I. INTRODUCTION

The choice of how much of society’s income to consume today and how much to save for future consumption is among the most important economic decisions. It can have implications not only for the well-being of the households taking the decisions, but also for the rate at which the economy invests and grows, and therefore the well-being of future generations. The way in which savings decisions are taken is a matter of ongoing research in the literature, involving such important and unresolved issues as the degree of foresight and rationality of households, the extent to which capital markets are complete and well-functioning, the importance of life-cycle versus bequest versus precautionary motives for saving, and more generally the weight, if any, that current savers put on the welfare of future generations.

Taxation is one of many policies that affect the level of savings. Other significant ones include the system of social insurance, especially pensions, but also health and disability insurance, welfare services and education; transfers of various sorts; and debt policy. None the less, the principles by which saving behaviour responds to these various policies are very similar. We begin our survey with a discussion of these principles, first at the household level and then in the aggregate. Then, the effects of taxation are addressed. This is done first from a theoretical perspective, followed by a discussion of some of the main currents in the empirical literature on taxation and savings. Finally, we address some important normative or policy issues that have arisen in the literature, focusing mainly on two issues. One is the optimal tax treatment of savings —
should capital income be taxed? If so, how? The second is the more general issue of what the optimal rate of saving in the economy is.

II. THE DETERMINANTS OF SAVING

There are several aspects of saving that differentiate it from other consumer optimisation decisions and that turn out to be important for tax analysis. For one thing, the ultimate object of choice is future consumption, of which savings is the value rather than the quantity. The price of future consumption in terms of current consumption is the after-tax discount factor. Thus, saving is the discounted present value of future consumption. Even though a fall in the price of future consumption (a rise in the after-tax interest rate) may cause a rise in future consumption, it may involve a fall in the level of savings since a given amount of future consumption can be obtained at a lower cost in terms of forgone current consumption.

A second distinguishing feature of saving concerns the nature of the budget constraint. An individual’s income may take the form of a stream of earnings (or other forms of non-capital income) in both present and future periods, so that the budget constraint involves a present value of earnings, referred to as lifetime wealth. This has two implications. First, a given stream of consumption over time will involve different levels of savings depending upon the time profile of the earnings stream. The later that earnings occur, the less savings (or more dissavings) will be required to finance a given amount of consumption. Second, a change in the price of future consumption (the after-tax discount rate) will cause a change in the present value of earnings or lifetime wealth. This change will be larger the more is the earnings stream skewed towards the future because the discount factor will be higher for earnings further in the future. This induced change in lifetime wealth caused by a change in the price of future consumption will give rise to an indirect effect on household saving behaviour, referred to as the human wealth effect, that is not present in standard consumer choice models (Summers, 1981).

The intertemporal nature of the saving decision gives rise to various other unique issues. For one thing, there will be uncertainty about the future, and varying amounts of it may be borne by the individual saver. For another, to the extent that the individual is a dissaver in certain periods, there may be liquidity constraints which will restrict the individual’s choice. Also, there will be a variety of instruments for converting present into future consumption, including financial assets (debt, shares in firms, annuities, pension funds, mutual funds, insurance policies), real property, unincorporated business assets, consumer durables (including housing) and even investment in human capital formation. These distinctions are relevant because different savings instruments are typically treated differently for tax purposes.
Finally, saving is often seen as a vehicle for achieving altruistic objectives, particularly altruism within the family. The existence of altruism turns out to have a profound effect on the way in which taxation affects savings, as well as on the way in which we conduct our economic analysis of saving. Saving to account for altruistic motives gives rise to bequests and gifts to younger cohorts.

1. Individual Savings

The simplest case to consider is that of a selfish individual whose economic life can be divided into two periods, who receives earnings $Y$ in the first period only, faces perfect capital markets with full certainty and whose only decision is how to divide the earnings into first-period consumption, $C_1$, and second-period consumption, $C_2$, so as to maximise a lifetime utility function in $C_1$ and $C_2$. The budget constraint facing the household states that the present value of consumption, discounted at the after-tax interest rate, equals earnings. The combination of present and future consumption chosen will be such that the marginal rate of substitution between the two is just the relative price of future consumption, $p = 1/(1+r(1-t))$, where $r$ is the interest rate and $t$ is the tax on interest income.

The amount of saving, $S$ (i.e. $Y - C_1$), will depend upon the earnings of the household and the relative price, $p$. An increase in $Y$ will typically increase $S$, since the household will want to increase both present and future consumption (assuming them both to be normal goods). However, an increase in $p$ (reduction in $r(1-t)$) will have offsetting effects on savings. The higher price will tend to make the consumer substitute present consumption for the more expensive future consumption (the substitution effect), thus reducing savings. At the same time, the higher price of future consumption makes the consumer worse off since less of both goods can be bought with a given amount of income. This reduction in real income tends to cause the household to reduce both present and future consumption and so to save less (the income effect). Although future consumption will definitely fall, whether saving rises or falls with a fall in the interest rate depends upon the relative strengths of the income and substitution effects. Broadly speaking, the more the consumer is willing to substitute present for future consumption (the less the curvature in the household’s indifference curves between $C_1$ and $C_2$), the greater will be the substitution effect relative to the income effect, and the larger positive (smaller negative) will be the responsiveness of saving to increases in the after-tax interest rate. The issue becomes an empirical one.

The above discussion assumes that all earnings are in the first period. If there are second-period earnings, there will be additional effects. The household problem now involves maximising lifetime utility subject to a budget constraint which says that the present value of consumption must equal the present value of earnings, or lifetime wealth. An increase in earnings in either period will
increase lifetime wealth, and thus will increase both present and future consumption if both are normal goods. However, the effect on savings will depend upon the period in which the earnings increase. Increases in first-period earnings will increase savings as above. However, increases in second-period earnings will reduce savings since the individual will need to borrow against the future earnings increase in order to increase first-period consumption. This life-cycle timing effect is of crucial importance to the analysis of the effect of taxation on savings.

An increase in $p$ (reduction in $r(1 - t)$) will have the usual income and substitution effects as above, but there will now be an additional effect, the human wealth effect. The increase in $p$ will increase the present value of future earnings, thereby causing lifetime wealth to rise. This will increase the demand for $C_1$, thus reducing $S$ unambiguously. In other words, the human wealth effect will unambiguously increase the interest elasticity of saving. For a given initial value of lifetime wealth, the interest elasticity of saving will be higher the higher is the proportion of earnings accruing in the second period.

Human Capital Investment

There are various extensions that could be added to this two-period model of savings. The first involves human capital investment, which effectively makes the earnings stream endogenous. Suppose individuals can convert forgone present earnings into increased future earnings by devoting some of their time to education or training rather than working. The stream of earnings will then be chosen as that at which the marginal rate of return to human capital formation (the increment in future earnings from a marginal reduction in current earnings) just equals the after-tax discount factor defined as above. A reduction in $r(1 - t)$ increases investment in human capital, causing current income to fall and future income to rise. This causes saving to fall, thereby reinforcing the human wealth effect in increasing the interest elasticity of saving. In effect, the individual has substituted human wealth for asset wealth (Davies and St.-Hilaire, 1987).

Bequests

Another important extension to the two-period model is to allow for bequests. There are three main reasons for bequests: uncertain lifetimes, altruism, and strategic behaviour with respect to children.

**Uncertain lifetimes.** Uncertain lifetimes will result in involuntary bequests if the individual has to self-insure against the possibility of living a long life. In the two-period model, individuals may or may not live to the second period, but precautionary wealth must be held to finance consumption in case they do. If

---

2 Of course, earnings could also be made endogenous by allowing the labour supply to vary. Since we are concentrating on savings decisions, we ignore this possibility for simplicity.
they die early, the wealth is passed on to the next generation. This motive for saving would disappear if there were well-functioning annuity markets in which the individuals could insure themselves; individuals would behave as if their lifetimes were known with certainty and the above analysis of the effects of earnings and after-tax interest rate changes would apply. An additional possibility might now be forms of income received in the second period which are contingent on being alive, such as public pensions. Since these are like annuities, they reduce the need for precautionary saving (Abel, 1985).

Altruism. Altruism towards one’s heirs implies that households take decisions not only with their own lifetime utility in mind, but also with the well-being of their heirs. They will allocate their lifetime earnings between their own consumption and that of their heirs such that at the margin they are indifferent between the last pounds of own and heirs’ consumption. An increase in income will partly go to consumption for one’s heirs as well as one’s own consumption. Thus, there will be an additional factor tending towards an increase in saving out of income increases. An interesting case is that in which all households care for the level of utility attained by their immediate heirs. In this case, indirectly they care for their heirs’ heirs (since this affects the utility of their heirs), and for their heirs’ heirs’ heirs, and so on into the indefinite future. In taking decisions about how much to leave to their immediate heirs, they behave as if their own utility function included the consumption of all future heirs. This is referred to as a dynastic utility function, or a utility function with an infinite time horizon. The intertemporal allocation of consumption across various members of the dynasty can be determined by maximising the dynastic utility function subject to the present value of wealth available to the entire dynasty into the infinite future. To the extent that this is the case, it has an important policy implication. Consider an increase in an individual’s income which is offset by a reduction in the income of their heir (or their heir’s heir etc.) such that dynastic wealth is unchanged. In this case, the pattern of consumption across members of the dynasty will remain unchanged. The income transfer will be exactly offset by changes in bequests in the opposite direction. This idea that intergenerational transfers imposed on households by, say, the government will be completely undone by changes in bequests is known in the literature as the Ricardian Equivalence Theorem (Barro, 1974).

Strategic bequests. Finally, bequests may be ‘strategic’ in nature, given by selfish parents conditional on their children performing certain duties or services, such as staying in the family business, caring for parents in old age etc. (Bernheim, Shleifer and Summers, 1985). The greater the promised bequest, the more services might the children be expected to perform. Strategic bequest models focus attention on the effect of tax and other policies on the bargaining

---

3 There is an enormous literature on the issue of Ricardian equivalence and the circumstances under which it applies. A summary of opposing positions can be found in Broadway and Wildasin (1993).
strength of parents and children within the family; whereas altruistic parents
would simply pass incremental social security benefits on to their children as
bequests, strategically- motivated parents would use the prospect of higher
benefits to elicit additional assistance or other transfers from their children. The
strategic-bequest model may also have important implications for the effect of
estate taxation on savings. Estate taxes effectively increase the price of child
services relative to current consumption, reducing the demand for these services
and hence the level of parental saving.

The Multi-Period Case

The results of the two-period model apply in the multi-period case, along with
some additional results. Consider the selfish consumer who obtains an
exogenous stream of earnings over a given (working) portion of the life cycle
and converts it to a stream of consumption. The basic results can be illustrated
supposing that lifetime utility is just the sum of utilities of consumption achieved
in each period, discounted at the rate $\delta$. In this case, the marginal utility of
consumption falls at the proportional rate $r(1 - t) - \delta$, which is generally taken to
be positive. If the elasticity of the marginal utility of consumption is constant and
denoted $\eta$, consumption rises at the proportional rate $(r(1 - t) - \delta)/\eta$. All saving is
for life-cycle smoothing purposes. The savings profile is hump-shaped, rising in
the early part of the life cycle, then falling, and eventually becoming negative as
wealth is run down in retirement. Saving may also be negative at the beginning
of the life cycle. Define lifetime wealth in any period as the sum of asset wealth
and human wealth (the present value of future earnings). Consumption in each
period can be shown to be proportionate to lifetime wealth in that period, where
the propensity to consume rises with age, but can rise or fall with the after-tax
interest rate (Beach, Boadway and Bruce, 1988). Since the present value of
future earnings falls with a rise in the interest rate, the human wealth effect from
a rise in the interest rate causes a decrease in current consumption.

Suppose income increases unexpectedly in some period. Lifetime wealth then
rises, as does the lifetime consumption profile. Saving will increase in the short
run as well as in some following periods, and the asset accumulation profile will
shift upwards. If income rises in one period and falls in another, there will be a
change in saving even if the lifetime consumption profile is unchanged. For
example, an increase in income in some future period accompanied by a decline
in income now which leaves the present value of lifetime income unchanged (as
would occur with a funded public pension scheme) will cause no change in
lifetime consumption. However, saving will fall this period to finance the same
stream of consumption out of a smaller amount of income. Or, a transfer of
income from a young person to an old person will cause a rise in total

---

4 If instead the increase in wealth is anticipated, consumption will rise in earlier periods and saving will fall to
finance the increase in consumption.
consumption because the propensity to consume out of wealth is higher for the old than for the young.

Changes in the interest rate have a somewhat more complicated effect on the consumption and asset accumulation profile. An increase in the interest rate will cause the consumption profile to become steeper so that over the life cycle more assets would have to be accumulated to finance future consumption. However, in the short run, savings could rise or fall since the propensity to consume out of current wealth could increase or decrease with the interest rate. Note that the effect of an interest rate change on saving is likely to differ with age. The short-run interest elasticity of savings should be higher for younger persons than for older ones because the human wealth effect is greater.

Finally, in the multi-period model, desired savings and wealth holdings may well be negative in the early part of the life cycle when earnings are relatively low. However, liquidity constraints may prevent an individual from holding negative wealth, so that consumption cannot exceed earnings when young. Later in the life cycle, earnings will be high enough that the individual becomes a saver. From then on, the consumption stream follows the standard increasing pattern. With liquidity constraints, fluctuations in income in the constrained part of the life cycle result in one-for-one changes in consumption. Policies which change the timing of tax liabilities across the life cycle will thus have quite different effects when the household is liquidity-constrained. Moreover, a decrease in the interest rate which would ordinarily cause the individual to increase consumption early in the life cycle can no longer do so. It simply causes the individual’s constraint to become tighter.

2. Aggregate Savings

Determining aggregate savings involves more than simply aggregating the savings of households. One should also account for corporate savings (retained earnings). As well, public sector savings should be included if the government budget is not balanced.

Corporate Savings

With perfect capital markets and no capital income taxes, corporate and personal savings would be perfect substitutes. If capital markets are characterised by imperfect information, corporate savings may have some advantages to shareholders. For example, internal financing may be less costly to a firm if its managers have better information about the profitability of the firm than do outsiders. There is a large literature on the implications of imperfect information on the financial structure of the firm, much of which has only limited relevance for tax issues. The structure of capital taxes (corporate and personal) typically introduces advantages to saving within the corporation as well. Unless the corporate tax is fully integrated with the personal tax and all corporate-source
income imputed to shareholders as it accrues, there will be a tax advantage from retaining funds within the corporation rather than paying them out as dividends (Boadway and Bruce, 1992). This is referred to as the *trapped equity effect*, and can be used to explain not only the preference firms will have for retained earnings versus outside finance but also the incentives for the take-over of immature firms by mature ones (or their merging).

In taking account of corporate savings, the issue is whether or not households ‘see through the corporate veil’ and treat saving done on their behalf by corporations as part of their own saving (as they should). Assuming they do, then one should be able to think of individual saving as including retained earnings on their shares. This means that the tax applicable on these savings should include both personal and corporate taxes. In this paper, we ignore the complications introduced by the corporate tax and corporate savings.

The Consequences of Public Sector Savings

To see the implications of public sector saving (or dissaving) in the determination of aggregate savings and its changes, consider a simple economy consisting of overlapping generations of identical individuals who leave no bequests. At a given time, individual behaviour differs only by age: younger cohorts will be saving and older ones dissaving. The sum of individual savings plus government saving will equal aggregate investment; equivalently, the capital stock will be the sum of household wealth holdings less government debt. Most of the issues involved in aggregation can be seen by concentrating on policies which involve income changes alone.

*Revenue-neutral income changes for each household* Consider first the hypothetical case of a policy which changes the after-tax income stream of a given cohort but keeps its present value the same. For example, each member of the cohort might pay lower taxes (receive transfers) early in life and higher taxes later, such that the present value of the changes is zero. The typical individual in this cohort will save more when young to spread the income change over the life cycle to keep the consumption stream unchanged. In the absence of any other changes, aggregate saving and the demand for assets would increase temporarily while that cohort was young, then fall temporarily while it was old, but remain unchanged after that cohort had passed away. However, from the point of view of the government, net revenue would be lower while the cohort was young and higher when it was old. To make good the tax–transfer change without affecting other households in the economy, the government would have to borrow money. The amount of the public dissaving would exactly offset the increased private saving, and total saving would remain unchanged. In effect, the increased saving would go entirely to purchase the new debt issued by the government so that private capital investment would remain unchanged. More generally, the same principle would apply to a permanent tax reform which took increased taxes
Taxation and Savings

from the old and reduced them for the young. In order that the present value of taxes of all cohorts remain the same, this would have to be accompanied by an increase in government debt. The demand for assets by households would have permanently increased, but so would the supply of assets from the government by an equal amount. The economy would remain unchanged in real terms and all households would be equally as well off.

Revenue neutrality at each point of time. The picture is somewhat different if changes in government debt are ruled out. Consider a scheme in which transfers are made permanently from young persons to old persons, and the scheme is self-financing at each point in time. Examples of such policies include tax reforms which move tax liabilities earlier in the life cycle, unfunded public pensions, and in-kind transfers of public services to the elderly financed by taxes on the young. When the scheme is first introduced, all members of the older cohorts obtain an increase in their income and are better off. The incomes of individuals born after the policy change will be lower when young and higher when old; they will save less for retirement. Thus, the demand for assets will decrease and private investment will be lower. Whether such individuals are better or worse off depends upon the characteristics of the economy. The money they contributed to the scheme when young would have earned a return equal to the rate of interest $r$ (neglecting capital income taxes). On the other hand, a continual scheme of transfers from the young to the old will earn an implicit rate of return equal to the rate of growth of the economy $g$. If $g = r$, the net present value of the intergenerational transfer is zero so that individual welfare is unaffected in the long run, even though the level of private capital will be lower. The economy is said to be on the Golden Rule growth path. If $g > r$, the scheme will increase not only the welfare of the current older cohort, but also that of all future cohorts. In this case, the stock of capital is above the Golden Rule level; the economy is said to be over-capitalised. The intergenerational transfer scheme is Pareto-improving. And if $r > g$, the scheme will make young and future cohorts worse off while it makes the older better off. In this case, a scheme of intergenerational transfers from the young to the old will make the old better off at the expense of future generations. Of course, the reverse is true as well.

The above effects of intergenerational income transfers depend upon the assumptions of the model being used. If, instead of being selfish, individuals are

---

5 In this example, the same result would apply if households were altruistic towards their heirs. Since no intergenerational transfers are involved, there is no need for bequest behaviour to change in response to the tax–transfer scheme.

6 To see this, suppose population grows at the rate $n$ and that there is no technical progress, so $g = n$. In the two-period life-cycle case, if $T$ is the tax per worker, the transferee will be $T(1+n)$. The present value of the scheme over the life-cycle will be $-T/T(1+n)(1+r)$, which will be positive or negative according to whether $n$ is greater or less than $r$. More generally, the same result would hold if cohorts if cohorts live for several for several periods and if technical progress occurs. The implicit rate of return on the intergenerational transfer schemes would be $g$. 

altruistic towards future generations, at least part of the effects of intergenerational transfers will be offset by changes in bequest behaviour. Thus, a scheme which transfers from the younger to the older will induce the older cohorts to increase their bequests. Changes in bequests will fully offset the effects of the transfer if Ricardian equivalence applies (Barro, 1974). The effects will also differ from the above model if there are capital market imperfections. If very young workers face liquidity constraints, an increase in the tax on them will serve to reduce their consumption rather than their saving, and the adverse effect of the scheme on aggregate saving will be lessened (Hubbard and Judd, 1987). Similarly, if annuity markets are imperfect or absent, an intergenerational transfer from the young to the old will provide a sort of annuity. This will reduce the demand for precautionary saving and will cause aggregate saving to fall by even more than described above (Abel, 1985).

III. THE EFFECTS OF TAXES ON SAVING

The effect of taxes can be inferred readily from the above discussion. Taxes influence saving through two main mechanisms. First, they may affect the rate of return to saving and, second, they may affect the income stream which is converted into a stream of consumption (and possibly bequests) through saving. These effects can be seen by concentrating on three forms of taxation — wage taxation, consumption taxation and capital income taxation. (Income taxation is equivalent to capital income taxation combined with wage taxation.) For simplicity, we assume that taxes are at proportional rates.

1. The Effect of Taxes on Individual Saving

Consider an individual who obtains an earnings stream for a given number of working periods, and then is retired for a given number of periods. Capital markets are perfect so the individual is able to borrow freely against future earnings. As well, annuities are available to insure against uncertainty in the length of life, so that we can proceed as if the age of death is fixed and known. There are no bequests, so saving is for life-cycle smoothing purposes alone. The consumer maximises the discounted sum of utilities of consumption in each period subject to a budget constraint which says that the present value of lifetime consumption expenditures equals the present value of lifetime earnings after tax. All taxes are incorporated into the budget constraint. Consumption expenditures include taxes on consumption; the earnings stream is net of wage taxes; and earnings are discounted by the interest rate net of capital income taxes. In these circumstances, the effect on saving of each of the three taxes is straightforward.

Wage taxation. A tax on wages reduces lifetime wealth and leaves the after-tax interest rate unchanged. The time profile of consumption over the life cycle will shift downwards though its shape will remain the same. Similarly, the
time profile of earnings will shift downwards during working periods. Saving will fall during the working years since less consumption during retirement needs to be financed, and fewer assets will be demanded. The earlier in the life cycle are earnings obtained, the greater will be the decrease in saving. Equivalently, the later in the working part of the life cycle is a wage tax imposed, the higher the proportion of the tax will go to reduced savings.

**Consumption taxation.** A consumption tax causes the consumption stream to fall in exactly the same way as for a wage tax. It does so not by lowering the earnings stream, but by increasing the cost of consumption. However, because the tax liability occurs later in the life cycle, saving will not fall as much. In fact, in the case of a lifetime utility function which is the discounted sum of identical per-period utility functions, saving and asset accumulation will be unchanged by a consumption tax (Beach, Boadway and Bruce, 1988). The rise in the cost of consumption with given lifetime wealth causes consumption to fall in the same proportion in each period. Since the consumption tax also applies proportionately in each period, consumption expenditures inclusive of the tax remain unchanged; so the same amount of saving is required in order to finance them (since earnings are unchanged). The fact that a consumption tax leaves saving unchanged makes tax substitutions involving consumption taxes easy to analyse.

**Capital income taxation.** A capital income tax is equivalent to a reduction in the after-tax interest rate. In the short run, it will have an ambiguous effect on saving depending on the relative sizes of the substitution, income and human wealth effects. The magnitude of the human wealth effect will be larger the later in the remaining life cycle are earnings obtained (for example, the younger the person is when the tax is imposed). Over the longer run, the reduction in the after-tax interest rate causes the entire consumption profile to become flatter, so that over the life cycle, the saving needed for life-cycle smoothing will be reduced.

2. **Tax Substitutions with Individual Revenue Neutrality**

From these individual tax effects, it is straightforward to infer the effect of substituting one tax for another. First, consider tax substitutions which raise the same amount of revenue in present value terms from the individual. Since a consumption tax does not affect saving, substituting a wage tax or a capital income tax for a consumption tax will have the same effects as imposing those taxes alone as just discussed. The wage tax substitution will cause individual savings to fall, while the capital tax substitution will have an ambiguous effect on savings in the short run, but will reduce the demand for assets over the life cycle. Substituting a capital income tax for a wage tax will also have an ambiguous effect since both are likely to reduce the demand for assets in the long run.
These ambiguities disappear once we take public sector savings into account. If the present value of tax liabilities of all cohorts is to be kept constant, the time pattern of government tax receipts will change; it will need to borrow or lend. For example, when a wage tax is substituted for a consumption tax, the entire amount of the fall in private savings comes about from the change in timing of tax liabilities since the consumption stream remains unchanged. In this case, the fall in private savings will be exactly offset by an increase in public savings as a result of the fact that tax revenues are received earlier. Thus, total savings will remain unchanged, along with the path of consumption and capital accumulation for the economy as a whole and the welfare of all cohorts. In fact, when tax changes are revenue-neutral for all households, all changes in private savings which originate in changes in the timing of tax liabilities will be offset by changes in public savings in the opposite direction.

Only those sources of change resulting from the relative price effect of a change in the after-tax interest rate will cause saving to change. Thus, in the case of the replacement of a capital income tax for a wage tax, the substitution effect which causes saving to fall will remain. The fact that taxes are collected earlier in the life cycle under the wage tax will be of no consequence for aggregate saving; the increase in private saving this generates will be offset by a decrease in public saving. Moreover, given the changes in public saving, the substitution of a tax on capital income for a wage tax will have exactly the same aggregate effect as the substitution of the same tax for a consumption tax. That is, saving will fall in both cases by the substitution effect. The tax on capital income will represent a distortion in the capital markets; all individuals will be worse off as a result of substituting a capital income tax for either a wage tax or a consumption tax.

3. Tax Substitutions with Aggregate Revenue Neutrality in Each Period

The above exercises require the government to collect the same amount of revenue in present value terms from each cohort. They should be viewed as illustrative only since it is difficult to imagine the government being able to institute tax changes which are revenue-neutral for every cohort, even if public saving is allowed. It is more likely to be the case that some cohorts will gain and others lose. (Indeed, this is often the objective of tax reforms.) Suppose instead that we consider the same three sorts of tax substitutions when the same amount of total tax revenue is generated in each time period so total public saving remains unchanged. Again, capital markets are assumed to be perfect and households leave no bequests.

Suppose the economy is in steady-state equilibrium with a consumption tax used to finance an exogenous stream of government expenditures. The government then switches to a wage tax which generates the same stream of revenues per period. In the period in which the tax substitution occurs, members
of the older cohort (the retired plus the workers part way through their working life) obtain a windfall gain since they no longer have to pay taxes on their consumption and escape bearing the full burden of the wage tax. Younger workers and future generations face a wage tax rather than a consumption tax. They will reduce their savings since their tax liabilities occur earlier in the life cycle. Thus, the total stock of capital in the economy will fall. In the new wage-tax steady state, individuals will be worse off if $r > g$, as is normally the case. The tax substitution will be equivalent to a pure intergenerational redistribution, with older cohorts gaining at the expense of all future cohorts. It is purely redistributitional in the sense that there is no efficiency gain; the economy simply moves from one point to another on its intertemporal utility possibility frontier, though the process of transition can take several periods.

Next suppose the consumption tax is replaced by a capital income tax. If lifetime utility is the discounted sum of per-period utilities, the reduction in consumption tax will have no effect on saving. The capital income tax will cause the consumption profile to flatten out, so asset accumulation over the life cycle will fall, even though saving early in the life cycle could rise or fall. Thus, the capital stock should decline with the tax change. Long-run welfare will unambiguously decline if $r > g$, both due to the distortion imposed by the capital income tax and due to the further reduction in the capital stock. During the transition, however, older cohorts are made better off. For them, consumption tax liabilities will be higher later in life than capital income tax liabilities.

Finally, suppose a capital income tax is substituted for a wage tax. As with the previous case, this one will have a relative price effect and a timing effect. The relative price effect will be the same as for the previous case. The timing effect will be stronger in favour of increasing saving since the wage tax is collected earlier in the life cycle than was the case for the consumption tax. Thus, in the long run, aggregate saving could rise or fall, and households could become better or worse off. However, in the short run, since taxes are collected earlier in the life cycle under the wage tax, older cohorts are made worse off during the transition to the new long-run equilibrium.

These results rely on the assumptions we have made about individual behaviour and capital markets. Suppose that individuals are altruistic towards their heirs and that bequests fully reflect that, so Ricardian equivalence applies. Any purely redistributive transfer between cohorts resulting from tax reform will be undone by a change in bequests. In the case of substituting a wage tax for a consumption tax, which is exactly equivalent to a lump-sum transfer from younger to older cohorts, bequests will rise by the full amount of the transfer with no real changes in the economy. For the other two tax substitutions, the intergenerational transfer component will be offset by changes in bequests, leaving only the relative price effect. Since the relative price effect amounts to removing an inefficiency in the economy, utilities of all cohorts along the growth path will be reduced when the capital income tax is introduced.
If capital markets are imperfect, the effects will differ as well. For example, if individuals are liquidity-constrained, substituting a wage tax for a consumption tax cannot be expected to reduce saving for the youngest cohorts. Since they cannot dissave, the additional tax liability they face must be met by a fall in consumption rather than an increase in borrowing. Thus, the adverse effects of the tax substitution on saving will be lessened. The same applies for the substitution of a wage tax for the capital income tax. Similarly, if lifetimes are uncertain and annuity markets are imperfect, the substitution of a wage tax for a consumption tax will implicitly affect the amount of insurance implemented through the tax system. The consumption tax system increases the need for precautionary saving relative to the wage tax because consumption tax liabilities are incurred as long as the household is alive. Thus, the substitution of a wage tax will reduce saving more than it otherwise would.

4. Summary

One way to summarise this section is to review the ways in which tax policy could be used to increase aggregate saving. These will depend upon the appropriate model of the economy. For simplicity, consider only tax reforms that are revenue-neutral in each period, ruling out government debt. In the absence of bequests and liquidity constraints, saving is increased by the substitution of a consumption tax for a wage tax or by any other tax reforms that redistribute from older to younger generations (that is, which collect taxes later in the life cycle). A tax reform that changes the net rate of return on saving may increase or decrease the rate of saving depending on the relative magnitudes of the income and substitution effects. If annuity markets are incomplete, so that unintentional bequests exist, a tax reform that substitutes tax liabilities later in life for those earlier will reduce the effect on saving since this will be equivalent to providing annuities to the household (since the payment of the tax later in life is contingent on being alive). As well, if the household is liquidity-constrained early in life, the effect of the reform on saving will be further blunted since the household will want to consume more of the additional income obtained early in life.

By comparison, in an economy in which Ricardian equivalence applies, pure intergenerational transfers are completely undone by bequest changes and leave aggregate saving unchanged. On the other hand, increases in the after-tax rate of return on savings should unambiguously increase saving. Since the income effects of tax reforms are offset by bequests, all that is left is the substitution effect which is unambiguous in direction. The effects arising out of liquidity constraints and incomplete annuity markets should also apply here. However, in a world of Ricardian equivalence, these market imperfections are much less likely to occur. Transfers within the family ought to be able to overcome liquidity constraints, and family dynasties should be able to self-insure against the uncertainties of length of life.
IV. EMPIRICAL ANALYSIS OF SAVINGS

As we have seen, a number of different theories of consumption and saving behaviour have been put forward. In order to deal with practical problems of public policy, one would like to know what theory has the best ability to explain observed behaviour; beyond this, it would be desirable to have a good predictive model that would provide quantitative estimates of the behavioural responses of households to different policy initiatives. Broadly speaking, empirical research on savings can help us to test competing theories and to estimate the values of crucial parameters; simulation analyses, discussed in the next section, can be used to explore, in a more predictive spirit, the implications of different policies in models of the economy that build on the tests and estimates derived from empirical research.

Much of the empirical research on savings is not directed specifically to the analysis of the effects of taxation; rather, it is concerned with understanding the basic nature of household decision-making. As is clear from our previous discussion, however, the effects of taxes can depend sensitively on the nature of savings behaviour, so this empirical research carries important implications for the analysis of tax policy. A detailed review of the vast empirical literature on saving is beyond the scope of this paper, but it is useful to review the main lines of empirical research from the perspective of potential applications to tax analysis.

1. Studies of Aggregate Savings Behaviour

For the purposes of tax analysis, the effect of interest rates on consumption and savings is of particular importance, but early empirical research on aggregate consumption functions had not found interest rates to be important explanatory variables. Indeed, according to ‘Denison’s Law’, the proportion of national income devoted to saving is more or less constant over long periods. If this were true, the interest elasticity of the savings rate would be approximately zero, since interest rates are observed to fluctuate substantially over time. Boskin (1978), however, casts doubt on the validity of Denison’s Law, arguing that the stability of the gross savings rate masks considerable variation in the rate of savings net of depreciation. Boskin regresses the aggregate net savings rate on an estimate of the real after-tax rate of return on saving and concludes that the interest elasticity of savings is in the range of 0.2–0.4. Subsequent research has yielded mixed results; Howrey and Hymans (1978) conclude that interest rates have little effect on savings, while Blinder and Deaton (1985) find that nominal interest rates tend to depress consumption but real interest rates do not, a result not readily reconciled with theoretical models that suggest that consumers make decisions based on real trade-offs.

Instead of estimating traditional aggregate consumption functions, a number of studies follow Hall (1978) in attempting to estimate a model built on the
assumption that aggregate consumption is the outcome of intertemporal optimising behaviour by a representative household. Here the objective is to recover underlying structural parameters of the preference structure of the representative agent which could then be used, among other things, to estimate the interest elasticity of savings. In one such study, Mankiw, Rotemberg and Summers (1985) conclude that the data reject the testable implications of the underlying hypothesis of intertemporal optimising behaviour by a representative agent. Hall (1988) finds that consumers are not very willing to substitute between present and future consumption, implying a low interest-elasticity of saving.

Overall, it would be fair to say that empirical analysis of aggregate consumption data has been inconclusive in its findings about the effects of interest rates on savings. In part, at least, the divergent findings of the literature on aggregate consumption and savings can be traced to differences in the data used, as different investigators choose somewhat different concepts of savings, income and other key variables. Another perennial source of controversy concerns the choice of variables to include in an aggregate consumption function. In addition to tax policy per se, it has been suggested in various contexts (Bosworth, Burtless and Sabelhaus, 1991) that aggregate savings may be sensitive to fluctuations in stock market prices and real estate values, changes in capital market institutions (new types of borrowing arrangements, for example), increases in government deficits or demographic shifts (for example, changes in age structure, increases in the number of single-parent families). Carroll and Summers (1987) identify government borrowing and differences in the tax treatment of capital income (especially through tax-sheltered savings arrangements) and interest expense as probable contributors to recent divergences in savings rates in the US and Canada. A failure to incorporate these sorts of factors in an estimated aggregate consumption function not only harms the overall explanatory power of a model but may also result in misleading estimates of the effects of other variables such as interest rates and income. The correct specification of a macro-level consumption function to incorporate detailed institutional factors is usually far from clear, however. Undoubtedly this helps to explain the increased interest in the use of micro-level data in consumption analysis, as discussed further below.

As explained in earlier sections, tax policy affects not only the net rate of return on saving but the distribution of tax burdens over time. Many empirical studies use macro-level data to assess the impact on aggregate consumption of effects of changes in tax timing resulting from public pension or debt policy. In one well-known analysis of the effect of the US social security system on aggregate consumption, Feldstein (1974) calculates (net) social security wealth, i.e. the present value of benefits less taxes, aggregated across all age cohorts in a given year. Aggregate social security wealth in each year then becomes an explanatory variable in a time-series aggregate consumption regression equation.
Feldstein estimates that social security wealth has a large positive effect on aggregate consumption, reducing savings rates and thus depressing wealth accumulation. (Beach, Boadway and Gibbons (1984), however, distinguish between impact and long-run effects of social security policy and note that the long-run effects are likely to be quantitatively much smaller than the Feldstein impact estimates; for instance, social security is estimated to depress the stock of wealth by only about 5 per cent instead of the 30 per cent that Feldstein estimates.)

Exercises of this sort raise not only many aggregation and other econometric issues, but serious measurement problems. For example, the calculation of social security wealth requires a determination of households’ expectations of future taxes and benefits, which are not directly observable (Leimer and Lesnoy, 1982). A second major measurement problem concerns the effects of other policies that redistribute income across generations, such as explicit government borrowing, public investment in both non-human and human capital, changes in the structure of taxation, and a host of other policies (Boadway and Wildasin, 1993). Recent efforts to develop more comprehensive measures of intergenerational flows through the public sector (Kotlikoff, 1992) may, however, prove useful in overcoming some of these problems for aggregate consumption function studies.

2. Micro-Level Analysis of Savings

The tax systems of many countries provide preferential treatment for retirement savings, for owner-occupied housing and for human capital investment. The taxation of households and firms is often not very well integrated and the tax treatment of gifts and bequests can be quite complex. Furthermore, different forms of wealth accumulation have differing degrees of liquidity, ranging from highly liquid cash-balance savings in financial institutions to rather illiquid capital tied up in equity on owner-occupied housing. The increasing availability of micro-data sets has encouraged researchers to attempt to take these complexities into account in the study of savings behaviour, as described in the following paragraphs.

Retirement Savings

Many tax systems shelter pension and other retirement savings from taxation. Typically, contributions to pension plans, both by companies and by employees, are tax-deductible, and the returns to these contributions (whether they take the form of dividends, interest or capital gains) can accumulate free of tax. When benefits are paid out to retirees, however, they are included in the taxable income of recipients and subject to personal tax. Since the returns to pension savings are not taxed on accrual, savings in pension programmes earn the before-tax rate of return, and therefore enjoy significant tax preferences relative to other forms of capital accumulation. A number of other special retirement savings vehicles in
various countries provide individual-level tax-sheltered savings opportunities; these include Registered Retirement Savings Plans in Canada, Individual Retirement Accounts (IRAs) and, more recently, 401(k) plans in the US, and personal pensions in the UK.

Retirement saving is empirically very important. In the US, it accounts for more than half of national savings, but its composition has varied widely. Over the 1980s, pension contributions fell from being most of retirement savings to about half of the total; of pension savings, the proportion in defined-benefit plans fell relative to defined-contribution plans. Higher rates of return on pension fund assets (including appreciation of equity holdings) during the 1980s, making it easier for firms to maintain adequate funding for defined-benefit plans, may have contributed to this relative decline of defined-benefit contributions (Bernheim and Shoven, 1988); changes in the regulations governing defined-benefit pensions, such as rules concerning vesting of benefits, higher premiums to support government insurance of pension fund solvency and increased taxes on employers who attempt to withdraw ‘excess’ pension fund contributions may be another factor (Papke, Petersen and Poterba, 1993). IRA contributions grew to about one-fourth of retirement savings before dropping off toward the end of the decade; 401(k) contributions rose from a trivial amount in 1980 to about one-third of retirement savings by the end of the 1980s (Poterba, Venti and Wise, 1993). Retirement saving has also varied in magnitude and composition in the UK; for instance, the take-up of defined-contribution and personal pensions has increased markedly since the 1988 Social Security Act made these available as contracting-out options from the State Earnings-Related Pension Scheme (SERPS) (Banks and Blundell, 1992; Dilnot, Disney, Johnson and Whitehouse, 1994).

A number of studies have investigated the relationship between retirement savings and other forms of wealth accumulation. Poterba (1987) estimates that an increase in corporate saving is accompanied by a reduction in household saving of about 50–75 per cent. Households seem to ‘see through the corporate veil’ to some degree but do not regard corporate saving as a perfect substitute for personal saving. IRAs would not necessarily be expected a priori to have very significant effects on aggregate savings (Gravelle, 1991), since IRA contribution limits are relatively small. Further, IRAs might substitute for employer-provided pension contributions, or otherwise change the form of household saving without changing its level. However, in a series of studies based on surveys of individual households, Venti and Wise (1992 and references therein) present evidence that IRA participation is higher for higher-income households and that households that make IRA contributions have not reduced their accumulation of other forms of financial assets. They find that households covered by pension plans do not appear to have levels of IRA contributions different from those that are not covered. Households appear to regard IRA savings as a distinct type of asset that is only imperfectly substitutable for other savings. Venti and Wise conclude that
most IRA savings represent a net increase in total household savings. Evidence on the effect of personal pensions on savings in the UK is not yet available (Banks and Blundell, n.d.), but participation rates are high; as more UK data accumulate, it will be of interest to compare results with US research on IRAs. A challenge for future work on retirement savings is to develop models that can integrate the crucial but complex institutional details of tax-sheltered retirement savings vehicles within an overall framework for the analysis of total household savings.

Savings and Housing

The accumulation of housing equity over time due to mortgage repayment and appreciation in house value increases a household’s net worth and is thus a component of household saving, comprehensively defined. Indeed, housing equity can be a major component of personal saving for most households, who often retire with very little saving in the form of financial assets. The illiquidity of housing equity could significantly affect life-cycle savings behaviour. Pre-retirement households whose housing increases in value would presumably wish to consume more but might find it difficult to translate appreciation in the price of a house into liquid resources that can be used for (non-housing) consumption. Older households usually experience large reductions in cash-flow income at retirement and might wish to liquidate some of their housing wealth, but doing so typically involves various transactions costs, so that their (non-housing) consumption might be cash-constrained.

Skinner (1989) studies this issue by regressing the consumption of a sample of home-owners on a measure of housing wealth. It appears that increases in housing value have little effect on consumption, which Skinner interprets as evidence of altruistically-motivated bequests to subsequent generations. Manchester and Poterba (1989) find that households that take out second mortgages have lower overall net worth positions, which could mean that second mortgages permit home-owners to liquidate housing equity to finance consumption. Alternatively, it could be that some households experience unexpected reductions in cash flow or increases in desired consumption (for example, due to unemployment or medical emergencies) and that these households disproportionately seek out second mortgages. A negative correlation between second-mortgage borrowing and household net worth could then simply reflect the impact of a sudden need to increase consumption rather than any independent effect of the availability of second mortgages on consumption and savings. The link between savings and housing decisions for the elderly is discussed further below.
Savings Behaviour of the Elderly

Since the life-cycle savings model emphasises the importance of savings for retirement, the behaviour of elderly households is of particular interest. If most saving is motivated by a desire to provide for one’s old-age consumption, and if capital and insurance markets function perfectly, then one would expect older households to engage in substantial dissaving. In the absence of bequest motives for saving, the ideal life-cycle consumption path is one that totally exhausts household wealth at the time of death. Hence, the life-cycle theory would lead one to expect a negative wealth–age relationship in old age. As explained previously, however, this prediction must be qualified if insurance and capital markets are imperfect; uninsured risks concerning health and the age of death may give rise to precautionary motives for preserving wealth in old age, and to ‘unintended’ or ‘accidental’ bequests.

A substantial amount of empirical research has focused on understanding patterns of saving and dissaving in old age and on the purchase of annuities and insurance by the elderly. Conflicting evidence has emerged. Bernheim (1987), for instance, finds that ‘bequeathable’ wealth, i.e. non-annuitised wealth, including housing wealth as well as financial and other assets, tends to decrease after retirement, as suggested by the life-cycle theory. A more comprehensive measure of wealth that includes both private pension benefits and social security benefits, however, seems to remain approximately constant over time. Hurd (1992) concludes that households do generally draw down their wealth at rates consistent with the pure life-cycle theory; he also observes that parents and non-parents tend to dissave at approximately equal rates, casting doubt on the importance of bequest motives for wealth accumulation. Friedman and Warshawsky (1989) analyse retirement annuity contracts and find ‘load factors’ (mark-ups above actuarially fair prices) of 20–40 per cent. Though substantial, these are similar to the load factors on other commonly purchased types of insurance, suggesting that the low level of wealth annuitisation reflects a desire to preserve wealth for bequest purposes rather than a serious market imperfection.

Much of the wealth of the elderly takes the form of somewhat illiquid housing equity. Venti and Wise (1990 and references therein) examine the housing wealth of older households, comparing those who stay in their existing dwellings with those who move. If the consumption of older households is generally liquidity-constrained, movers would presumably choose less-costly housing, freeing up liquid wealth with which to sustain current consumption. Instead, it appears that aged movers acquire new housing that is just about as costly as their previous dwellings, so that housing equity remains relatively constant. The underlying motivation for the preservation of housing wealth is not entirely clear, but to the extent that it occurs, it casts doubt on the hypothesis that consumption of the elderly is liquidity-constrained. Such behaviour also suggests
that the relatively limited use of reverse-annuity mortgages may be attributable more to a simple lack of demand rather than to any institutional barriers on the supply side.

Bernheim (1991) observes that many older households purchase term life insurance which, in terms of its effects on the household’s budget constraint, is basically equivalent to the sale of an annuity, with the proceeds of the sale accruing to the insurance policy beneficiaries. Such insurance increases the level of bequests at the expense of current own consumption, suggesting that the elderly do value planned bequests and that pure life-cycle consumption motives cannot, by themselves, explain observed savings behaviour.

Heterogeneous Savings Behaviour and Liquidity Constraints

There is no compelling a priori reason why all households should necessarily follow the same behavioural rules in their consumption and savings decisions. Several authors (Diamond, 1977; Diamond and Hausman, 1984; King and Dicks-Mireaux, 1982) use household-level data to examine the pattern of wealth-holding by age and permanent income level. Households with low lifetime incomes frequently save so little that they cannot sustain a level of post-retirement consumption close to that achieved during the working lifetime, indicating that a significant fraction of the population acts more myopically than the life-cycle model of consumption smoothing would suggest. On the other hand, a relatively small but important fraction of households save a great deal, as evidenced by the fact that mean levels of wealth, by component and in total, are frequently far higher than median levels. The behaviour of these households also cannot easily be explained in terms of consumption smoothing over the life cycle, but for the opposite reason: the savings of these households appear to be based on much longer-term objectives, involving perhaps the transfer of wealth to their children or to others.

Further evidence of the heterogeneity of household consumption behaviour, if not of underlying savings motives, comes from studies that suggest that a significant proportion of consumers are credit-constrained (Hall and Mishkin, 1982; Hayashi, 1985; Mariger, 1986; Zeldes, 1989b; Wilcox, 1989; Japelli, 1990). The data and empirical tests differ among these studies, but indicate variably that consumption may be more sensitive to fluctuations in current income than is consistent with intertemporal optimising behaviour, that younger households with small amounts of liquid assets save less and have more difficulty obtaining credit than other consumers, and that the behaviour of households with significant asset holdings does conform to the restrictions implied by intertemporal optimising whereas this is not true for households with few assets. An international comparison by Japelli and Pagano (1989) suggests that the proportion of income accruing to credit-constrained households may be as low as 12 per cent in Sweden and as high as 50–60 per cent in Spain, Italy and
Greece, with estimates of about 21 per cent for the US and 40 per cent for the UK. Evidently, liquidity constraints do not interfere with the consumption of all households, but they are not trivial in extent either.

Intergenerational Transfers

There is considerable controversy in the literature about the magnitude of intergenerational transfers. Data on bequests — the amount of bequests, their disposition and the attributes of households that make them — are quite limited. A few authors (Menchik, 1988; Tomes, 1988) have analysed court records pertaining to individual estates, but such data are not easily obtained and often lack important information on other household attributes Aggregate-level data on the magnitude of bequests are generally unavailable. Furthermore, bequests are only one form of intra-family transfers. Parental nurturing of children begins at infancy and involves the expenditure of both parental time and pecuniary resources at least until young adulthood. Parents may provide financial support for their children’s higher education or for purchases of housing or durable goods, while older children may support their parents in times of financial need and may spend significant amounts of time caring for their parents. Few data sources are available that measure these sorts of intergenerational transfers.

In the absence of direct observations on the magnitude of intergenerational transfers, Kotlikoff and Summers (1981) attempt to infer how important they may be by estimating the extent of aggregate savings that might be reasonably attributed to pure life-cycle motives. The residual — all savings other than for life-cycle purposes — they attribute to intergenerational transfers, whether accidental, altruistic or strategic in nature. They estimate that as much as 50–80 per cent of wealth accumulation in the US arises from intergenerational transfers, though this estimate is sensitive to the resolution of some subtle conceptual questions (Modigliani, 1988; Kotlikoff, 1988).

What motives underlie the intergenerational transfers that do occur? Altonji, Hayashi and Kotlikoff (1992) examine the consumption response of related households to changes in the resources of any one household. If parents and their adult children are altruistically linked, the resources of each should be effectively merged in an extended-family budget constraint, such that the consumption of each household unit depends on the resources of the extended family but not on the separate contributions of each to the family’s resources. The evidence indicates that most households are not strongly altruistically linked, in the sense that own resources do influence own consumption, but that the households are still somewhat linked in the sense that own consumption does depend on the resources of other related households. Cox (1990) and Cox and Japelli (1990) find that parents and children are linked through _inter vivos_ transfers that offset liquidity constraints facing the latter. Other studies
Taxation and Savings

(Bernheim, Shleifer and Summers, 1985) favour the strategic bequest motive. Additional work is needed to sort out these competing hypotheses.

In summary, empirical research does indicate that tax policy can significantly affect savings decisions; this much at least is apparent from the studies of retirement savings through tax-sheltered accounts. There can be little doubt that the extent of pension fund savings is attributable in significant part to the fact that pensions are tax-sheltered, but available evidence does not strongly support the notion that total private (personal and corporate) saving can be treated as a meaningful aggregate. Some but not all households seem to be subject to liquidity constraints; some but not all households may be linked through intergenerational transfers; some but not all savings behaviour seems consistent with a precautionary savings motive; annuities markets do exist but are not heavily utilised, either because the annuities are too expensive or because households wish to leave bequests; and some households appear to make very little provision at all for retirement. The availability of new data, especially new micro-data, has greatly facilitated empirical research, but data limitations continue to plague researchers and to spawn conflicting interpretations. In the face of all this uncertainty, convincing empirically-based quantitative analyses of the effects of tax policy on savings are difficult. Simulation methods provide one means by which the effects of policy can be assessed quantitatively, conditional on the assumed structure of a simulation model. We turn to simulation analysis next.

V. SIMULATION ANALYSIS OF TAX POLICY

Economic models of the savings behaviour of individual households can become rather complex. Even simple life-cycle models with fixed labour supply and perfect capital markets involve analysis of optimal consumer choice over as many commodities as the number of periods in the life cycle. Adding such plausible complications as bequests and other intergenerational transfers, human capital formation, liquidity constraints or uncertainty about earnings or mortality makes it very difficult to obtain much detailed insight from the theoretical analysis of general models. Understanding the effect of taxation on individual savings behaviour alone does not reveal the effect of taxes on relative factor supplies, factor prices and the dynamic evolution of the economy — the effects which are often of greatest interest in policy evaluation. To examine these effects, it is necessary to analyse tax policies in a general equilibrium framework. The development of tractable dynamic general equilibrium models for tax analysis is a most challenging undertaking, however.

In the face of these complexities, researchers have come to rely increasingly on simulation methods. Simulation analysis is now an important adjunct to empirical research, enabling investigators to ‘test’ whether particular hypotheses about savings behaviour are consistent with empirical regularities, to gauge the
likely empirical importance of certain types of savings behaviour or to guide econometric analysis. Simulation models make it possible to study more complex and realistic policy questions, and to gain more insight into the relative practical importance of key parameters, than would otherwise be the case.

1. Precautionary Savings

As an example of the use of simulation models of savings behaviour, consider the case of precautionary savings. Previous sections discussed the effects on savings of imperfections in the market for annuities and of liquidity constraints. Still a third type of market imperfection arises from the difficulties that consumers may have in insuring themselves against the risk of uncertain earnings. Earnings uncertainty can arise from the risk of unemployment, from industry or occupational shocks to wages, or from individual health and other risks. These earnings risks are not readily insurable, since such insurance creates obvious disincentives to labour supply. Thus, households must ‘self-insure’ using precautionary savings.

A number of authors have developed models of ‘precautionary’ savings in which households accumulate wealth as a form of insurance against earnings risk. In general, these models are technically complex, since they involve optimisation in a multi-period uncertainty setting. However, they can usefully be simulated, using data on observed earnings levels and risk and with plausible preference specifications. One use of such simulations is to determine how much saving households might undertake because of this precautionary motive. Zeldes (1989a) estimates that precautionary savings could account for roughly 10–20 per cent of observed savings in the US. (Guiso, Japelli and Terlizzese (1992) estimate that earnings risk in Italy is such that the precautionary motive can only explain a very small fraction of savings there, however.)

Simulation methods can also uncover qualitative properties of precautionary savings models that may be difficult to establish using general theoretical analysis. For example, models with precautionary savings generate excess sensitivity of consumption to current income and low rates of wealth decumulation on the part of the elderly, both empirical observations that have been interpreted by some as inconsistent with standard life-cycle/permanent income hypothesis models, and as possibly providing evidence of credit rationing or of a bequest motive for saving. If these simulations do not constitute formal empirical tests of the precautionary savings motive, they at least provide a convincing conjecture for what may lie behind some of the apparent anomalies in the empirical studies.

2. Dynamic Tax Analysis

The work of Auerbach, Kotlikoff and Skinner (1983) and Auerbach and Kotlikoff (1987) illustrates well the potential policy applications of simulation
models. These authors develop an overlapping-generations model in which each household engages in life-cycle saving over a deterministic lifetime, choosing a consumption stream and labour supply to maximise lifetime utility subject to a lifetime budget constraint, taking factor prices (i.e. wage and interest rates) as given. Given a path of factor prices over the life cycle, it is possible to calculate the utility-maximising choices of each household. Making appropriate assumptions about the demographic trends in the economy, one can aggregate across households of different ages to determine aggregate consumption, savings and labour supply conditional on factor prices; the supply of savings and the supply of labour determine the economy’s capital/labour ratio. But the factor prices themselves depend on this ratio through its effect on the marginal productivity of each factor; the exact relationship depends on the production technology. A dynamic equilibrium in this model is a path of factor prices over time such that the factor supplies forthcoming, conditional on these prices, give rise to that factor-price path under marginal-productivity factor pricing. Of course, this means that factor prices and factor quantities are simultaneously determined. Since households are assumed to be forward-looking intertemporal optimisers, it also means that today’s factor supplies, and thus today’s factor prices, depend on tomorrow’s factor prices. In equilibrium, the entire dynamic path of price and quantity decisions by all households must be consistent. One main task in a simulation analysis is to devise computational procedures that find these equilibrium prices and quantities.

A special case of a dynamic equilibrium in this model is the steady state, in which factor prices and factor proportions are unchanging over time. Steady states are typically used as starting-points for policy analysis. For example, the economy might be in a steady-state equilibrium with a proportional or progressive income tax. The tax system drives a wedge between gross and net factor prices, the former still being determined by marginal-productivity factor pricing and the latter entering into household lifetime budget constraints. Since lifetime consumption/saving and labour/leisure paths are calculated explicitly, it is straightforward to determine the value of lifetime utility in the initial steady state, under the initially given tax policy. Having established a bench-mark equilibrium, one can then change the tax structure, for example by imposing a switch from income to consumption or wage taxation.

A change in tax policy causes the equilibrium of the economy to change, eventually leading to a new steady state. Early simulation studies (e.g. Summers, 1981) compared the steady-state equilibrium under one tax structure with that under another. More recent work carries the analysis further by examining the transition path from an initial steady state to a new steady state, making it possible to trace out the effects of a change in tax policy over time. This turns out to be quite important for policy evaluation, as becomes apparent when analysing a switch from a comprehensive income tax to either a consumption tax or a wage tax. Both the consumption and the wage tax eliminate the income tax
distortion of the consumption/savings decision, and both result in a significantly higher steady-state capital/labour ratio. The initial impact on households of different generations that results from a switch to one or the other of these taxes is quite different, however. A switch to a consumption tax effectively captures some of the previously accumulated wealth of those who are old at the time of the tax change, since their wealth buys fewer consumption goods than would have been the case under the income tax. The taxation of the stock of accumulated wealth is lump-sum in nature (assuming that the tax reform was not anticipated) and entails no efficiency loss. On the other hand, a switch to a tax on earnings eases the tax burden on the elderly, who would have had to pay some taxes on their interest income under the income tax, while the young, for whom earnings are a more important source of income, have to pay higher taxes.

Simulation analysis not only reveals these qualitative effects of tax policy, it allows one to quantify various effects. For instance, in Auerbach, Kotlikoff and Skinner (1983), those who are 18 or more years into the life cycle at the time of the policy change are harmed by a switch to a consumption tax, while all younger households and future generations benefit. The converse is true when switching to an earnings tax. Steady-state welfare is higher under the consumption tax than under the income tax, but a move from an income tax to a tax on wage income alone lowers steady-state welfare; expressed in terms of welfare-equivalent changes in lifetime wealth, the consumption tax raises steady-state welfare by 6 per cent, while the wage income tax reduces it by 4 per cent. These results illustrate clearly the differences between the short- and long-run effects of tax substitutions, and the way that these policy issues raise the issue of intergenerational income distribution. Moreover, as is common in this literature, these numbers are relatively large, suggesting that issues surrounding the effect of taxation on savings can be quantitatively important.

To separate the efficiency and distributional effects of tax policy, Auerbach et al. introduce the fiction of a lump-sum intergenerational redistribution mechanism through which some generations can compensate others. If the cohorts alive at the time of a switch from an income to a consumption tax are compensated so that their welfare is unchanged, subsequent generations can still enjoy an increase in welfare equivalent to that which would result from an increase in lifetime wealth of 1.73 per cent. Evidently, the ‘long-run’ welfare gains from consumption taxation are more than adequate to compensate those who suffer during the transition. In moving from an income to an earnings tax, there are short-run gains but long-run losses. In this case, if the welfare of those alive at the time of the tax change is held constant, the welfare of succeeding generations falls, in wealth-equivalent terms, by 2.33 per cent. Those who gain from this policy do not gain enough to compensate the losers.

Simulation analysis of dynamic fiscal policy has been extended to incorporate multiple capital assets. Gahvari (1984) considers housing as a second type of capital good, and finds that a revenue-neutral tax reform that imposed a tax on
housing would increase the capital intensity in the industrial sector. If the pre-tax rate of return on industrial capital exceeds the growth rate of the economy, this will enhance steady-state welfare. Lord (1989) analyses the revenue-neutral substitution of consumption taxation for wage taxation in an economy with both human and non-human capital. As before, wealth accumulation increases, but the magnitude of the effect is dampened, and much of the increase goes to human capital rather than physical capital.

3. Policy Analysis with Imperfect Markets

When capital markets are imperfect, the welfare analysis of taxation and savings is likely to change substantially. However, there are few routes open to the analyst who wishes to understand the implications of tax policy in such an environment. Market imperfections interact with tax policy in complex ways that make purely theoretical analysis extremely difficult. A number of studies have shown how simulation models can offer a tractable framework for the study of tax policy while incorporating capital market imperfections such as liquidity constraints or incomplete annuities markets.

Abel (1985) studies the introduction of a public pension system in an overlapping-generations model where households have uncertain lifetimes and where private annuities markets do not exist, so that accidental bequests occur. The introduction of a fully funded (actuarially fair) social security system reduces risk and thus the need for private precautionary saving, which is welfare-improving. If the capital stock is below the Golden Rule level, however, steady-state consumption levels fall, so the effect on social welfare is ambiguous. The transitional impact of a fully funded scheme is to benefit the first generations (since accidental bequests are still high), but subsequent generations could be made better or worse off.

Hubbard and Judd (1987) investigate the simultaneous impact of both liquidity constraints and an absence of annuities markets. A funded public pension programme in this model exacerbates liquidity constraints that limit the consumption of the young, even though it provides valuable insurance; liquidity constraints thus reduce any welfare gains from such a policy, or may result in welfare losses. This problem is even greater if the scheme is unfunded; one way around it is to make contributions age-dependent, with higher contributions later in the working life when the liquidity constraint is not binding.

Altig and Davis (1992) have developed simulation models incorporating both capital market imperfections and altruistically-motivated intergenerational transfers. Altig and Davis consider various forms of liquidity constraints that might face young households, such as a complete inability to borrow or interest rates on loans that exceed the interest obtainable on savings, for instance because of asymmetric tax treatment of interest income and interest expense. They
emphasise that young workers are likely to face rising earnings profiles and that they would therefore wish to borrow in order to smooth consumption over time; altruistic parents might therefore wish to make *inter vivos* transfers in order to help their children overcome the losses that they suffer from capital market imperfections. (Altruistic parents can get a larger psychic return from a given wealth transfer to their children if they provide the transfer at a time when the marginal utility of consumption of the children is unusually high, which is the case when they are credit-constrained.)

These models are naturally more complex than those with perfect capital markets and no transfers, because there are many possible equilibrium regimes: depending on the parameters of the model, the young may or may not be liquidity-constrained, and parents may or may not make transfers to them. Altig and Davis examine the effects of changes in implicit tax subsidies to borrowing and taxes on interest income for different possible equilibrium scenarios. If the young do not receive transfers from their parents and are not liquidity-constrained (essentially a traditional life-cycle environment), a reduction in the implicit subsidy to borrowing from 11 per cent to zero raises the steady-state capital/labour ratio by about 10 per cent. If instead the children are liquidity-constrained and receive transfers, the estimated response is closer to 12–14 per cent; if the children are not liquidity-constrained and receive transfers (a situation close to the Barro model), the policy has no effect at all. By comparison, a reduction in the tax rate on interest income from 22 per cent to 11 per cent raises the capital/labour ratio by 4–8 per cent, 0–1 per cent and 16–18 per cent in each of these cases. Notice that the magnitudes differ by type of policy change within each regime, even though the policy change in each case is a tax reform that raises the rate of return on net saving. Notice also that the magnitudes differ by type of regime for each policy change.

As a final example of an interesting policy application of simulation analysis, it is instructive to consider the evaluation of debt and tax policy in a model where households face uninsurable income risk and therefore have an incentive to engage in precautionary saving. Barsky, Mankiw and Zeldes (1986) and others analyse the effect of a reduction in an earnings tax that results in a temporary increase in government borrowing; this tax cut is made up by subsequent increases in taxation. In Barsky et al., the experiment is structured so that the burden of debt repayment falls on the households which enjoyed the initial reduction in taxes. In a certainty environment with life-cycle or dynastic savers, such a policy change would be neutral because it would not change the lifetime budget constraint of any household. In a world with uncertain earnings, however, a certain tax cut now coupled with an increase in future tax rates changes the distribution of income not only across time periods but across states of nature. As a result, changes in the intertemporal structure of taxation are not neutral in their impacts on consumption and savings; the marginal propensity of households to consume out of the higher disposable income resulting from a
current tax cut, coupled with a future tax increase, can well be much closer to the values that characterise myopic Keynesian consumption behaviour rather than to the value of zero predicted by the Barro model. The magnitude of the marginal propensity to consume is calculated for tax changes that extend over one or more periods into the future, and for varying degrees of earnings risk. For example, a current tax cut coupled with a tax increase five periods later may generate additional consumption equal to around 20 per cent of the tax reduction if the level of uncertainty is in a range that corresponds approximately to the variability found in observed earnings distributions. However, the marginal propensity to consume approaches zero as the duration of the fiscal policy diminishes and as the amount of risk in the economy falls.

VI. NORMATIVE ISSUES IN THE ANALYSIS OF TAXATION AND SAVING

In this section, we focus on three normative issues that have arisen in the context of taxing saving. The first looks at what can be learned for the tax treatment of saving from applying the theory of optimal taxation to a dynamic setting. The second concerns the optimal rate of saving and the role of taxation in achieving it. Finally, we look at what has been the main tax policy issue in this context — the choice between income and consumption as a tax base.

1. Optimal Taxation in a Dynamic Setting

Optimal taxation involves the selection of a welfare-maximising tax system in a world in which non-distorting taxes are either not feasible or do not satisfy the desired equity properties. As such, it is an exercise in applied second-best theory. It is now well-known that in a second-best world, where prices are distorted and no longer reflect social marginal values, welfare analysis becomes very complicated. We adopt the strategy of restricting ourselves to simple cases and searching for sufficient conditions for appealing and understandable results. Since the analysis in the literature is fairly technical, we do little more than summarise the main results in words. In most cases, intuitive explanations are not possible. Much of our discussion will be in terms of present and future consumption; the consequences for saving can be inferred readily.

The bench-mark case is that of a single household in a two-period setting with variable labour supply and no bequests. The individual chooses present and future consumption \( (C_1 \text{ and } C_2) \) and labour supply \( (L) \) to maximise lifetime utility subject to a lifetime wealth constraint. Assume initially that labour is supplied in the first period only so that all income is earned then. The government must raise a given amount of lifetime revenue from the household, fixed in present value terms, using taxes on \( C_1, C_2 \text{ and } L \). This encompasses all
Moreover, given that there are only two relative prices in the economy, we can arbitrarily set one of the three taxes equal to zero and rely on the other two to achieve all feasible equal-revenue outcomes. Assuming that the tax on labour is zero, we can focus on taxing \( C_1 \) and \( C_2 \) alone. (In other words, any tax on labour is equivalent to an equal-rate tax on \( C_1 \) and \( C_2 \) so can be incorporated into the taxes on the latter.)

The problem is like a standard three-good optimal commodity tax problem. According to the ‘Ramsey rule’ (Heady, 1993), optimal taxation implies that the demands for all three goods should fall by approximately equal proportions. An alternative characterisation of the three-good optimal tax structure asserts that the tax rate should be higher on the good which is more complementary with (or less substitutable for) leisure (Harberger, 1964). Thus, in the two-period context, if \( C_2 \) is more complementary with leisure than \( C_1 \), it should be taxed at a higher rate; in other words, there should be a tax on capital income, though not likely at the same rate as the tax on labour income. If both are equally complementary with leisure, so that a change in the price of leisure leaves the ratio \( C_1/C_2 \) unchanged, the tax rate should be the same for both goods. Either a proportional tax on consumption or a wage tax would be optimal; there should be no tax on capital income. It might be argued that these conditions are especially unlikely to be satisfied in an intertemporal context in which the variability of leisure occurs in period 1, while consumption is spread over two periods.

These results depend on several strong assumptions concerning household and government behaviour and the nature of the economy. Various extensions to the simple model have been studied in varying degrees of detail, and we summarise the results obtained for each.

**Steady-State Utility Maximisation**

As a first attempt to introduce elements of a dynamic economy, suppose the above household is put into the context of an economy of overlapping generations, and the government is assumed to maximise per capita utility in the steady state. The level of steady-state utility now depends not only on tax distortions, but on how much capital households choose to accumulate relative to the Golden Rule level. If the government is allowed to borrow or lend, variations in the stock of debt can control the stock of capital in the economy. In this case, taxation can be directed solely at revenue-raising and the tax structure will be given by the standard Ramsey rule as above. On the other hand, if the government is unable to deploy debt to achieve the desired steady-state capital stock, its choice of tax structure will have to address two objectives — minimising the dead-weight loss and influencing the amount of capital.

\[ \text{A general consumption tax is a uniform tax on } C_1 \text{ and } C_2. \text{ A wage tax is a tax on } L, \text{ or, equivalently by the budget constraint, a uniform tax on } C_1 \text{ and } C_2. \text{ A capital income tax is a tax on } C_2. \text{ An income tax is a system of taxes on } C_1 \text{ and } C_2 \text{ with that on } C_2 \text{ being higher. And } C_1 \text{ and } C_2 \text{ can be selectively taxed by excise taxes.} \]
accumulated. For example, if private saving generates a level of capital such that \( r > g \), the tax structure should be altered so as to encourage more saving. King (1980) shows that when the government is restricted to using taxes on wage and capital income only, it is optimal to tax capital income more heavily. The reason is that capital income taxes are incurred in the second period so households must save to pay for them. However, if taxes on consumption are used as well, the need for capital income taxes is reduced since (as discussed above) consumption taxes generate more steady-state saving than wage taxes.

This steady-state analysis ignores transitional problems. To get to the optimal steady state by increasing the capital stock (if that is required), the older generations will have initially to forgo some consumption. The time path of adjustment will then depend upon the relative weights given to the lifetime utility of different generations. We return later to the issue of choosing the time path of taxes to maximise an intertemporal social welfare function.

Many Goods

If there are many goods in each period, the analysis becomes quite complicated, especially if the government is constrained to levy the same tax rate on each good independent of the age of the household, as would be the case when taxes are indirect. The effect of this is to introduce as many second-best pricing constraints as their are types of goods. As is well-known from the theory of second best, very few general results can be obtained, and none have been obtained in this context. In the absence of these pricing constraints, the problem would be a standard optimal tax one. However, even in this case, there are conditions under which wage (or proportional consumption) taxation would be optimal. That will be the case if the utility function is such that increases in income cause equal proportionate increases in the demand for all goods and changes in the wage rate do not affect the proportion in which goods are consumed.

Variable Leisure in Both Periods

Alvarez, Burbidge, Ferrall and Palmer (1992) have analysed the taxation of labour and capital income in the two-period case where utility in each period depends upon leisure and a consumption good. They show that, if wage rates and the interest rate are constant over the life cycle, optimal wage tax rates decline with age if the interest rate exceeds the utility discount factor (and vice versa).

---

8 One would expect some further results could be obtained in the case in which lifetime utility was a separable function of the first-period consumption bundle, the second-period consumption bundle and leisure. Then per-period consumption could be treated as a composite commodity, and one could rely on wage and capital income taxes alone.

9 Technically, this is so if utility is homothetic in goods, and if the marginal rate of substitution between all pairs of goods is independent of the level of leisure (i.e. goods are separable from leisure in the utility function).
More generally, the ability to condition wage and consumption taxes on age eliminates the need to tax interest income. On the other hand, if the wage tax (or consumption tax) cannot be conditioned on age, interest income becomes optimal if the interest rate exceeds the utility rate of discount; this is because an interest income tax is like a declining wage tax, though imperfectly so.

Many Consumers per Cohort

The presence of many consumers in each cohort raises issues of redistributive equity if the consumers have differing incomes. In standard models of optimal income taxation, income differences arise from both different levels of ‘ability’ across households and different amounts of labour supply. The government is assumed to be able to observe incomes, but not underlying abilities, and can levy a non-linear income tax as well as indirect taxes at different rates on different commodities. In a static context, Atkinson and Stiglitz (1976) have shown that if the consumption bundle is separable from leisure in the utility function (i.e. if changes in the amount of leisure taken do not affect the marginal rate of substitution among the various pairs of commodities), non-linear wage taxation will suffice; differential commodity taxes are not needed. In the absence of separability, the government should combine the non-linear wage tax with an indirect tax system which imposes a higher tax rate on goods which are complementary with leisure (Edwards, Keen and Tuomala, 1992). By extension, in a dynamic context, if the government is able to tax present and future consumption differentially, the latter would be taxed more heavily if it is more complementary with leisure. Stiglitz (1987) has extended the results of Atkinson and Stiglitz (1976) to an overlapping-generations setting in which households consume a vector of goods in each of two periods and supply labour only in the first. If debt policy can be used to control the level of capital stock, and if the relative wages between skilled and unskilled workers are independent of the capital stock, progressive wage (or lifetime expenditure) taxation is the most efficient redistributive instrument if consumption and leisure are separable in the utility function. However, if relative wages vary with the capital stock or if the government cannot use debt policy, interest taxation or subsidisation becomes a useful supplementary instrument. If it wishes to encourage saving either to increase the capital stock towards its optimum or to induce a fall in the relative wage of the high-ability persons, an interest subsidy is called for, and vice versa.

10 See Heady (1993) for a non-technical survey of this literature.

21 One way to do this would be to combine a tax on income with one on consumption. By varying the rates on the two, any combination of a tax on present and future consumption could be attained. It is interesting to note that many countries’ income tax systems are essentially of this sort, given that they shelter part of capital income.
Altruistic Preferences

If altruism is such that Ricardian equivalence applies, the representative household can be viewed as being infinitely-lived. Chamley (1986) has investigated the optimal tax problem in this context assuming that leisure and consumption are variable in each period and the government must collect a given stream of revenue through a combination of taxes on wages and capital income. Given an initial stock of previously accumulated capital, the optimal tax plan will be one in which the government initially imposes a relatively high tax on capital income, but gradually replaces it by a wage tax; in the long run, only a wage tax is used. The latter result reflects the fact that, given the additive intertemporal utility function, the steady-state optimal tax structure involves only a wage tax. The initial use of capital taxation occurs because at the time the tax problem is solved, some capital has already been accumulated, so it yields a fixed stream of income which can be taxed with no dead-weight loss. However, the government cannot distinguish ‘old’ from ‘new’ capital and must tax the return to all capital identically. As time goes by, the proportion of old to new capital rises, so that the distorting effect of capital income taxation increases, so the taxation of capital income has to be gradually phased out and replaced by wage taxation.

Note that the structure of this problem gives rise to a time-inconsistency. If the government were to redo its optimal tax problem in any future period, given the stock of capital that had then been accumulated, it would renege on its previous plan and impose a high capital income tax once again rather than abiding by the previously announced lower and falling rate.

Time-Consistent Taxation

The notion of the time-inconsistency of capital income taxation in dynamic models was eloquently stated by Fischer (1980). He considers a two-period model in which a single household receives an endowment of wealth, saves some of it for second-period consumption and supplies labour in the second period. The government has to raise revenue in the second period through a combination of labour and capital income taxation. Viewed from the beginning of the first period, the optimal tax plan is one which uses a wage tax and, depending on the utility function, perhaps a capital income tax. However, if the government could re-optimise at the beginning of period 2, after the saving decision had been taken, it would opt to get as much revenue as possible from capital income taxes and as little as possible from wage taxes. The government’s announced optimal tax plan is therefore not credible. A time-consistent policy outcome will be one in which the household behaves according to an expected tax structure that is the same as the one that is actually chosen. It will have the highest capital income tax rate and the lowest wage tax rate consistent with government budget balance. Household saving and welfare will be lower than in the optimal tax solution.
Given the excessive capital income taxation resulting from the time-inconsistency of optimal taxation, it is natural to consider whether alternative policies might be used to mitigate its effects. One obvious policy is to provide savings incentives ‘up front’ knowing that the income from those savings will be taxed too heavily. An alternative is to constrain the government in its use of policy instruments. In an extension of the Fischer model to allow for variable labour supply in both periods, Rogers (1987) has argued that given a particular set of household preferences, consumption taxation might be preferable to wage taxation in a time-consistent equilibrium, even though the opposite is true in the optimal tax equilibrium. Mitigating the consequences of time-inconsistency by constraining the instruments available to the government seems an attractive alternative, though it is not at all clear how such a commitment could be made binding.

2. Is the Saving Rate Too Low?

The importance of the saving rate has been at the forefront of policy debates in recent years, and has been responsible for many policy recommendations, such as instituting full funding of public pensions, retiring the public debt, and replacing the income tax with a consumption tax. Is the saving rate too low? Feldstein (1977) argued for the US that, because \( r > g \), it is. (Recall that \( r \) is the rate of return on capital, while \( g \), the rate of growth in the economy, is the implicit return on intergenerational transfers.) He argued that, in these circumstances, the loss to the current old from increasing saving (reducing intergenerational transfers) is more than offset by the discounted gain to future generations. His argument involved not only the difference between \( r \) and \( g \), but also capital market inefficiency and intergenerational discounting. To evaluate it, we investigate what we can infer about the optimality of the saving rate from the fact that \( r \neq g \). Three main aspects of the optimality of the saving rate can be identified — dynamic inefficiency, capital market inefficiency and intergenerational equity.

Dynamic Inefficiency

An economy is dynamically inefficient if it is possible to increase the welfare of a member of at least one generation without decreasing the welfare of any person in any generation. The circumstances in which the economy is dynamically inefficient are quite limited. Basically, \( g \) must exceed \( r \) now or at some time in the future and continue that way indefinitely (Starrett, 1972). Dynamic

---

12 Specifically, he derived that the present value to society of an increase in saving by £1 was given by \( (r - g)(1+\delta)/(\delta - g)(1+\delta) \), where \( \delta \) is society’s discount rate between present and future consumption (reflecting differences in the marginal utility of consumption caused by growth in consumption) and \( \delta \) is the individual’s discount rate (which is less than \( r \) because of capital income taxation).
efficiency does not occur if \( r \) eventually becomes greater than \( g \) and stays that way indefinitely, even though \( g > r \) now and for some time into the future. Also, dynamic inefficiency cannot occur in an economy with altruistic bequests in which Ricardian equivalence applies; if there were dynamic inefficiency, bequests would adjust so as to increase dynastic utility.

Implications for saving policy cannot be derived from the dynamic inefficiency argument since it is virtually impossible to know whether the economy is on a dynamically inefficient growth path; that would involve seeing into the indefinite future. For example, we cannot infer from the fact that \( r > g \) now that the economy is dynamically efficient, since it is possible that \( g > r \) eventually.

**Capital Market Inefficiencies**

Inefficiencies on capital markets can arise either from distortions, such as capital income taxes, or from externalities. Evaluating the inefficiencies caused by capital income taxes involves studying the optimal tax system as outlined above. If capital income taxes are set 'optimally', the existence of positive taxes does not reflect under-saving. One obvious reason why capital income taxes might be set non-optimally in practice is that tax systems treat different types of assets differently. Imputed income on consumer durables is typically untaxed, mainly because of measurement problems. The same is true for capital income on human capital investment. Saving for retirement is usually treated preferentially as a matter of policy (for example, to encourage persons to save for their own retirement so as to reduce the need for future state support). Moreover, different types of capital income generated by investment in capital goods are treated differentially by the business tax system: corporations are taxed differently from unincorporated businesses, tangible assets from intangibles, manufacturing from resources, etc. In these circumstances, capital is likely to be allocated inefficiently among uses in the economy. Moreover, given the preferential treatment given to durables and human capital, there is too little investment in capital goods. Given the difficulty of taxing these other forms of investment, a case can be made for encouraging physical capital investment.

Externalities can arise on the saving side or on the investment side. In the former case, the argument is that, if saving for bequests is motivated by intergenerational altruism, it may be like a public good in the sense that all members of the current generation benefit from the amount that each person saves for future generations (e.g. Sen, 1967). In an ingenious argument, Bernheim and Bagwell (1988) have argued that saving for bequests becomes a public good even in a Ricardian world in which each person only cares about their own direct heirs. In this world, people care indirectly about all their future descendants. But, by the natural process of intermarriage among persons of different family lines, all persons become interrelated in the long run and care about each other’s heirs.
therefore exceed the private return, so that we might expect the rate of saving to be too low; indeed, given the extent of the free-rider problem involved, persons may well leave no bequests even though they are altruistic towards future generations. The implication is that government redistribution from current to future generations would be Pareto-improving.

The argument about investment externalities has been given prominence recently with the advent of the ‘new growth theory’ (Romer, 1986), though it goes back to the growth theory of the 1960s. Investment is said to embody technical progress (for example, new knowledge, learning by doing), the returns of which are not fully appropriated by the firms undertaking it. This implies that social returns are higher than private returns so that investment is too low. This constitutes an argument for intervention on the investment side of the market rather than the saving side, at least in an open economy. For both saving and investment externalities, there exists a difficult measurement problem of verifying the extent to which external effects are present.

Intergenerational Equity

If the economy is dynamically efficient, if capital income taxes are set optimally and if externalities are corrected, gains to future cohorts from higher saving can only be attained by redistributing from current cohorts. Naturally, this involves making a value judgement involving the comparison of utilities of different generations. The conventional way to analyse such policies is by the use of an intergenerational social welfare function which explicitly incorporates the trade-off between the welfare of different generations. A convenient form to use is the weighted utilitarian form:

\[
W(u_1, u_2, \ldots) = \sum_{t=0}^{\infty} \frac{u_t}{(1 + \alpha)^t}
\]

where \( u_t \) is the per capita lifetime utility of generation \( t \) and \( \alpha \) is the utility discount rate. Two sorts of value judgements are involved in \( W(\cdot) \). One concerns the way in which the utility function \( u_t \) varies with the lifetime consumption of cohort \( t \). The more rapidly the marginal utility diminishes with consumption, the more inequality-averse is the social welfare function and the greater the tendency to equalise consumption across generations. The second involves the utility discount factor which determines the weight to be given to future generations. For \( \alpha = -\infty \), only the welfare of the first generation matters. In the other extreme, \( \alpha = 0 \) and equal weights are given to the welfare of all generations.

Suppose that there is no technical progress, that population grows at the rate \( n \) and that a planner wishes to maximise \( W(\cdot) \). The optimal path of capital
Taxation and Savings

accumulation can be characterised as follows. If the optimal capital/labour ratio is sufficiently low, so that $r - n > \alpha$, the level of consumption per capita and the capital/labour ratio (saving rate) should be increasing over time, and vice versa. (Note the difference with Feldstein (1977).) The economy approaches a long-run optimum in which $r - n = \alpha$; because of the discounting of future generations' utilities, the long-run optimum involves a smaller capital/labour ratio and a smaller level of consumption than is true of the Golden Rule optimum. The higher the discount rate $\alpha$, the lower are the long-run optimal levels of the capital/labour ratio and of consumption. The speed of adjustment of the economy to the long-run optimum depends upon the extent of inequality aversion in the social welfare function. The more rapidly does the marginal utility of consumption diminish, the more quickly is the capital stock adjusted for a given difference between $r - n$ and $\alpha$. This analysis thus implies that the savings rate is too low if $r - n > \alpha$, and it should be increased more rapidly the greater is the rate at which the marginal utility of consumption diminishes as consumption rises.

3. The Consumption versus Income Tax Debate

From a tax policy point of view, a key issue concerning taxation and savings is whether the base of the direct tax system should include capital or not. This section addresses that issue.

There are two polar forms of personal tax bases which have been advocated in the literature: comprehensive income and consumption, or their equivalents. Comprehensive income ($Y$) is defined to include two components — current consumption ($C$) plus all net accruals to wealth ($\Delta W$, or real saving, $S$). Using the single-period budget constraint of the household, comprehensive income can be written:

$$Y = C + \Delta W = E + rW + A$$

where $E$ represents labour earnings, $r$ is the real rate of return on wealth and $A$ includes autonomous receipts, such as inheritances, gifts received and windfall gains received during the period.

14 The weighting of welfare levels in different generations could reflect their population. For example, if the intergenerational social welfare function is the population-weighted sum of total (rather than per capita) utilities and if total utilities are discounted at the rate $\alpha'$, then per capita utilities will be discounted at $\alpha = \alpha' - n$. The economy converges to an optimum in which $r = \alpha'$.

15 There is an issue as to whether bequests and gifts given ought to be deducted from the right-hand side of the equation. If the giving of gifts and bequests is considered an act of consumption, they should not be deducted. Of course, the gift will then be double-counted to the extent that it gives rise to consumption by the recipient. However, this may be reasonable in that the gift gave rise to two acts of consumption. This issue need not concern us here since precisely the same problem arises under consumption taxation as under income taxation.
There are some problems involved in applying a comprehensive income base. Not all sources of consumption are included in (2), such as consumption from household production and leisure. Moreover, there are several difficulties encountered in measuring capital income, \( rW \). In principle, this should include all forms of returns to assets including interest, dividends, accrued capital gains, capital income from unincorporated business, imputed rent on consumer durables (especially housing) and the imputed return on assets such as transaction balances and insurance. These should all be indexed for inflation and should include an appropriate risk premium. Unfortunately, the measurement of these items is difficult and impractical. Perhaps even more problematic is the fact that, in principle, the return to human capital investment ought to be included as capital income. Including all earnings in \( E \) is equivalent to treating investment in human capital on a cash-flow accounting basis rather than capitalising such expenditures, as is the case with other forms of capital investment. Needless to say, it would be extremely difficult to correctly impute a rate of return to human capital investment. To do so would require, for example, that forgone earnings be capitalised rather than being deducted on a cash-flow basis. This would require information on the depreciation of human capital. Thus, a truly comprehensive income tax is not feasible from a purely administrative point of view.

The alternative, a personal consumption (expenditure) tax, first advocated by Kaldor (1955), avoids some of these problems. The consumption tax base can be obtained from rearranging (2) to give:

\[
(3) \quad C = E + rW - \Delta W + A.
\]

Thus, consumption is simply income less saving. From this, it would appear that all the problems of measuring income reappear in (3). However, except for the difficulty in measuring consumption, this is not the case. It is no longer imperative to measure capital income on an accrual basis or to index capital income for the effect of inflation on asset values. Capital gains which accrue, but are not spent, add both to \( rW \) and to \( \Delta W \) so cancel out. Similarly, inflationary losses in the value of wealth will be offset by the fact that nominal capital income is included in the base. Thus, all accounting for tax purposes can be done on a cash-flow basis, which is relatively easy to administer.

Furthermore, unlike with a comprehensive income tax, returns to capital which take an imputed form, such as imputed rent on housing, need not be measured under a consumption tax. To see this, note that the present value of \( rW - \Delta W \) is simply zero, so omitting both the deduction for saving and the capital income on a given asset from the tax base will not change its present value. This is referred to in the literature as treating assets on a *tax-prepaid* basis, as opposed to that of equation (3), which is the *designated asset* basis (US Treasury, 1977;
Meade, 1978). The tax-prepaid method eliminates the need to measure capital income.

In implementing a consumption tax, some assets could be treated on a designated basis and others on a tax-prepaid basis. Those assets whose returns take the imputed form, such as consumer durables, are natural candidates for the tax-prepaid method. For others, the designated method is more suitable. For example, unincorporated business earnings are best treated on a designated, or cash-flow, basis, because of the difficulty of differentiating capital income from labour income within the firm, and because of the fact that pure rents would go untaxed under the tax-prepaid method. Similarly, investment in human capital is best treated on a cash-flow basis to avoid the problem of having to include forgone income in the tax base. For other assets, such as financial assets, either method would be suitable. In fact, there are advantages in allowing households to decide which assets to treat on a tax-prepaid basis and which to designate. Combining the two bases allows them to arrange their time stream of tax liabilities as they choose, thereby allowing the smoothing of their tax base over time for averaging purposes (i.e. self-averaging). At the same time, since records must be kept of designated assets to ensure that they are taxed when run down, there is an administrative cost involved in designating assets, which households will want to minimise.

In practice, direct tax systems do not tend to conform to either a consumption or a comprehensive income tax, but contain elements of both. While financial asset income tends to be taxed (though not always uniformly), many forms of capital income escape taxation. Housing and other consumer durables tend to be treated as tax-prepaid assets, if not even more generously through mortgage interest deductibility. Pension saving is treated on a cash-flow or designated basis, as is human capital investment. It can be argued that most income tax systems are actually closer to a consumption tax system in the sense that a larger proportion of assets are non-taxable than are taxable.

While the administrative arguments favour consumption taxation, what of the economic arguments? Since the difference between the two concerns the taxability of capital income, the issue revolves around the efficiency and equity arguments for taxing capital income.16

Efficiency Arguments for a Consumption Tax

The choice between a consumption and an income tax concerns whether or not to tax future consumption more heavily than current consumption. Since neither tax applies to leisure, we are in a second-best world in which the outcome is not easily predictable. In the simple case in which there are two periods and labour is

---

16 It should be noted that the choice between an income tax and a personal consumption tax revolves around the base only. In principle, virtually any degree of progressivity can be attained for a given base by the choice of a suitable rate structure.
supplied in the first only, the optimal tax results discussed earlier for the single-household case apply. If first- and second-period consumption are equally substitutable for leisure, a consumption tax would be efficient. If second-period consumption is more complementary with leisure, it should bear a higher tax. In the latter case, this does not necessarily imply that an income tax should be used rather than a consumption tax, that is, that capital and labour incomes be taxed at exactly the same rates; that would require a particular degree of complementarity. Generally, if second-period consumption is complementary with leisure, the optimal tax would require some combination of an income and a consumption tax to ensure that the ideal differential tax on second-period consumption is attained. Things get more complicated once one takes account of the fact that under income taxation it will be impossible to tax all capital income on a par; there will be inter-asset distortions implying that capital is allocated inefficiently among alternative uses. This presumably weakens the case for income taxation. More generally, there has been no optimal tax model developed to date that gives a comprehensive income tax as the preferred outcome. Thus, the case for income taxation must be based on other considerations, such as the ease with which labour income may be converted into capital income by taxpayers.

Equity Arguments for a Consumption Tax

The original proponent of consumption taxation, Kaldor (1955), used essentially an equity argument to make his case. He suggested that persons should be taxed according to what they take out of the ‘social pot’ rather than what they contribute to it. In utilitarian terms, the notion is that one’s well-being is determined by consumption rather than by income. However, as with efficiency considerations, variability of leisure complicates matters as well. Given that neither tax base includes leisure, it is not obvious whether consumption is a better indicator of utility than is income. One might expect by analogy with the efficiency case that income may be a better index of utility if future consumption is sufficiently more complementary with leisure than is current consumption. This turns out to be the case. As we have seen earlier, in a multi-consumer world with non-linear taxation, if the consumption bundle (including both present and future consumption) is separable from leisure, a non-linear consumption tax will be optimal. In the absence of separability, the analysis is more complicated and results have yet to be derived for the multi-period setting. We have seen earlier that with non-linear consumption taxation and linear commodity taxation, the linear tax rate will be higher on goods which are more complementary with leisure. One might expect that with non-linear taxation possible on future consumption, as with a progressive income tax, if future consumption is more complementary with leisure than current consumption is, the optimal redistributive tax system will involve a progressive income tax, generally
Taxation and Savings

alongside a progressive consumption tax. One complicating feature of income taxation is that it does not satisfy horizontal equity. Persons with the same lifetime wealth but with different time profiles of earnings will be treated differently under an income tax. Those whose earnings occur earlier in the life cycle will pay higher taxes.

Thus, the choice between a consumption and an income base involves many considerations, some of which are difficult to verify. Administrative considerations favour consumption taxation, while efficiency and equity arguments can go either way. However, in order to justify income taxation, complementarity of future consumption with leisure is required for both equity and efficiency criteria to be satisfied. Even if such complementarity exists, it is likely that income taxation alone will not be optimal. Given the imperfections of an income base, and the administrative costs of running a joint consumption and income tax system, it is not surprising that the US Treasury Blueprints (1977), the Meade Report (Meade, 1978) and the Economic Council of Canada (1987) all opted for progressive consumption tax. Of course, there are various hybrids of consumption and income taxation that are used in practice. An interesting one is that used in Nordic countries combining flat taxes on capital income with a progressive labour income tax. This avoids many of the administrative difficulties of taxing capital income on a par with labour income while at the same time retaining some capital income taxation for distributive or revenue-raising reasons (Sørensen, 1994).

REFERENCES

Beach, C. M., Boadway, R. W. and Bruce, N. (1988), Taxation and Savings in Canada, Ottawa: Minister of Supply and Services Canada.
Economic Council of Canada (1987), Road Map for Tax Reform, Ottawa: Minister of Supply and Services Canada.


Taxation and Savings


