owner’s use) they cannot be sold and therefore will not be produced in a private enterprise economy. One policy reaction would be for research to be conducted as a publicly-financed activity, with results made available for all.

4. Extending static tax theory to dynamic questions

4.1. The dynamic interpretation of the static model

In the preceding two sections we set out the basics of received static tax theory and of normative growth theory. For most of the latter, tax questions are not central. In this section we discuss issues, insights and problems that enter into a dynamic analysis of taxation but which do not figure prominently in static analysis.

\[ \text{This section draws on chapter 5 of Newbery and Stern (1987).} \]
Our point of departure is the reinterpretation of the standard Arrow–Debreu model of static general equilibrium as intertemporal where goods appearing at different dates are seen as different goods. There are a number of problems that arise with this reinterpretation, but it provides a useful organising framework in that we can examine these problems as consequences of failures of the assumptions of the Arrow–Debreu approach. We begin in this subsection by setting out the reinterpretation and by noting problems and some market failures that arise in a static context and that appear, although possibly more severely, in a dynamic context too. In the two later subsections we focus on problems that are essentially dynamic.

We discuss, in subsection 4.2, dynamic policy issues that have been treated with fairly minor modifications to the standard framework: the tax base (income, expenditure, capital and so on); the process of adjustment from one regime to another; and pensions and inter-generational transfers. In subsection 4.3 we consider problems that can involve radical departures from the dynamic reinterpretation of Arrow–Debreu: incorrect expectations; the development of knowledge and entrepreneurship; revision of policy and the honouring of contracts; the influence of future generations on current markets; and the dynamic modelling of individual choice. What is radical and what is minor are, however, matters of judgement and subsections 4.2 and 4.3 can be read together without insisting on any fundamental break between them.

The standard Arrow–Debreu model of general equilibrium can be interpreted as a full intertemporal model provided that goods are distinguished not just by their physical characteristics (and location) but also by their date of availability (and the state of the world). In this model, and under the intertemporal interpretation, the economy will be on an efficient growth path, and government intervention will be associated with efficient revenue raising and redistribution rather than with direct concern with the rate of growth. The Diamond–Mirrlees efficiency theorems would tell us that under their assumptions (see section 2), indirect taxation should not interfere with production efficiency, neither dynamic nor static. There should be no taxation on transactions between producers in capital markets, although there is no reason why consumers and producers should not face different intertemporal price vectors.

The assumptions of the above model under the dynamic interpretation are very strong (including existence of markets, perfect competition and rational expectations) and a discussion of some of the implications of their failure for the analysis of policy forms the basis of much of this section. Two problems that arise in both static and dynamic models may be of particular importance for dynamic tax policy. The first is that a number of markets, both insurance and forward, may not exist. It is difficult to know whether the analysis of policy for those goods, insurance, futures, and capital
markets that do exist is rendered seriously misleading by the absence of other markets. We know from the standard second-best theory that imperfections elsewhere in the system should generally influence policy in particular markets, but just how this might work out in this case is an open question. There is some theoretical experience – for example, we can discuss the taxation of externalities when the full set of personal taxes and markets does not exist [e.g. Diamond (1973a)] – but it is fairly limited for dynamic issues. We could, however, speculate as follows, giving one example in a static context and another in a dynamic context. In many developing countries there are constraints on women’s ability to work. This might mitigate in favour of a tax-cum-subsidy which is redistributed from men to women (for example, child support payments paid directly to women financed by taxes on men). From a dynamic perspective the ability to spread income over the future is limited by poor capital and insurance markets. This might provide an argument for a progressive income tax system where the government provides some ‘smoothing’ over time. A difficulty here is that the problem is most severe precisely where income tax systems are in a weak position to provide a contribution. Compulsory pension schemes or unemployment benefit contributions provide examples of relevance here.

Second, we have the difficulties of separating production and consumption decisions. Where household production is prevalent some production activities will be subject to the consumer prices, \( q \), rather than the producer prices, \( p \). The Diamond–Mirrlees efficiency theorems must then be reinterpreted as applying to that part of the production sector that is separate from the consumers rather than to production as a whole. This does not, however, prevent an analysis of optimal taxation which is now interpreted as applying to net demand functions from consumers and producer-consumers rather than ordinary household demand functions. The ‘jointness’ of production and consumption decisions is perhaps more marked in an intertemporal context where many small firms may be financed largely by household saving. The taxation of interest, capital and property income of households may then directly affect the cost of firm finance as opposed to indirectly through the tax wedges between savers and investors which are prominent in the theory of finance. It is possible that this might lead to a presumption in favour of lower taxes on interest income than if these issues were ignored, but the question should still be regarded as open.

4.2. Applications and extensions of the standard framework

In this subsection we consider three dynamic policy issues which can be set within the standard framework given by the dynamic reinterpretation of the static model. We discuss in turn: (i) the tax base; (ii) the process of adjustment; and (iii) pensions and transfers across generations. There is a
large literature and we shall not attempt a detailed survey [for extensive references, see Kotlikoff (1984) and Auerbach and Kotlikoff (1987)].

4.2.1. *The tax base*

The standard optimal tax arguments described in section 2 have been extended to simple intertemporal models by a number of authors [for example, Diamond (1973b), Atkinson and Sandmo (1980) and King (1980)]. In these models we can again consider the question of the uniformity of taxation across commodities, in this case consumption of goods in different periods. The proposition that taxes should be uniform then becomes the statement that an appropriate tax base is expenditure, because a tax on expenditure at the same rate in each period has the uniformity property in this context. Not surprisingly, separability between leisure and goods is again crucial, and we can show in special models – for example, those in steady state, with individuals working for only one period and consuming in two [see, for example, Atkinson and Sandmo (1980) and King (1980)] – that uniformity is optimal.

As should be clear from the discussion in section 2, this result is very sensitive to assumptions and would not usually hold if individuals work for two periods rather than one (recall that the uniformity results depend on there being a single factor that is supplied). In this context it is possible that a proportional income tax may be superior to a proportional consumption tax. We cannot, therefore, settle the issue between expenditure and income taxation on theoretical grounds [for further discussion, see Atkinson and Sandmo (1980), King (1980) and Auerbach, Kotlikoff and Skinner (1983)]. This theoretical ambiguity has led to two rather different approaches to the question of the appropriate tax base in a dynamic economy. The first [see Meade (1978) and Kay and King (1983)] notes the advantages of an expenditure tax on practical grounds – its main attraction is that it does not require a distinction between capital and income, a distinction that lies at the source of many administrative difficulties and tax dodges. Second, the absence of clear-cut analytical rankings has stimulated a number of economists to simulate the outcomes of using different tax bases in order that they may be compared for policy purposes.

The simulation models are based on the same set of principles as the static models in that they both assume optimising behaviour for individuals in order to derive and evaluate the consequences of tax changes. The optimising choices of individual households concern the allocation of consumption over time in response to relative intertemporal prices in an analogous manner to the static allocation of goods given prices and a budget constraint. We can then ask how a tax reform that changes intertemporal prices will affect revenue and welfare. The most popular type of exercise has been the replacement of the income tax by a consumption or wage tax [see, for
example, Auerbach, Kotlikoff and Skinner (1983) and Fullerton, Shoven and Whalley (1983)]. The simulation studies differ in their focus, but the introduction of dynamics usually involves the sacrifice of some of the detail of static studies, in particular distribution within a generation. In each of the studies cited movements from an income tax to a consumption tax generate annual gains (averaged over the path) of two or so percentage points of GNP. The main determinant of these gains is the increase in capital stock brought about by the big response of savings to the higher post-tax interest rates arising from the switch to consumption taxation. The interest elasticity of savings is, therefore, crucial to the results.

As Fullerton, Shoven and Whalley (1983) remark at the end of their paper (p. 22): 'Additional analysis with the model indicates considerable sensitivity to the elasticity of savings with respect to the real after-tax rate of return. Further efforts to narrow the professional consensus on the value of this elasticity would clearly aid policy evaluation.' More recently Auerbach (1989) considered differential taxation on different capital goods in an overlapping generations model with many capital goods. Uniform taxes are not optimal, but the benefits of deviations from uniformity are small. However, questions of the distribution of welfare, principal motivations for differentiation in the static context, are ignored.

For developing countries the tax base is largely on expenditure through indirect taxation, but this is mainly due to administrative inability to tax personal incomes. The indirect tax route to expenditure taxation does not allow as much tailoring for equity considerations as would a household or individual based tax system. And where incomes are taxed in developing countries it is mostly via corporations [for further discussion see Ahmad and Stern (1991) and Burgess and Stern (1992)].

4.2.2. The process of adjustment

Economists have considerable experience in the theoretical analysis and comparison of different policy regimes. These include some of the issues of tax design discussed in section 2. They have rather less experience and success in analysing the problems of transition from one regime to another. Many countries considering the dismantling of central planning are now facing the question of not only where they would like to go, but how they want to get there. The World Bank, no doubt fully aware of the efficiency theories of section 2, is fond of advising countries both to abandon tariffs and to broaden tax bases to increase government revenue thus raising the obvious question of how quickly an effective domestic tax base can be established. Much of the analytical work described in the previous subsection on tax bases has been confined to steady states.

There has, however, been some progress, both theoretical and empirical, in analysing the fiscal problems of transition and it will no doubt be an
important area for further work. The simulation analysis of Auerbach and Kotlikoff (1987), based on U.S. considerations, is, for example, much concerned with timing of taxes. They produce interesting examples of the way in which expectations, timing and duration of policies can themselves have important consequences with different short- and long-run effects. For example (p. 4): 'Deficits arising from income tax cuts of short duration “crowd in” saving and investment in the short run even though saving and capital formation are crowded out in the long run by such policies.' Harris, Heady and Mitra (1989) examined an intertemporal optimising model of taxation and investment, calibrated using Turkish data, and were interested in the process of growth and adjustment rather than steady states. They found that (p. 289):

The main policy implications are that it is inappropriate to use time-invariant shadow prices in cost–benefit analysis, that the response of shadow prices to changes in the economic environment [debt and oil 'shocks'] may not be unidirectional [i.e. they may be of different direction in the short and long run] and that the accounting rate of interest cannot simply be seen as a regulator of investment that responds to shortages of investment finance.

See also the contributions by Heady and Mitra (1992) and by Mitra (1992).

There has now been considerable experience of IMF/World Bank and other adjustment packages and these have been reviewed recently and instructively in the volume edited by Blejer and Chu (1989). The problems of transition pose important theoretical and immediate policy challenges and no doubt will be lively topics of further research.

4.2.3. Pensions and transfers across generations

Many governments have felt a responsibility to provide state pensions. There is, however, a basic question of whether these should be left to private markets. Individuals allocate income across private goods, it is argued, so why should they not choose their own intertemporal allocation? The question is an important one and there are a number of responses one can make. There may be, for example, as mentioned above, real problems involved with intertemporal capital markets and the administration of a large number of personal pension plans may be less costly for governments. Governments may be less risk-averse in long-term contracts than private firms. And finally the government may feel that it has to act to reinforce the prudent, pension-providing higher self in the divided individual personality against the more profligate, short-sighted lower self. No doubt this attitude can be derided as that of the paternalist or the nursemaid, but I suspect that it is genuinely, although perhaps implicitly, part of the argument in practice.
For a discussion of some aspects of these problems, see Akerlof and Dickens (1982) and, for modelling, Thaler and Shefrin (1981) and Shefrin and Thaler (1988).

It has long been argued that state pensions discourage private-sector saving. The debate was reopened by Feldstein (1974) who argued that the effect of the U.S. social security programmes was to dramatically lower savings and the capital stock. There has been a subsequent voluminous debate in which many types of model and sources of data (cross-section, time-series and panel) have been used. This debate has been thoroughly reviewed recently by Atkinson (1987, pp. 869–880) and thus will not be discussed extensively here. He presents a summary table of over 20 articles in which Feldstein’s conclusion has been challenged and many authors have come to different conclusions. Like the interest elasticity of saving it is an important question that remains substantially open, notwithstanding a large amount of research. And, similarly, it seems unlikely that we shall be able to get settled answers in the near future.

An incentive relevant in developing countries might operate through size of family. If another institution (for example, company or state) can provide for old age, then there may be less incentive to have large families – it is possible that this is one of a number of factors in the explanation of China’s birth rate standing at a lower level than other developing countries.

4.3. Problems for the standard framework

In this subsection we draw attention to five dynamic modelling issues, and discuss the difficulties they pose and some of the implications for taxation and planning. We discuss in turn: (i) incorrect expectations; (ii) the generation of knowledge; (iii) dynamic inconsistency; (iv) the absence of consumers from markets; and (v) modelling intertemporal behaviour. Whilst some aspects of these problems have analogues in the static theory, crucial elements of each of them are not only essentially dynamic in nature but can take us substantially beyond the Arrow–Debreu framework.

4.3.1. Incorrect expectations

The Arrow–Debreu reinterpretation assumes that all futures markets exist and trades can be made now. This is formally equivalent to trades being made later at prices that are perfectly foreseen, provided expectations are common to all agents and correct. There can, in principle, be a role for government price and planning policy, therefore, where expectations differ (so that not all of them are correct) or where there are shared, but incorrect, expectations. This could be via indicative planning [Meade (1970)] whereby surveys of intentions are made, aggregated and published to help individual
agents plan in the context of what others are planning, or by taxation or government direction. The attractiveness of policy motivated in this way does depend, however, on the government believing that it knows something that other people do not. In indicative planning that something is, in principle, a view of the sum total of individual decisions.

An example of a policy based on a different view of expectations might be a subsidy on air-raid shelters if the government believes the probability of war to be higher than perceived by individuals. Alternatively a government might impose special building standards arguing that the probability of an earthquake is systematically under-estimated by the population.

4.3.2. The generation of knowledge

The accumulation of knowledge and the rewards to arriving first or early would seem to be essentially intertemporal notions. In a number of theories the provision of knowledge is like an externality which flows, as a by-product, from investment [see, for example, Arrow (1962), Romer (1986) and Scott (1989)]. On these grounds the market might provide insufficient investment and there may be an argument for subsidy or for public investment. The argument does, however, look rather like the familiar static externality one, even though the learning may be essentially a dynamic process.

Knowledge may, however, also arise from investment going directly into research and development. Here matters can become more complicated. To the extent that knowledge is common property, private R&D investment would be seriously discouraged. These disincentives may be overcome by institutional changes such as establishing or strengthening intellectual property protection (patents, copyright, trademarks, trade secrets legislation) or offering prizes. If the government were to carry out such investment and make the results universally available, then, from the theoretical point of view, the situation would be fairly straightforward. But much of the activity might be better carried out in the private sector where researchers may be closer to markets, the production process and have better incentives. To encourage more private R&D investment the government might subsidise R&D expenditures, issue research contracts or establish venture capital funds or private production problems and activities.

The problems associated with search for knowledge may result in multiple equilibria, some of which may be superior to others. Diamond (1981, 1984), for example, points to this possibility. Government policy may then be oriented towards shifting from one to another.

4.3.3. Dynamic inconsistency

Of the problems special to dynamic economics perhaps the one that has received most attention in the policy literature is that of 'dynamic consist-
ency'. The issue was raised by Kydland and Prescott (1977). A clear description in the tax policy context is provided in Fischer (1980) and a survey of related macroeconomic work by Persson (1988). Regarding tax policy, the issue of dynamic consistency arises in a second-best tax world where first-period revenue cannot be raised in a lump-sum manner. The idea is that the future tax policy announced by the government last period will no longer be optimal when it comes to implementation this period not because the future has developed in an unexpected way, but because the passage of time makes certain disincentives now irrelevant. Thus, last period the government may have announced that capital taxation in the second (now current) period would be low in order to encourage accumulation. However, when the second period comes the government could simply announce a capital levy, thus reneging on its promise. This levy would be a lump-sum tax with, ex post, no distortionary implications. Ex ante, however, individuals may recognise that the government will be tempted in this way and adjust their first-period behavior accordingly. The analysis of optimal policy requires a careful specification of the game between individuals and the government.

Related issues may arise with changes in government. In this context incentive effects can arise that would not be present with a consistent government. Take, for example, the switch from direct to indirect taxation in the first budget (June) of the incoming Conservative Government in 1979 when VAT was raised from 8 to 15 percent and the basic rate of income tax was cut from 33 to 30 percent (with the top rate coming down from 83 to 60 percent together with increases in the personal allowances). It was suggested that there was a presumption that shifting taxes to spending and from income would provide an incentive to work more. As a general statement this is clearly wrong since an increase in the take-home wage and a corresponding increase in prices would leave real wages unchanged and, in the absence of wealth effects, would not change the incentive to work. If, however, one foresaw the prospect of a Labour Government (one would have had to be far-sighted in 1979), then an attractive strategy for an individual might have been to work during the Conservative administration and spend in a Labour one. Notice that the increase in consumer prices from the switch to indirect taxation does, in part, act like a capital levy.

Illustrations of the problems of asymmetric information and lack of trust, arising from incentives to renege, are associated with the bank finance of investment as discussed by Stiglitz and Weiss (1981) and de Meza and Webb (1987). The former suggest that there will be under-investment because bankers are risk averse and unable to judge the riskiness of different prospects with similar expected values. The latter point to the possibility of over-investment since borrowers may be happy to gamble because bankruptcy sets a floor to their losses.
4.3.4. The absence of consumers from markets

Consumers who do not yet exist cannot trade in current markets. The Arrow–Debreu model assumes that everyone can trade. We could assume that families foresee correctly the existence and preferences of future generations and take these preferences as their own. However, there are a number of problems associated with this 'dynastic' assumption.

First, the current generation may not perceive perfectly, or care fully, about the welfare of future generations in the sense that the welfare of future generations might get less weight in the dynastic welfare function than some exogenous ethical commentator might regard as appropriate from the point of view of social welfare. Second, the dynasty might worry about whether it would reap the full returns from bequests. For example, some might go to other dynasties through inter-marriage and some through taxation. This is the essence of Sen's so-called 'isolation paradox' [Sen (1967, 1982)]. From both these points of view there might be some presumption that the market-determined rate of saving is less than would be socially desirable.

Policy reactions to this might include subsidising, or providing special tax treatment for, saving. This is similar to the theory of policy towards externality. Alternatively it may be seen as the use of a tax tool to encourage saving to offset problems caused by the operation of tax tools elsewhere (the taxation of income or spending in the future). Furthermore, the government might give special weight to saving in appraising projects or other policies.

4.3.5. Modelling intertemporal behaviour

Individuals may find the planning of their lives over the indefinite future to be confusing and complicated. One of the difficulties is that preferences can depend on consumption experience. One reaction of the economist might be to say, 'Never mind if they may actually find it confusing; let us make the hypothesis that they behave "as if" they do their intertemporal optimisation thoroughly and then test the hypothesis.' Often, however, the theoretical and simulation models are constructed on this basis whether or not the hypothesis is tested and whether or not it receives support if tested. Unfortunately intertemporal optimising models are not always easy for the economist to solve either.

An alternative would be to suppose individuals follow simple rules of thumb such as a constant savings rate. The great advantage of such simple rules as models of behaviour is that they typically allow the positive model to be solved more easily. We must take care, however, with the normative models, because we may have to specify social welfare in the absence of a formal model of utility maximisation by individuals. Nevertheless, if we take the rule of thumb as being broadly consistent with an attempt by the individual to improve utility in a complex environment, then it is reasonable to write down a social welfare function that formalises the utility accruing to
the individual from that individual's consumption stream which arises in part from his or her own behaviour.

5. An assessment

The theory of optimal taxation has not had a great deal to say about dynamics and the theory of growth has been reticent on taxation. This has clearly not been because the developers of one have been ignorant of the other – the number of authors who have contributed to both is striking. Rather the reasons include the following: dynamic optimising models are more difficult analytically than static ones; we know much less about the dynamic behaviour of individuals; and there are important logical and theoretical problems that arise for dynamic models that do not appear in the static.

Our purpose in this paper has been to try to draw together the relevant basic elements of the two theories and then to discuss some of the lessons from and the problems of combining them. From the theory of optimal taxation there are four groups of results that carry fairly direct lessons for dynamic economics. First, we have results on the circumstances under which indirect taxation is optimally uniform. These focus on cross-elasticities with leisure and on whether a well-functioning direct tax system is in place. Uniformity can correspond to expenditure taxation if translated into a dynamic context. Second, we have the intertemporal analogues of the efficiency theorems which suggest that inter-firm dynamic transactions should not be taxed. Where the household and firm overlap, then this might suggest lower taxation on household savings than might otherwise be the case.

Third, we have taxation for externalities or to correct for distortions introduced elsewhere in the tax system. This type of idea has been behind discussions of the 'isolation paradox' and the desirability of encouraging saving beyond the levels achieved in the private market. The idea is that through inter-marriage and taxation others reap the rewards of my saving and hence the social return is higher than the private. Fourth, the government may need to take greater responsibility for future generations than current generations are prepared to do and, indeed, should take greater responsibility for an individual's own future than they are prepared to take themselves (a version of the merit-good argument).

The contribution of growth theory was to tutor us on the logic both of the determinants of growth – a simple but crucial example being the expression given by Harrod for the warranted growth rate $s/v$ – and of intertemporal optimisation, for example that the social discount rate should be equal to the social rate of return. Growth theory pointed also, however, to what we are bad at explaining, particularly technical progress. There have been some efforts in recent years to improve our understanding here. However, we have,
perhaps, not yet gone much further than pointing to the problems of the externality and publicness of ideas. These problems may mean that insufficient resources are allocated to the creation of ideas, either indirectly through the act of investment or directly through research and development. If these activities do influence the long-run growth rate, then the importance of institutions, government activities and tax-subsidy policies that encourage them may be substantial.

We should perhaps include in our conclusions a note of empirical humility. There are crucial empirical phenomena on which we are worryingly ignorant. On the production side the role of investment in stimulating growth is poorly understood both in terms of why it is that incremental capital–output ratios are much higher in some circumstances and countries than others, and in terms of the workings of technical progress that we have just mentioned. On the demand or savings side we remain unsure not only about the interest elasticity of savings, and how it varies within the population, but also about the appropriate models within which the elasticity should be estimated. We do not know then whether tax policy is very important in its effect on savings rates. I would be sceptical whether it has a lot to do with the rise in the savings rate in India in the decades following independence or in the fall in the rate in economies in sub-Saharan Africa.

The ways in which dynamic models raised issues that were intrinsically new relative to the static included: incorrect expectations; the development of knowledge; revision of policy; the influence of future generations; pensions and inter-generational transfers; the tax base; and the process of adjustment. Each of them could be of great significance for tax policy and together they imply that the simple intuition developed in static theory may have only a minor, though non-trivial, role to play in the understanding of dynamic tax policy. For example, the provision and funding of state pensions raises very large questions for tax and expenditure policies. Central reasons for the provision of pensions are poor capital markets, so that individuals are limited in their ability to reallocate resources over time, and the problems of human weakness and divided identity which may prevent the young from taking full account of future needs. Cultural and institutional arrangements such as large families might work in part to alleviate these problems. We can see, however, that a standard reinterpretation of the static model which treats goods at different dates as different goods but which does not change the tax theory, does little to capture these particular issues. Similarly, the limitations on credible policies introduced by the possibility of ‘dynamic inconsistency’ radically changes the nature of tax analysis.

We have seen, therefore, that there are some issues on which static tax analysis carries quite useful dynamic lessons and others where the issues with which it is concerned provide a very limited, possibly a rather small, part of the answer to the questions at hand.
References


