The Base for Direct Taxation[1]

James Banks and Peter Diamond

[1] This chapter was prepared as part of the Mirrlees Review: Tax Reform for the 21st century. We wish to thank Henry Aaron, Tony Atkinson, Alan Auerbach, Richard Blundell, Malcolm Gammie, Mike Golosov, Jon Gruber, Bob Hall, Dan Halperin, Ken Judd, Louis Kaplow, Gareth Myles, Jim Poterba, Dick Tresch, Aleh Tsyvinski, Iván Werning and Eric Zolt for helpful comments, Ben Etheridge, Catarina Reis, Johannes Spinnewijn and Maisy Wong for research assistance. Financial support from the National Science Foundation under grant SES-0239380 and from the ESRC and Nuffield Foundation through their funding of the Mirrlees review at IFS is gratefully acknowledged.

James Banks is Professor of Economics at University College London and Deputy Research Director of the Institute for Fiscal Studies

Peter Diamond is Institute Professor and professor of economics at Massachussetts Institute of Technology

Introduction

Chapter 2 of the Meade Report,[2] 'The Characteristics of a Good Tax Structure', is divided into six sections: Incentives and economic efficiency, Distributional effects, International aspects, Simplicity and costs of administration and compliance, Flexibility and stability, and Transitional problems. To consider direct taxation in the UK, the Meade Committee examined each of these issues separately and then combined the insights into a policy recommendation. It seems to us, as it seemed to Alfred Marshall, that this is an appropriate way to proceed.[3] While the capacity of computers to find equilibrium in complex models has grown apace since the Meade Report, the models available for analysis, like much of the underlying theory, are still quite limited and still too far from reality to proceed in any other fashion than that followed by the Meade committee. Whilst citing some simulations, this essay focuses on theoretical findings with regard to the tax base.[4]

The traditional starting place for a study of tax reform, such as the Meade Report, is a definition of an ideal tax base, one that reflects both horizontal equity (treating equals equally) and vertical equity (those with larger ideal tax bases pay larger taxes). This ideal tax base is then adjusted in light of the issues raised by the other five areas of concern identified in Chapter 2 of the Report.[5]

---

[2] The Structure and Reform of Direct Taxation, Report of a Committee chaired by Professor J. E. Meade, London: George Allen & Unwin, 1978.

[3] it [is] necessary for man with his limited powers to go step by step; breaking up a complex question, studying one bit at a time, and at last combining his partial solutions into a more or less complete solution of the whole riddle. ... The more the issue is thus narrowed, the more exactly can it be handled: but also the less closely does it correspond to real life. Each exact and firm handling of a narrow issue, however, helps towards treating broader issues, in which that narrow issue is contained, more exactly than would otherwise have been possible. With each step ... exact discussions can be made less abstract, realistic discussions can be made less inexact than was possible at an earlier stage. [Alfred Marshall, *Principles of Economics*, eighth edition. New York: The Macmillan Company. 1948, page 366.]

[4] For a recent optimal tax calculation and discussion of accomplishments and difficulties, see Judd and Su, 2005.

[5] Dedicated taxes for particular expenditures are a common feature of advanced countries (particularly in the context of social insurance) and can play an important political role. And there may be a direct normative gain from doing this in some circumstances. This chapter considers only individual (not corporate) taxation for general revenues.

Since the mid-1960's, there has been a great deal of analysis that considers both equity and efficiency in a single model, rather than discussing them separately. These studies analyze the maximization of a social welfare function that is defined in terms of individual utilities.[6] Equity issues are incorporated by having a heterogeneous population in the model rather than a single representative agent.[7] After arguing briefly in Part I (and further in Part VII.D) that an initial choice of an ideal tax base drawn from an asserted concept of fairness is not a good starting place for policy analysis, the primary purpose of this essay is to review the optimal taxation literature and draw inferences for policy that sets the tax base.[8]

Part II considers lessons from the optimal tax literature with regard to the taxation of income from capital in the presence of taxation of earnings. Part III considers the related issue of the tax treatment of savings. A succession of papers have shown that under certain conditions the optimal tax schedule should not include taxes on capital. This has led some analysts to favor taxing labour income but not capital income or taxing consumption by taxing labour income minus net savings. The analysis discusses both single cohort versions of this result (based on the Atkinson-Stiglitz (1976) theorem) and the infinite horizon result of Chamley (1986) and Judd (1985), the former addressing the problem from the perspective of decisions over the lifetime of a single generation, and the latter looking at an economy of multiple generations. In both cases, however, the required conditions for the optimality of zero taxation of capital income, however, are argued to be too restrictive and the finding of no role for capital taxation is therefore

[6] Some studies consider properties of taxes that result in individual utilities such that it is not possible to make everyone better off, given the set of allowable taxes. The set of such utilities is referred to as the second-best Pareto frontier.

[7] The standard basic model treats administrative costs of different taxes as zero or (implicitly) infinite and ignores tax evasion. See, for example, the textbooks by Myles, 1995, Salanié, 2003, Tresch, 2002, Tuomala, 1990, although there are articles that address administrative costs and evasion. There has not been integration with macro issues incorporating, for example, built-in stabilizers (Auerbach and Feenberg, 2000) nor has the incorporation of international issues (trade, investment, migration) included the macro dimensions of those issues.

[8] In terms of the Chapter 2 topics of the Meade report, we do not consider administrative costs (ignoring them for given tax bases), international aspects (analyzing closed-economy models), nor the use of taxes as part of discretionary fiscal policy for macroeconomic stabilization. Oddly, the Meade report ignores built-in stabilizers, which seem to us to matter. Other chapters in this volume contain discussions of issues not considered here, including tax rates, the presence of families, some administrative issues and corporate taxation. For some administrative issues in a consumption tax, see Bankman and Schler (2007).

considered not robust enough for policy purposes. Hence there should be some role for including capital income as a part of the tax base. However, the conclusion that capital income should be taxed does not lead to the conclusion that the tax base should be total income, the sum of labour income and capital income. At present, the literature has only a little to say about how to combine the two sources of income to determine taxes.

In Parts II and III, the rate of return is assumed to be fixed and known. Part IV examines some issues when there are alternative investment opportunities with safe and risky rates of return. Part V discusses age-dependent taxes (for example different taxation of earnings for workers of different ages). Part VI examines some implications of recognizing diversity in individual savings behavior. Part VII touches on a number of issues including a further discussion of the use of a social welfare function (VII.A.), government commitment (VII.B.), some modeling assumptions (VII.C.), and horizontal equity (VII.D.). Part VIII presents some empirical underpinnings for two key elements in determining the desirable taxation of capital income – differences in savings propensities and the shape of earnings (and uncertainty about earnings) over the lifetime. Part IX sums up and concludes.

This chapter leaves to other chapters in this volume discussion of the provision for the very poor and concern about inheritances. It also leaves to another chapter discussion of taxation that recognizes the existence of families. And the chapter assumes that annual measurement of wealth is not available and so considers annual capital income taxation instead.[9] While the Meade Report was part of a tradition contrasting taxation of annual income with taxation of annual expenditures, the Report's inclusion of annual taxation of wealth along with taxation of expenditures in its policy recommendation represented a departure from previous debates based on choosing between either income or expenditure taxation. This chapter shares the Meade Report

[9] While the values of some types of wealth are readily measurable, others are not. Of course the same is true for accruing capital income. In practice, this is addressed by taxing realized incomes. Such taxation could be, but is not, adjusted to offset the difference between accrual and realization taxation. We are not aware of a literature exploring the relative advantages of wealth and capital income taxation (with the latter supplemented by wealth taxation at death) as part of optimal taxation. Our conjecture is that capital income taxation could do better, but that is just a conjecture awaiting analysis.

framing of the potential simultaneous use of several tax bases and focuses on three questions.

- If there is annual non-linear (progressive) taxation of earnings, how should annual capital income be taxed – not at all, linearly (flat rate, as in the Nordic dual income tax[10]), by relating the marginal tax rates on capital and labour incomes to each other (as in the US[11]), or by taxing all income the same?
- If there is annual non-linear taxation of earnings, should there be a deduction for net active savings?[12]
- If there is annual non-linear taxation of earnings, is it worth having a more complex tax structure, particularly age-dependent tax rates? Would greater use of age-dependent rules in capital income taxation be worthwhile?

The chapter reaches the conclusions that neither zero taxation of capital income nor taxing all income the same are good policy conclusions. The chapter leans toward relating marginal tax rates on capital and labour incomes to each other as opposed to the Nordic dual tax. In parallel, the chapter reaches the conclusion that there should not be a full deduction for all of net savings. And the chapter concludes that age-dependent tax rates seem to offer enough advantages to justify the added complexity, although more research is needed to support this conclusion.

[10] On the Nordic dual tax, see Sørensen, 2001, 2005.
[11] In the US, the rate of tax on capital gains and dividends, generally 15 percent, is lowered for individuals whose marginal tax rate is 15 percent or less. In the past, half of capital gains were included in taxable income, also resulting in a marginal rate that varied with overall taxable income.
[12] Active savings are savings from earnings. Thus earnings minus net avtice savings equals income minus net savings.

Part I. Horizontal equity and the choice of tax base

Going back at least to Adam Smith, economists have asserted what the base for taxation should be (along with the degree of progressivity, given the chosen tax base).[13] [14] The Meade Report states:

> No doubt, if Mr Smith and Mr Brown have the same 'taxable capacity', they should bear the same tax burden, and if Mr Smith's taxable capacity is greater than Mr Brown's, Mr Smith should bear the greater tax burden. But on examination 'taxable capacity' always turns out to be very difficult to define and to be a matter on which opinions will differ rather widely." (Page 14.)

This is a definition of an ideal tax base, in the sense that it is underpinned by a direct view or argument about what is ideal. But it still relies on a further definition of taxable capacity, and reflecting the acknowledged difficulty in defining taxable capacity, the Report goes on to ask: "Is it similarity of opportunity or similarity of outcome which is relevant?" and "Should differences in needs or tastes be considered in comparing taxable capacities?"[15] Historically, the debate over the appropriate base for annual taxation has been an argument between two approaches. One is that total (Haig-Simons)

[13] "The subjects of every state ought to contribute towards the support of the government, as nearly as possible, in proportion to their respective abilities; that is in proportion to the revenue which they respectively enjoy under the protection of the state." Adam Smith, Wealth of Nations, New York: The Modern Library, 1937. page 777.

[14] Historically there have been two different approaches to an ideal tax base – one drawn from ability to pay and one drawn from the benefits received from government spending. Discussion of the pattern of benefits received from government spending programs that affect the entire population did not achieve any consensus on its distributional significance and has disappeared from discussion of an ideal tax base. For example, it is hard to see how to allocate the benefit of military spending by income level in a way that is not too arbitrary to be useful. For historical discussion, see Musgrave, 1959.

[15] The Meade report is not the only examination of taxation that concludes that taxable capacity is hard to define in a way to compel wide acceptance, as is needed for the role as an agreed-on normative basis. For example, Vickrey (1947) writes: "In a strict sense, 'ability to pay' is not a quantity susceptible of measurement or even of unequivocal definition. More often than not, ability to pay and the equivalent terms "faculty" and "capacity to pay" have served as catch-phrases, identified by various writers through verbal legerdemain with their own pet concrete measure to the exclusion of other possible measures. Ability to pay thus often becomes a tautological smoke screen behind which the writer conceals his own prejudices." [footnote omitted] (page 3-4.)

income[16] is the best measure of ability to pay and therefore horizontal equity calls for Haig-Simons income as the tax base. The other, argued particularly in Kaldor (1955), is that annual consumption is the best measure of ability to pay and therefore horizontal equity calls for consumption as the tax base. This latter view is generally supported by the further argument that it is better to tax people on what they take from the economy (consumption) than a measure of what they provide (income).

We agree with the Meade Report that "'taxable capacity' always turns out to be very difficult to define and to be a matter on which opinions will differ rather widely." We conclude that the consideration of an ideal tax base lends itself to too many concerns and conflicting answers to be viewed as a good starting point for the consideration of taxation. An alternative start is by examining the economic equilibria that occur with different tax structures.[17] That is, for any tax structure (assuming it generates enough revenue to cover government expenditures), there is an economic equilibrium,and that equilibrium will result in particular levels of lifetime wellbeing for all the people in the economy. Given a social welfare function relating aggregate benefit to the distribution of individual lifetime utilities, these lifetime utilities can therefore become the basis for evaluating the normative properties of the various alternative equilibria. This is the starting place of an optimal tax approach to tax policy. Thus, optimal tax theory is based on a consequential philosophy. For each tax structure it describes the economic equilibrium, and thus the utility levels of the different economic agents. Then it asks which of these equilibria offers the utility levels judged best by a social welfare function (an increasing function of individual utilities, which thereby incorporates concern about distribution in terms of utilities, not incomes).

[16] Haig-Simons income is labour income plus accrued capital income – Haig (1921), Simons (1938). Shaviro (2002) notes that "the spirit in which this hypothetical measure [relevant to distributive justice] is discussed (or, rather, deliberately not discussed) was well illustrated by Henry Simons (1938, 31), when he argued that attempts to poke too far behind the supposed objectivity of an income definition "lead directly back into the utter darkness of 'ability' or 'faculty' or, as it were, into a rambling, uncharted course pointed only by fickle sentiments.""

[17] Traditionally economics has been consequentialist in this sense, as shown, for example, by the centrality of the Fundamental Welfare Theorem examining conditions under which there is equivalence between competitive equilibrium and Pareto optimality. A Pareto optimal allocation is one from which it is not possible to increase the utility of one household without decreasing utility for another.

With an optimal tax approach, some aspects of horizontal equity can be addressed by viewing horizontal equity arguments as providing limitations on the set of allowable tax policies, as has been argued by Atkinson and Stiglitz (1980). This chapter accepts the view that tax tools should be limited by such equity considerations and that policies should be restricted to ones that are uniform over their stated tax base, i.e. tax systems in which those with equal circumstances in the relevant dimensions are treated equally.[18] Tax tools should also reflect administrative and political feasibility. One would need a great deal of faith in the political process not to want some protections against arbitrary tax assessments under the guise of "better taxation." A complication in structuring protections lies in the definition of arbitrary. If one actually can increase social welfare by drawing distinctions between individuals, are the distinctions still arbitrary? A concern with actual and possible motivations in the political process should lie behind restrictions on tax policies, and the concept of horizontal equity is likely to be very helpful in addressing this issue, without necessarily being the starting place for tax analysis.

Although much has been learned about earnings taxation in one-period models since the pioneering work in Mirrlees (1971), one-period models lack an intertemporal dimension suitable for considering the relative tax treatment of capital and labour incomes. When one moves to intertemporal settings a source of concern about the formulation of the objective function individuals are assumed to maximize arises to the extent that some people may not exhibit time consistency in their behavior.[19] Since this issue is indeed central to the analysis of the relative taxation of capital and labour incomes, the chapter returns to it in Part VI, after first exploring implications of models

[18] The condition of uniform taxation given the base rules out randomized taxation, which, under some circumstances, can raise social welfare. Nevertheless, randomized auditing of returns does not seem unfair to us or, apparently, to the public as long as the probabilities are suitably selected and the audits are not unduly unpleasant.

[19] Time consistency is the property of making the same decision when given the same choices under the same circumstances at different times. Time inconsistency occurs when different choices are made even though the circumstances are the same. Analyses with time-inconsistent quasi-hyperbolic preferences and with the simple assumption that some people do no savings at all do not reach the same conclusions as the usual full rationality model where individuals are consistent in their desire to borrow and save in anticipation of future events. A similar issue of the appropriate objective function for social evaluation arises if the analyst is concerned that individuals discount the future excessively even if they are time-consistent.

with fully-rational agents. For now, the chapter simply proceeds with preferences that are assumed to be fully rational, time-consistent. This approach is based on the idea that a good starting place for policy is the policy for fully-rational agents, a policy that can then be adjusted in recognition of the inadequacy of the assumption that all individuals show fully-rational behavior. For example, in considering the taxation of capital income, the chapter first asks how that should be done in an economy with only fully-rational agents and then asks (in Part VI) about adjustment in recognition that some fraction of agents do not appear to save enough for their own good and others accumulate vast sums, not aimed at later consumption. Even the first step, with fully rational agents, is complex given the many relevant aspects of the economic environment, which are modeled separately in optimal tax analyses because of the difficulty in making inferences if the model has many complications at the same time.

The focus in this chapter is on the relative taxation of labour and capital incomes, not the relative merits of taxing total (Haig-Simons) income and taxing consumption, as has commonly been the focus of analyses.[20] In the end, the Meade Report effectively did the same – the Report closes with a section entitled "ULTIMATE OBJECTIVES:"

> We believe that the combination of a new Beveridge scheme (to set an acceptable floor to the standard of living of all citizens), of a progressive expenditure tax regime (to combine encouragement to enterprise with the taxation of high levels of personal consumption), and of a system of progressive taxation on wealth with some discrimination against inherited wealth, presents a set of final objectives for the structure of direct taxation in the United Kingdom that might command a wide consensus of political approval and which could be approached by a series of piecemeal tax changes over the coming decade." (Page 518.)

Thus with a tax on expenditures and a tax on wealth, the Meade report did not keep a simple measure of taxable capacity as the basis for taxation, although it argued that wealth and consumption were both relevant for measuring taxable capacity. The chapter discusses equity further in VII.D.

[20] See, for example, Aaron, Burman, and Steuerle (2007), Bradford (1986), Pechman (1980).

Part II. Optimal taxation of capital and labour income

Optimal tax theory uses simple general models and calculated examples to draw inferences about how taxes should be set in order to strike a balance between equity and efficiency concerns. Different weights on the concern for equity naturally lead to different taxes.[21] So the theory is designed to show a relationship between normative concerns and tax bases and rates. The approach is to consider economic equilibria under different tax structures and to examine which tax structure gives an equilibrium with the highest social evaluation of the lifetime utilities of the participants in the economy. The specific optimal taxes from any particular model are not meant to be taken literally, but insights from the modeling, when combined with insights from other sources, can help lead to better taxes. That is, just as the Meade Report had multiple concerns beyond its concern with taxable capacity, so too, the optimal tax approach is a starting place, to be combined with concerns that are not in the formal modeling. One additional concern of particular relevance is the complexity of the tax structure. A desire to avoid complexity comes from seeking simplicity in the tasks of taxpayers, tax collectors and tax-setting legislatures. There are many papers that analyze optimal taxes; and they differ in many ways. This chapter is not a survey of methods and model results, but a selective drawing of some key policy inferences from the literature.

In each year, there are taxpayers with labour income and taxpayers with capital income and taxpayers with both. Apart from previously deferred compensation, labour income comes from time spent working during the year. Earnings are also influenced by earlier decisions about education, on-the-job training, job location, and job history. Capital income within the year comes primarily as a result of the previous accumulation of assets and liabilities on which capital income is earned and paid. Savings and portfolio decisions during the year are influenced by anticipated taxes in future years. Anticipated

[21] Formally, differing concerns about equity are incorporated by the choice of a particular cardinalization of ordinal preferences and the degree to which the social evaluation of an individual's utility varies with the individual's level of utility.

future taxes have some relevance for earnings as well, with future earnings being a substitute for current earnings in financing lifetime consumption. Focus on taxation in a single year, without consideration of both earlier and later years, is thus incomplete. This incompleteness is more significant for consideration of taxes on capital income than on labour income. This distinction between the roles of the two types of income on a lifetime basis is the basis for consideration of intertemporal models, even when considering taxation levied on an annual basis.[22]

Taking a lifetime perspective, some policy analysts have called for ending the taxation of capital income.[23] This position is based, at least in part, on optimal tax modeling that reaches this conclusion. This chapter presents separately the two arguments for zero taxation of capital income that have been important for the thinking of many economists, and then shows their lack of robustness to changes in the underlying assumptions, changes that are empirically important. The analysis also serves as background for considering the polar opposite policy of basing taxation on total income, the unweighted sum of labor income and capital income. Why this alternative has not received support from optimal tax analyses is discussed briefly below.

A. A simple two period model of work and retirement

Our starting place for consideration of the taxation of both labour income and capital income is a model with two periods, with labour supply in the first period and consumption in both the first and second periods.[24] Suppressing a role for taxing initial wealth (discussed briefly in II.C and VII.B), savings from first-period earnings, used to finance second-period consumption, generates capital income that is taxable (in the second period). Since there is only a single period of work, the model can be viewed as shedding light on the taxation of savings for retirement. For an analysis of issues relating

[22] The analysis in this chapter ignores the existence of a corporate income tax and reasons for having one. The focus is on taxing individuals. The presumption is that the suitable role for a corporate income tax builds on the desired role of taxation of individual capital income, not vice versa.
[23] See, for example, Atkeson, Chari, and Kehoe (1999), Weisbach (2006), and Bankman and Weisbach (2006).
[24] Interpreting the solution from such a model should be in terms of the total taxation that falls on the tax base, not just the particular form of tax used in describing the model.

to the taxation of early life savings that are intended for possible consumption during mid or late working life one would need a model with two separate labour supplies, representing labour supply at different times or ages. Such models are considered in II.B.

A good place to start considering this class of models is the well-known Atkinson-Stiglitz theorem (1976) which states that when the available tax tools include nonlinear earnings taxes differential taxation of first- and second-period consumption is not optimal if two key conditions are satisfied: (1) all consumers have preferences that are separable between consumption and labour and (2) all consumers have the same sub-utility function of consumption.[25] The first condition states that the marginal benefit derived from consumption over the life-time should not depend on labour supply, and the second requires all consumers to be similar in their desire to smooth consumption across their life-cycle and across potentially uncertain states of the world. Like the Fundamental Welfare Theorem, this theorem can play two roles – one is to show that limited government action is optimal in an interesting setting, and the second is to provide, through the assumptions that play a key role in the theorem, a route toward understanding the circumstances calling for more government action (in this case distorting taxation of savings and therefore implicitly taxing (or subsidizing) consumption in the second period relative to consumption in the first period). While we present the intuition behind the first use, our focus is on the second use as we identify in differing tastes and uncertainty about future earnings two strong reasons for finding the theorem not a good basis for policy, for finding that some taxation of capital income is part of a good tax system.

The theorem refers to not "differentially taxing first- and second-period consumptions." That is, a tax on consumption that is the same in both periods (a VAT or retail sales tax) is equivalent to a tax on earnings since the choice between first- and second-period consumptions financed by net-of-tax earnings does not alter the total taxes paid (on a present discounted value (PDV) basis). It is different tax rates that matter for

[25] Separability between labour and the vector of consumptions and the same subutility function for all individuals can be expressed as $U^n\left[x_1, x_2, z\right] = \tilde{U}^n\left[B\left[x_1, x_2\right], z\right]$, with x_1 and x_2 being consumption in each of the two periods and z being earnings. A special case is the convenient and widely used additive function $U^n\left[x_1, x_2, z\right] = u_1\left[x_1\right] + u_2\left[x_2\right] - v\left[z/n\right]$.

efficiency by introducing a "wedge" between the intertemporal marginal rate of substitution (MRS) and the intertemporal marginal rate of transformation (MRT) between consumer goods in different periods.[26] Two ways of having differential taxation of consumption in the two periods are through different tax rates on consumption in the two periods and through taxation of the capital income that is received as part of financing second-period consumption out of first-period earnings. That is, if taxes should not distort the timing of consumption (if the MRS should equal the MRT), then the optimum is not consistent with taxing these consumer goods other than with equal rates, and thus inconsistent with taxing savings at the margin. The theorem extends to having multiple periods of consumption with a single period of labour.

The underlying logic of the theorem extends to additional settings beyond the full optimization of social welfare. Konishi (1995), Laroque (2005) and Kaplow (2006a) consider distortionary taxes in environments with the same preference assumptions, and any earned income tax function. They show that one can always move to a system of non-distorting consumer taxes coupled with an appropriate modification of the earned income tax and generate more government revenue whilst leaving every consumer with the same utility and the same labour supply.[27]

The underlying logic behind the Atkinson-Stiglitz result starts with the observation that the incentive to earn comes from the utility achievable from consumption purchases with after-tax earnings. With separable preferences and the same subutilities for everyone, differential consumption taxation can not accomplish any distinction among those with different earnings abilities beyond what is already accomplishable by the earnings tax, but would have an added efficiency cost from distorting spending. Thus the use of distorting taxes on consumption (MRS unequal to

[26] The intertemporal consumption MRS captures the consumers valuation of consumption in the second period relative to consumption in the first period. The matching MRT represents the ability of the economy to produce more of the latter by producing less of the former and would be typically reflected in the price of moving consumption between periods. When these ratios are not equal, a change in production can increase utility, if everything else is held constant.

[27] If labour supply is smooth in response to uniform transfers to all consumers (no jumps in labour supply), then this revenue gain can be used to make a Pareto improvement.

MRT) is a more costly way of providing the incentives for the 'optimal' earnings pattern in equilibrium.

Of course, an argument that a better policy is available should only be used as an argument against a particular policy proposal if the available alternative is actively pursued. As with the inadequacy of the Hicks-Kaldor-Scitovsky criterion,[28] hypothetical alternatives that would not be adopted are not legitimate arguments against a policy that would increase social welfare. That is, one can argue against a distorting consumption tax that would increase progressivity in taxation by preferring an alternative of increasing the progressivity of the income tax if the increased income tax progressivity is more efficient. However, arguing on the basis of the existence of a dominating proposal is somewhat hypocritical if the dominating proposal is not supported and will not be adopted or pursued for adoption in the future.

The logic behind the Atkinson-Stiglitz theorem gives insight into several changes in assumptions, discussed below, that would no longer lead to the conclusion in the Atkinson-Stiglitz model that there should be no taxation of capital income.[29] Considered first are two changes to preferences – non-separability and then non-uniform separability. Further changes, some of which involve two periods of work are then also analyzed.

[28] The Hicks-Kaldor-Scitovsky criterion is that a policy change can be considered worth doing if those made better off could fully compensate those made worse off by the policy change. Hence the policy change could lead to a Pareto improvement. The original version was faulted in that a policy change can pass the test but, having been implemented, canceling the policy could also pass the test. The refined criterion is therefore that a policy change can be considered worth doing when a policy passes the test and canceling the policy does not pass the test. The criterion can be faulted for being hypothetical if the compensations do not occur as part of the reform. We agree that hypothetical alternatives do not have the ethical standing needed to support a normative use of the criterion. A similar view is implicit in the condition of the Independence of Irrelevant Alternatives in the Arrow Impossibility Theorem.

[29] The theorem assumes no restriction on the allowable shape of the taxation of earnings. Deaton (1979) notes that if the income tax is constrained to be linear, then the Atkinson-Stiglitz conditions that are sufficient for the non-taxation of capital income with optimal nonlinear taxation are no longer sufficient for the result. A further condition is needed when the income tax function must be linear even when preferences are weakly separable between goods and leisure (as in Atkinson-Stiglitz) - that all consumers have parallel linear Engel curves for goods in terms of income. Thus, even with weak separability and uniformity of preferences, different savings rates for different earners because of nonlinear or nonparallel Engel curves prevent the general holding of the result. Note that this argument applies as well to each piece of a piecewise linear tax function, with application of the condition to those on a single linear stretch of the tax function. That is, with a linear income tax and differing savings rates, a change in the income tax rate cannot reproduce the tax pattern from taxing savings and without the ability to reproduce a change in the tax rate can not generally be a dominant policy change.

One obvious change would be that preferences do not exhibit separability between consumption and labour. Then the Corlett-Hague (1953) style analysis in a representative agent 3-good model (current work, current consumption, and future consumption) can examine whether a move towards taxing savings or towards subsidizing savings raises welfare.[30] The key issue is the pattern of the cross-elasticities between labour supply and consumptions in the two periods. However, we do not know much about these cross-elasticities and thus do not have clear policy implications. Although the commonly-used assumptions of atemporal and intertemporal separability[31] strike us as implausible, that does not lead to a straightforward conclusion about the cross-elasticities. In particular, those in the second period (who are retired) have more time to do home production (and so less reason to value financing from first-period earnings) than those in the first period, but also more time to enjoy consumption opportunities that are time-intensive (and so more reason to value financing from first-period earnings). It is not clear which of these two effects dominates, and hence which cross-elasticity is higher. Consequently, it is not clear whether savings should be taxed or subsidized because of this issue.[32]

Even were seperability to be preserved, a second consideration would be that the subutility functions of consumption are not the same for everyone. Saez (2002b) presents an argument against the policy applicability of the Atkinson-Stiglitz theorem based on differences in desired savings rates across individuals with different skills. Saez argues that it is plausible that there is a positive correlation between labour skill level (wage rate) and the savings rate and cites some supporting evidence. [33] (We review some of the

[30] Results in models with a representative agent are not necessarily the same in many-person models with heterogeneous agents. Nevertheless the results are suggestive that some results will continue to hold, possibly with modified conditions.

[31] For atemporal additivity, utility within a period can be written as a sum of a utility of consumption and a disutility of work. For intertemporal additivity, utility over a lifetime can be written as a sum of utilities in each period.

[32] Recognition of home production is an argument for differential taxation of different goods at a point of time (Kleven, Richter and Sørenson, 2000), but does not appear to help clarify the issue of intertemporal taxation.

[33] Dynan, Skinner and Zeldes (2004) report that those with higher lifetime incomes do save more in the US, but that the full pattern of savings requires considerable complexity in the underlying model (including uncertainties about earnings and medical expenses, asset tested programs, differential availability of savings vehicles, and bequest motives) to be consistent with the different aspects of savings at different ages that they discuss. Thus the higher savings rates are consistent with the behavioral assumption of Saez,

evidence on individual savings in VIII.A.) In the Atkinson-Stiglitz two-period certainty setting with additive preferences, this pattern of savings rates is consistent with those with higher earnings abilities discounting future consumption at a lower rate.[34] In terms of the conditions of the Atkinson-Stiglitz theorem, Saez preserves separability in preferences but drops the assumption that the subutility function of consumption is the same for everyone. With the plausible assumption that those with higher earnings abilities discount the future less (and thus save more out of any given income), then taxation of savings helps with the equity-efficiency tradeoff by being a source of indirect evidence about who has higher earnings abilities and thus contributes to more efficient redistributive taxation.[35] In the context of this issue, how large the tax on capital income should be and how the marginal capital income tax rates should vary with earnings levels has not been explored in the literature that has been examined. The optimal rate would depend on the magnitude of the differences in savings propensities and on the elasticities that matter for distortions.

1. Allowing for uncertain earnings

In the Atkinson-Stiglitz model, a worker is assumed to know the return to working before deciding how much to work and, since work is in the first period only, knows full life-time income before doing any consumption. Uncertainty about earnings from a given labour supply does not influence optimal taxation of savings if the

---

but not, by themselves, a basis for necessarily having the discount rate pattern that Saez assumes, since these other factors are also present. From the perspective of this essay, it seems to us more plausible that there is the assumed correlation in parameters than that it is absent, and so the implication for taxes from this class of models is supportive of positive taxation of capital income, not zero.

[34] Saez works with the utility functions $U^n\left[x_1, x_2, z\right] = u_1\left[x_1\right] + \delta_n u_2\left[x_2\right] - v\left[z/n\right]$, with δ_n increasing in n.

[35] Saez derives a condition for the impact of introducing a linear tax on capital income in a setting of optimal taxation of earnings. He shows that this impact is generally nonzero, implying that a zero tax is not optimal. He gives conditions to sign the direction of improvement. In a setting of generally nonlinear taxation and two worker types, the optimum involves positive (negative) marginal taxation of capital income when the optimum has positive (negative) marginal taxation of labour income. A parallel condition holds for the introduction of a small linear tax on capital income. Positive taxation is the relevant case.

Within the standard discounting framework there appears to be considerable heterogeneity in the population in discounting of the future. For example, see Hausman (1979) on different discount rates for air conditioner purchasers, or Samwick (2006) on the distribution of discount rates that can rationalize the distribution of retirement saving wealth.

uncertainty is resolved before first-period consumption - the Atkinson-Stiglitz result carries over. But were consumption decisions to be taken before earnings uncertainties are resolved then this would impact the Atkinson-Stiglitz result. This point can be illustrated in a model with a single period of work before turning to the more relevant models with work in successive periods.

Modifying the model so that earnings occur only in the second period (with probabilities but not exact information as to future earnings known in the first period) would imply that the first-period consumption decision is made before the uncertainty about future earnings is resolved, while second-period consumption occurs after.[36] [37] The Atkinson-Stiglitz result no longer holds and second-period consumption should be taxed at the margin relative to first-period consumption (Cremer and Gahvari, 1995). This result holds whether there is general taxation of earnings and savings or only a linear tax on savings with a nonlinear tax on earnings.

We can see the underlying logic of this result by comparing it with that of taxing savings when higher earners have smaller discount of the future. To do that, it is useful to consider the problem of welfare maximization in terms of "incentive compatibility constraints." A natural starting place for optimizing taxation is to consider alternative tax structures by first determining the equilibrium that happens with each tax structure. Then the social welfare at the different equilibria are compared. In mathematical vocabulary, social welfare is maximized subject to the constraint of the equilibrium that occurs with individual behavioral responses to the chosen tax structure. There is a mathematically equivalent way of setting up the maximization which is helpful for intuition, even though it does not comply with how a government would naturally approach choosing a tax structure.

---

[36] Formally, the skill level, n, is a random variable, with distribution $F[n]$. First-period consumption must be chosen independent of the as-yet unknown skill level, while earnings and second-period consumption depend on the skill level, which becomes known before these decisions are made. With additive preferences expected utility is written as $\int \left(u_1[x_1] + u_2[x_2[n]] - \mathrm{v}[z[n]/n] \right) dF[n]$, with a separate budget constraint for each value of n and taxes depending only on the realized level of earnings.

[37] With annual taxation, consumption during the year is happening before earnings levels later in the year are known, at least for some workers. This parallels analyses of the demand for medical care with an annual deductible or out-of-pocket cap.

Consider the mathematical problem of a government deciding how much each person should earn and how much each person should consume in each period (with the relationship among these being an implicit description of the taxation of earnings). The government decision is subject to the resource constraint of the economy. If this is to be mathematically equivalent to the effects of a tax structure, the relationship between consumer spending and earnings (the implicit tax function) cannot be different for individuals with the same earnings. Given that uniformity, the government's consumption and earnings plan will be an economic equilibrium with a tax function if each person is willing to have his earnings and consumption under the government's plan rather than having the earnings and consumption pair of anyone else. Uniform rules for everyone is referred to as allowing each person to imitate the consumption and earnings of any other person, within the bounds of the individual's feasible earnings levels. The constraint on the government's plan that no one prefers to imitate someone else is referred to as an incentive compatibility constraint. This equivalent formulation allows a discussion of optimal taxes in terms of affecting the ease of imitating someone else. A change in implicit taxes that makes it less attractive for someone with high earnings skills to imitate someone with low earnings skills allows the government optimization to be more effective, that is, improves the equity-efficiency tradeoff (weakens the impact of the incentive compatibility constraint).

After that mathematical digression, let us return to comparing the results about taxing savings with random earnings and when higher earners discount the future less. In the latter case a worker choosing to imitate someone with less skill (by earning less than he would otherwise) saves more than that worker with less skill since the discount of future consumption is less for the potential imitator. Thus taxing savings eases the incentive compatibility constraint, having a bigger impact on the would-be higher skill imitator than on the lower earner potentially imitated. That is, it makes such imitation less attractive. In the uncertainty case, a worker planning to earn less than the government planned amount in the event of high opportunities has a higher valuation of savings than if the worker were planning to earn more by following the government plan (assuming normality of consumption). Thus, again, taxing savings eases the incentive

compatibility constraint. One example is that retirement tends to be at an earlier age for those with more accumulated savings (earnings opportunities held constant). Thus, discouraging savings encourages late retirement. This logic only holds for workers doing optimal savings, a point to which we return in Part VI.

Next, the chapter considers models with labour supply in both periods. Then, in parallel with this section, with uncertain second-period wages, first-period consumption is occurring after first-period opportunities are realized but before second-period opportunities are known. The same advantage of differential tax treatment of first- and second-period consumptions naturally occurs in this setting.

B. A two period model of working life

While the model with a single labour supply decision can shed light on the relative tax treatment of consumption when working and when retired, a model with two labour supply decisions addresses issues about consumption and earnings during a career. It also raises some issues of the sensible degree of complexity of tax structures, that are not present in the single-labour supply model.

Consider a setting where individuals work in each of two periods and consume in each of two periods. In the certainty setting with a single period of work discussed above, the starting place was a model where people differed only in their wage per hour of work. To extend the certainty analysis, we now characterize people by a pair of wage rates, representing the wage rates in each of the two periods. As above, we take wage rates to be the only differences across workers in the population. In light of the diversity in age-earnings trajectories, it is natural to assume diversity in the growth of wage rates.[38]

The Atkinson-Stiglitz result, that with separability and uniform subutilities of consumption[39] there should not be a distortion in the intertemporal consumption

[38] We continue to ignore worker decisions that influence future wage rates (investments in human capital).
[39] Separability between labour and the vector of consumptions and the same subutility function for all individuals can be expressed as $U^{n_1,n_2}\left[x_1, x_2, z_1, z_2\right] = \tilde{U}^{n_1,n_2}\left[B\left[x_1, x_2\right], z_1, z_2\right]$, with x_1 and x_2 being consumption in each of the two periods and z_1 and z_2 being earnings. A special case is the

decision, extends to this case provided that the taxation of earnings over a lifetime depends in a fully general way on earnings in both periods. That is, in the first period of a lifetime, there is taxation of earnings that can be thought of as withholding of taxes while waiting for the determination of lifetime taxes, which will depend on earnings in both periods.[40] With the Atkinson-Stiglitz preference assumptions and an optimal lifetime tax structure, it remains the case that the marginal rate of substitution between first- and second-period consumptions should equal the marginal rate of transformation. This corresponds to an absence of taxation on savings out of after-tax first-period earnings.

As with the analysis of models with a single working period, the result of zero taxation of capital income does not hold if discount factors vary with skill or if there is uncertainty about second-period earnings, both of which seem empirically important. Beyond the theoretical result that there should be positive taxation of capital income in a model with uncertain later-period earnings, we can look at simulation results to see how important and how large such a tax might be. Conesa, Kitao and Krueger (2007) have done a complex simulation of the asymptotic position of an empirically calibrated overlapping generations (OLG) model with uncertain individual wages and lengths of life. They have a three-parameter earnings tax (the same for each age), a 100 percent estate tax financing poll subsidies, a pay-as-you-go social security system, a linear tax on capital income and no government debt or assets. They choose taxes to optimize the long-run position of the economy and find a capital income tax rate of 36 percent, while the tax on labour income is nearly linear at 23 percent.[41] Golosov, Tsyvinski and

---

convenient and widely used additive function
$U^{n_1,n_2}\left[x_1,x_2,z_1,z_2\right]=u_1\left[x_1\right]+u_2\left[x_2\right]-v_1\left[z_1/n_1\right]-v_2\left[z_2/n_2\right]$.

[40] Writing lifetime taxes (in present discounted value) as $T\left[z_1,z_2\right]$, the budget constraint for a worker is $x_1+R^{-1}x_2=z_1+R^{-1}z_2-T\left[z_1,z_2\right]$, where R is one plus the rate of return on capital. If there was tax collection in the first-period, $T_1\left[z_1\right]$, it would still be the case that the tax collected in the second period, $T_2\left[z_1,z_2\right]$, would depend on both earnings levels, and the budget constraint would, equivalently, be written as $x_1+R^{-1}x_2=z_1+R^{-1}z_2-T_1\left[z_1\right]-R^{-1}T_2\left[z_1,z_2\right]$.

[41] Optimizing a long-run economic position is different from looking at the long-run position of an optimized economy. Increasing the capital stock has additional costs in a full optimization that are not

Werning (2007) examine a two period model where there is a wide range of worker productivities in the first period and each worker has a probability of one-half of losing half of first-period productivity in the second period. They allow a fully general tax structure, referred to as a mechanism design optimization.[42] Given the special nature of the economy (with no attempt to resemble an actual economy), the level of implicit marginal taxes (referred to as wedges) are not of direct interest, but the pattern of implicit marginal taxes may have robustness. They find a higher implicit tax on second-period consumption (on savings) the higher the wage rate of the worker in the first period.[43] While this model is very special, there is little else that casts light on the best pattern of a capital income tax.[44]

Beyond the two arguments detailed above, there is also an issue of the complexity of the tax structure needed for the zero tax result. The extension of the Atkinson-Stiglitz theorem to the setting with two periods of earnings (with separability and uniform subutility functions) potentially requires a complex tax structure with the marginal taxes in any year dependent on the full history of earnings levels. For example, in a setting of two periods with two labour supplies, lifetime after-tax consumption spending can depend in a nonlinear way on both first-period and second-period earnings including an

present when considering only the asymptotic position (Diamond, 1980a). This is similar to the difference between the golden rule and the modified golden rule.

[42] The standard optimal tax analysis begins with a set of allowable tax structures and optimizes the tax rates in the allowable structure. The mechanism design approach only rules out taxes that are assumed to require information that the government does not have. Thus, taxing skills is ruled out by the assumption that skills can not be directly inferred from the available information on earnings (without information on hours worked). Beyond this constraint, there are no further restrictions, allowing complex structures that might be assumed as unavailable for being too complex in an optimal tax setup. That is, individuals choose from the allowable set of complete lifetime consumption and earnings levels. From the marginal utilities at the chosen point, one can infer the wedge, the implicit marginal tax rate.

[43] They assume that there is zero interest rate and zero utility discount rate. Thus we can not map the implicit marginal tax on second period consumption (on the savings level), which ranges from .01 to .05, into a tax on capital income.

[44] These simulation studies and the theoretical results discussed have modeled labor supply with only an intensive margin (with a smooth response of labor supply to taxes) and have been primarily focused on marginal tax rates. In contrast, with an important extensive margin (lumpy decisions whether to work or not), average tax rates matter and results on tax rates differ. See, e. g., Choné and Laroque, 2001, 2006, Diamond, 1980b, Saez, 2002c for the case of personal incomes or Griffith and Devereux, 2002 for the case of multinational corporations.

interaction term.[45] Once one envisions modeling longer lives, this degree of interaction becomes implausible to implement in a general form.[46]

The Atkinson-Stiglitz theorem assumes that individuals are able to solve the complex choice problem of how much to earn in each period and the tax collector and legislature are able to cope with setting up and enforcing such a complex structure. These assumptions are problematic and, in practice, the taxation of labor income in a year is usually dependent only on what happens that year, with some exceptions involving averaging over a relatively short number of years.[47] So it is natural to consider the issue of what happens to the Atkinson-Stiglitz theorem in the context of a limited tax structure that resembles those commonly used. As far as we are aware, this problem has received little attention with a heterogeneous population.[48] Weinzierl (2007) has done simulations contrasting labour income taxation that is the same for everyone each period with labour

[45] The theorem needs to allow any function giving the PDV of lifetime taxes as a function of earnings in both periods, $T\left[z_1, z_2\right]$. Thus it is not generally the case that this involves simply adding separate tax functions each period, $T\left[z_1, z_2\right] \neq T_1\left[z_1\right] + R^{-1}T_2\left[z_2\right]$. Framing the problem in terms of a PDV of taxes fits with a restriction that everyone has the same safe rate of return on savings. Otherwise we would also track capital income to see the impact of the timing of tax collection on different individuals.

[46] One strand of the literature has explored assumptions under which the optimum can be implemented with tax structures that are not so complex. These findings arise in models that limit worker heterogeneity greatly. Thus they are an interesting starting place for exploring results as the population is made more diverse, but do not seem to lead directly to policy at present. For example, Golosov and Tsyvinski (2006) examine a role for asset testing, which would be interesting to explore in a more diverse model where asset testing can improve the allocation but does not achieve the mechanism design optimum. Asset testing for access to programs for the poor is widespread even though general taxation of wealth is not. On use of the latter, see Albanesi and Sleet (2006) and Kocherlakota (2005).

[47] It is common in public pension systems to base benefits on a long or full history of earnings records. In contrast to what is needed for mechanism design taxation, basic pension benefit formulas are usually fairly simple, although there is often complexity in special rules.

[48] Erosa and Gervais (2002) have examined the most efficient taxation of a representative consumer (Ramsey taxation) with intertemporally additive preferences in an OLG setting. If the utility discount rate differs from the real discount rate, individuals will choose non-constant age profiles in both consumption and earnings, even if period preferences are additive and the same over time and the wage rate is the same over time. Thus the optimal age-dependent taxes on consumption and earnings are not uniform over time, resulting in nonzero implicit taxation of savings. They also consider optimal taxes that are constrained to be uniform for workers of different ages. It remains the case that the taxation or subsidization of savings is then generally part of such an optimization.

Gaube (2007) examined the difference between general and period tax functions. He did not consider taxing capital income, but showed that the one-period result of a zero marginal tax rate at a finite top of the earnings distribution, which applies to the highest earner with general taxation, does not apply to the two-period model with separate taxation each period when there are income effects on labor supply since additional earnings in one period would lower earnings, and so tax revenues in the other period.

income taxation that can vary with the age of the worker.[49] (The issue of age-dependent earnings taxes is discussed in Part V.) While the paper only reports results for the case without a capital income tax, it does mention a similar calculation for a capital income tax of 15 percent. In personal communication, Weinzierl has reported that social welfare is slightly higher with a 15 percent capital income tax than with a zero tax in both cases – uniform and age-dependent labour income taxation. Weinzierl's model has no physical capital - the benefit of the capital tax in his analysis is that it discourages the use of saving to exploit the redistributive design of the tax system, as discussed above. Thus there is no presumption of the optimality of zero taxation of savings in general, although evidence on the desired structure of taxation with a diverse population and general earnings taxation in each period is very limited.

We have focused on the gap between MRS and MRT for consumption over time, referred to as a wedge, in this case the intertemporal consumption wedge. We have found circumstances in an economy such that this wedge should not be zero, as it is if the Atkinson-Stiglitz theorem holds. There is a similar wedge to consider between earnings in different periods. The presence of non-constant taxation on earnings in the two periods implies that a difference between MRS and MRT for earning in period one relative to earning in period two. If the disutility of labor is a power function[50] and everyone has the same age-wage rate profile, then there should not be an intertemporal earnings wedge (Werning, 2005). But if those with higher earnings have steeper age-earnings profiles, as appears to be the case on average, then the marginal taxes on earnings should rise with age and there should be a wedge on the implicit savings done by increasing early earnings and decreasing later ones, consumption held constant (Diamond, 2007). Taxing consumption implies no tax distortion between earnings in different years. While this does not appear to be part of an optimal plan, desirable aspects of this wedge have not received much attention.

[49] Allowing age-dependent labour income taxation in a two-period OLG model would involve two separate tax functions, $T_1[z_1]$ and $T_2[z_2]$, rather than the same tax function each year, $T[z_1]$ and $T[z_2]$.

[50] A power function is a constant times the variable raised to a power - ax^b.

The models discussed above had perfect capital markets – no borrowing constraints. But borrowing constraints are relevant for tax policy, providing another reason for positive capital income taxation in the presence of taxes on labour income that do not vary with age (Hubbard and Judd, 1986).

In the models reviewed above, the wage rates in the two periods are parameters for each worker. It is clear that later earnings depend on both education and earlier work decisions. The costs coming from efforts to increase future earnings come from leisure, foregone earnings, and expenditures. Some spending, such as tuition, is clearly linked to education and referred to as verifiable spending (although the mix of consumption and investment in an individual's education experience is not verifiable). Other spending, such as higher living costs while at school, are hard to distinguish from consumption spending and are referred to as non-verifiable spending. With constant tax rates on labour income, there would be no implicit tax on the foregone earnings portion of the investment to increase future earnings. With progressive labour income taxes and a rising age-earnings curve, there would be such an implicit tax. Verifiable spending, such as tuition, could be directly subsidized (and widely is). The optimal degree of subsidy depends on the effects on atemporal choices as well as the intertemporal human capital decision, and so may not be set optimally from the narrow perspective of human capital investment. Non-verifiable spending involves goods that also have consumption uses and so can not be subsidized without distorting other consumption decisions. The literature has considered models with no subsidy of non-verifiable spending and full subsidy of verifiable spending with a focus on education. Bovenberg and Jacobs (2005b) consider a three-period model of education, work, and retirement. After showing the desirability of taxing capital income despite the preference assumptions of the Atkinson-Stiglitz theorem, they calibrate the model and conclude that the optimal linear capital income tax rate approaches the optimal linear labour income tax rate. While the rejection of the optimality of a zero tax seems likely to be robust, it would be interesting to see a

calibrated calculation in a setting with more periods and thus on-the-job training as well as formal education.[51]

C. Additional issues:

income shifting, taxing total income, general equilibrium effects, initial wealth

Standard modeling assumes perfect observation of capital and labour incomes. This omits issues of tax evasion (Allingham and Sandmo, 1972, Sandmo. 1981, 2005, Slemrod and Yitzhaki, 2002) and the ability of some workers, particularly the self-employed, to legally transform labour income into capital income (and vice versa). Pirttilä and Selin (2007) found significant shifts of labour income to capital income among the self-employed after the 1993 Finnish tax reform to a dual income tax with a lower rate on capital income.[52] On a more widespread basis, labour effort devoted to earning a higher return on savings also represents a shifting from labour income to capital income. Christiansen and Tuomala (2007) examine a model with costly (but legal) conversion of labour income into capital income. Despite preferences that would result in a zero tax on capital income in the absence of the ability to shift income, they find a positive tax on capital income. As noted below, the Chamley-Judd result of zero capital income taxation also does not hold in a model with an inability to distinguish between entrepreneurial labour income and capital income.[53]

Consideration of income shifting supports marginal taxes on capital income that are higher for people facing higher marginal taxes on labour income. Indeed, taxing total income annually would avoid this issue (apart from the greater possibility of tax deferral with capital income). Apart from this consideration, there is no apparent reason why an

[51] Additional studies with two-period models, with education in the first and earnings in the second period, relate optimal incentives to the mix of opportunity costs and out-of-pocket costs (Hamilton, 1987, Bovenberg and Jacobs, 2005a). On the link between the taxation of financial capital income and the return to human capital see Nielsen and Sørensen, 1997.

[52] Gordon and Slemrod (1998) have argued that a large part of the response observable in the tax return was due to income shifting between the corporate sector and the individual sector.

[53] Income shifting is also an issue in the conversion of labour income into corporate income, which has received attention in the literature on the corporate tax (e. g., Gordon and MacKie-Mason, 1995).

optimal tax calculation would find an optimum with the same marginal tax rates on capital and labour incomes. The discussion below, accompanying Table 1, points out how different the tax wedges are from taxing labour and capital incomes at the same rates. Without extensive analysis of elasticities, one cannot make conclusions about optimal rates in light of this pattern of tax wedges. However, we see no reason to expect that studies would generate results close to uniformity in the relative taxation of the two types of income. This is particularly the case with capital income after retirement, for which the Atkinson-Stiglitz theorem has more relevance because of the absence of relevant uncertainty about earnings abilities. Indeed, we are not aware of any optimal tax study calling for taxing total income.

In addition to uncertainty about future earnings, there is uncertainty about future preferences. There may be uncertainty about how much consumption will be enjoyed when older – either from an inability to fully appreciate future preferences[54] or from shocks that are not fully insured – such as health shocks or spending shocks (medical or legal expenses) or an inheritance.[55] One example of significant uncertainty is in the length of life. Moreover, longer expected lives are positively correlated with earnings abilities (e. g., as proxied by education) for both men and women. Modeling this interaction would need to explore the use of and properties of the annuities market. In the absence of a range of models to draw from, it is not clear what sign to put on the optimal taxation of savings from this consideration.

Following the setup in Mirrlees (1971), the relative wage rates of different workers are exogenous in the Atkinson-Stiglitz theorem, although the absolute wage rates can be endogenous. Naito (1999) has shown that with endogenous relative wage rates of skilled and unskilled workers, the Atkinson-Stiglitz theorem does not hold.[56] If the production of consumption for period one makes different relative uses of skilled and unskilled labour than production of consumption for period two, then a change in the

[54] See, for example Gilbert, 2006.
[55] Another source of uncertainty comes from uncertain future relative prices. This is present even with savings in real assets based on a price index that is not precisely the right one for a given individual.
[56] This is similar to the failure of the Diamond-Mirrlees (1971) aggregate efficiency theorem with restrictions on the taxation of some commodities, for example, when different commodities must be taxed at the same rates (Diamond, 1973).

savings rate alters the relative demands for the two types of labour, changing their relative wages. This is an alternative approach to redistribution, one that is in principle a useful supplement to progressive earnings taxes. That is, there is an aggregate production set involving first-period consumption, second-period consumption, skilled labour and unskilled labour. If, by shifting consumption demand between periods, one can shift relative wages, then the incentive compatibility constraint can be weakened, breaking the dominance of the earnings tax over the non-proportional taxation of consumption. Empirical work supports the finding that increased capital (in the form of equipment) raises skilled relative to unskilled wages (Krusell et al., 2000), supporting taxation of capital income, although the importance and magnitude of this consideration are unclear.

The models considered above have variation in the population in earnings ability, and sometimes in preferences, but not in wealth at the start of the first period. With variation in initial wealth holdings and an ability to tax initial wealth, the optimum may call for full taxation of initial wealth, particularly when higher wealth is associated with higher earnings abilities. If immediate taxation of initial wealth is ruled out, the presence of capital at the start of the first period, which can earn a return when carried to the second period, can also prevent the optimality of the non-taxation of capital income if there are no fairness issues further limiting the desirability of taxation of initial wealth. As a modeling issue, one needs to ask where such wealth came from. Presumably gifts and inheritances are a major source. But since these might themselves be taxed and since gifts and bequests might be influenced by future taxation of capital income, a better treatment of this issue would be embedded in an OLG model that incorporates the different ways that people think about bequests.[57] A similar issue arises in tax reform given past savings under a previous tax regime.

D. Overlapping Generations (OLG) models

The analysis above considered the intertemporal dimension of direct taxation for a single cohort. A natural question is the impact of the reality of overlapping generations

[57] See, for example, Boadway, Marchand, and Pestieau, 2000, Cremer, Pestieau, and Rochet, 2001. That optimal taxation depends on bequest motivation is brought out in Cremer and Pestieau, 2003.

on such analyses. The OLG literature models choice by successive cohorts of workers, with the basic model having no bequests at all. There are two key aspects of the connection between analysis for a single cohort and OLG analysis. One is the government's role in affecting the lifetime budget constraints of different cohorts (and thus the aggregate capital available to different cohorts). The other is the extent to which taxes can vary with age and so with cohort in a single period.

If the government is free to use public debt and public assets as part of intergenerational redistribution, thereby altering national capital, and if taxes are age-dependent, then a full optimization in the OLG model can be divided to include suboptimizations for each cohort, as above (Diamond, 1973).[58] That is, from the intergenerational optimization there is a constraint on the net contribution to national capital from each cohort. Using this net contribution as a constraint on optimization of taxes for a cohort, then the type of optimizations we have analyzed above hold in the basic case where there is no direct concern about relative prices. The analyses with a concern about relative prices, particularly a concern about relative wages, do not generally have this full separation. Presumably our analysis above remains strongly suggestive. Other links would naturally arise, particularly related t education, since parents look after children.

Thus, with the assumptions on preferences that are sufficient for the Atkinson-Stiglitz theorem for a single cohort, the theorem still holds in the setting of overlapping generations with no constraints on government debt policy and on age- (and so cohort-) specific taxes. The reasons for the inapplicability of the theorem discussed above carry over to the OLG setting. A separate issue is whether the government does not adjust debt policies but then uses tax policies to affect capital stocks instead. That is, if the government follows policies, such as too much debt, that reduce capital below optimal levels, then tax policies to increase individual savings may become more attractive as a substitute (third-best) policy (Atkinson and Sandmo, 1980). Such analysis is likely to be sensitive to the way the determination of government debt policy is modeled. It is not

[58] If the government wants to give higher consumption to an early cohort, financed by lower consumption for later cohorts, it can do this in a pay-as-you-go pension system, or by borrowing to finance transfers to the early cohort and financing the debt from taxes on later cohorts.

clear how best to describe the determinants of UK debt/public capital policy, whether such political behavior is best thought of as stable over time, and how robust any findings about tax policy would be. There is also a natural suspicion that such third-best arguments can be a cover for other motives.

In practice, taxes do not vary (much) by cohort – that is they are period-specific rather than age-and-period specific. Above, we briefly discussed the issue of taxes for a single cohort that did not vary with age. The same issues arise with period-specific taxes affecting people of different ages. Thus recognition of the OLG setting emphasizes the importance of this consideration and of the possibilities in age-dependent taxes.

E. Models with infinite horizon agents

These OLG models have an infinite horizon for the economy, but have no direct links across the finite-lived cohorts. Redistribution across cohorts (with its induced change in the capital stock) is then important for capital growth and can be done without having to distort individual savings decisions. Conversely, distorting individual savings decisions can be done without necessarily changing aggregate capital by also redistributing across cohorts. In contrast, if agents optimize over an infinite future, altering the timing of their consumption does require distorting individual savings decisions. That is, a key implication of infinite horizon agents is that a shift of tax collection over time, which would influence capital accumulation when the shift involves different cohorts in an OLG model, is fully offset for infinite horizon agents. Thus the taxation of capital income plays a role in intertemporal allocation that is stronger than in the OLG model because of the lack of effect of this intertemporal redistribution policy tool. Infinitely-lived agents are naturally interpreted as doing optimization for a dynasty, and so making bequest decisions. Moreover, recognizing overlapping generations as opposed to sequential ones as part of the infinite horizon planning, the agents are also adjusting incomes of contemporaneous members of a single dynasty.[59]

[59] The empirical evidence on the consumption patterns of parents and adult children alive at the same time is strongly contradictory of the idea that people typically behave as if there were a single dynastic utility

The central finding in this literature, due to Chamley (1986) and Judd (1985), is the optimality of zero taxation of capital income in the long-run. We begin by considering the intuition generally put forth for this result. After discussing its relevance and considering generalizations that imply that optimal taxation of capital income is not zero, we consider a generalization of the basic result in Judd (1999).

Above, we have examined the relationship between the intertemporal consumption MRS and intertemporal MRT that would be optimal in different settings. We start this discussion by noting the relationship between them if there is a constant tax rate on capital income. Assuming an interest rate (marginal product of capital), r, which is constant over time, then a unit of consumption today can be converted into $(1+r)^T$ units of consumption T periods from now (in period T+1, if we denote today by period 1). Thus the MRT_{1T+1} is $(1+r)^T$. If an investor is subject to a tax at rate τ on capital income, then the investor can convert one unit of consumption today into $(1+(1-\tau)r)^T$ units of his own consumption after T periods. The ratio between the MRS and MRT between consumption today and consumption T periods from now is $\{(1+(1-\tau)r)/(1+r)\}^T$. This gives the fraction of the available social return that goes to the investor. With a positive rate of tax this expression goes to zero as T goes to infinity. And it gets small for long, finite time spans. Some examples, are given in Table 1.

Table 1. Ratio of MRS to MRT - $\{(1+(1-\tau)r)/(1+r)\}^T$.

T	r =.05, τ =.15	r =.10, τ =.15	r =.05, τ =.30	r =.10, τ =.30
1	.993	.986	.985	.973
10	.931	.872	.866	.758
20	.866	.760	.750	.575
40	.751	.577	.562	.331

function being jointly maximized. Moreover, taking this literally and recognizing marriage (which links dynasties to each other) leads to absurdities (Bernheim and Bagwell, 1988).

60	.650	.439	.422	.190
80	.564	.333	.316	.109

Comparing the table to a tax on labour earnings makes several points. A 30 percent tax on earnings puts a 30 percent wedge between contemporaneous earnings and consumption. A 30 percent tax on capital income puts only a 3 percent wedge between consumption today and consumption in a year (when the rate of return is 10 percent). But it puts a 67 percent wedge between consumption today and consumption in 40 years. The difference comes from the shifting relative importance of principal and interest in the financing of future consumption as we look further into the future. The table makes clear that the intertemporal consumption tax wedge depends on whether nominal or real incomes are being taxed. This table raises the issue of how far into the future people are thinking when making consumption-saving decisions. It suggests that if people have a long enough horizon, capital income taxation that impacts distant consumption will be inefficient, a suggestion we examine in detail. And it points to potential welfare gains from tax-favored retirement savings, since that saving tends to be for longer times.

When agents have long horizons, modeling their current decision-making using an infinite horizon model can be mathematically more tractable than a long finite horizon, while doing little violence to conclusions from the analysis that relate to current behavior. However, when considering the evolution of an economy over time, a model with a fixed number of infinitely lived agents behaves very differently from an OLG model, even one with long lives.[60] And that can matter for drawing conclusions about incentives that matter primarily for future behaviors, such as capital income taxes in the distant future.

Let us start with the basic interpretation of the model before turning to detailed modeling assumptions. In the standard OLG model, individuals have no concern for the future after their deaths and leave no bequests. This is empirically inaccurate - most people leave some bequests and we think that some people adjust earnings and/or savings

[60] Immigration of new dynasties makes a model with infinite-lived agents have some of the properties of a finite-lived OLG model (Weil, 1989).

in light of planned gifts and bequests.[61] Results vary in models that extend the basic OLG model for bequests, depending on how bequest decisions are modeled. Models with "accidental bequests" because of incomplete insurance/annuitization and models with planned bequests arising from motivations that can influence earlier decisions generate different positive and normative tax implications. [62] Empirically, how important bequest considerations are for behavior is unclear and widely varying in the population. A further complication in interpreting behavior as dynastic is the sizable tendency to make charitable bequests. Also key to further analysis is how to form a social welfare function since counting both the utility of a donor and the utility of a donee in a social welfare function has implications that can be questioned as being normatively unattractive.

In contrast, the standard infinite horizon agent model is viewed as a dynasty model with incorporation of future utilities in the decision-making of earlier cohorts and a normative evaluation of the utilities of consumption of each generation exactly as they are viewed by the existing generation. This is typically done as if there were only one generation alive at a time and lasting only a single period, rather than the multiple overlapping generations that are actually present. In terms of the normative issue raised above, this can be viewed as counting the utility of the donor and ignoring the utility of the donee, and is one way to approach the concern about overweighing the consequences of concern for others.[63]

---

[61] Part of the debate on the importance of intergenerational links for the evolution of the capital stock relates to the treatment of the financing of education and other gifts that occur well before the time of a parent's expected death. This is ignored in this discussion which focuses on the transfer of financial wealth at death or at a time when remaining life expectancy is small.

[62] The role of saving for bequests appears to be diverse in the population and unclear (Hurd, 1987). As an example of the importance of motivation, if all bequests are accidental from incomplete annuitization and also unobservable, then there is a case for capital income taxation when assumed preferences and technology would have a zero tax rate be optimal without the bequests (Boadway, Marchand, and Pestieau. 2000). On the other hand, with the same assumptions, if bequests are given from a utility motivation and if the utility motivation is fully respected in the government objective function, then the optimal tax on capital may be positive or negative (Cremer, Pestieau, and Rochet. 2003).

[63] Farhi and Werning, 2005, consider the case of respecting individual dynastic preferences and also giving weight to the dynastic preferences of later generations. As in Kaplow (1995) the thrust of such modeling is to subsidize gifts and bequests since they benefit both the donor and the donee. The results would change if the social welfare function treated dynastic concerns differently from the utility of own-consumption in the social welfare function, an issue considered in the context of charitable donations in Diamond (2006).

It is useful to complement OLG models that unrealistically ignore bequests with models that give bequests a larger role in decision-making than they have in reality – at least until we have better empirics and analytics about bequests. So an evaluation of the role of other assumptions in reaching the Chamley-Judd no-capital-income-taxation conclusion is appropriate. This widely cited result is that when such an economy is in a steady state, there should be no taxation of capital income (with a linked convergence result that the tax rate converges to zero as the economy converges to a steady state). As Chamley (1986) explained: "The main property of the model which is used in the proof is the equality between the private and social discount rate in the long run." (page 608) and, in the altruistic dynasty interpretation: "When the social planner uses the same discount rate for the future life cyclers as the discount rate applied in the altruistic families, the long-run tax rate on capital income is zero. This property … requires that individuals not be constrained at a corner solution for their bequest." (page 613) or "This assumes that the social planner and the individuals use the same relative utility weights for intergenerational transfers." (page 619). Once the weights differ, then the result changes.

As with the Atkinson-Stiglitz result, a key question is how robust the conclusion is to realistic changes in the model. We reach the same conclusion in this case as in the earlier analysis – the finding is not robust for policy purposes.

In the single-cohort model, Naito (1999) has shown that endogeneity of relative wages, together with a uniform earnings tax function, contradicts the optimality of zero capital income taxes when relative wages can be influenced, even with the Atkinson-Stiglitz separability assumptions. Correia, (1996) has shown a related result in the infinite horizon model with endogenous relative wages. She assumed two kinds of labour and an inability to tax one kind. The adjustment of capital to offset the absence of taxation of this labour results in a long-run equilibrium with non-zero taxation of capital, with the sign of the tax depending on the details of the technology. A similar result holds if the two types of labour must be taxed the same (and capital affects relative wages). A directly relevant result holds if one of the two types of labour must be taxed the same as capital income is taxed, reflecting an inability to tell apart capital and some labour incomes, which is relevant not only for the self-employed but also in the case of

successful corporations with large maintained control by the founders, as with Microsoft or Google. In this case the inability to distinguish between entrepreneurial compensation and the return to capital implies that capital income should be subject to a positive tax (Reis, 2007).

Also, as in the one-cohort model, uncertainty about the future earnings of those alive and already working as well as about the earnings of those not yet in the labour market or not yet born implies the optimality of positive taxation of capital income (Golosov, Kocherlakota and Tsyvinski, 2003).[64] Aiyagari (1995) and Chamley (2001) considered borrowing constrained agents in an uncertainty setting. In these models, precautionary savings are high in anticipation of future borrowing constraints, which implies that a positive capital tax is welfare improving in the standard setup.[65]

Additional considerations arise when there is human capital as well as physical capital in an infinite horizon model. In the presence of both physical and human capital, labour is supplied jointly with human capital, which means that a positive labour tax is also a tax on human capital if its cost is not just foregone earnings and subsidizable spending (such as tuition). In this setup, it is optimal to converge to zero capital and zero labour taxes (Jones, Manuelli and Rossi, 1997) unless human capital is observable. If a direct subsidy on human capital is available, then it is optimal to have positive labour taxes in the long run accompanied by a subsidy on human capital and zero taxes on physical capital (Judd, 1999). The result with unobservable human capital suggests that the accumulation of sufficient government resources, relative to expenditures, is a key part of the result on the optimality of asymptotic zero taxation. Thus, at a time of tax reform from a non-optimal tax structure, it is not clear whether the result that long run taxation of capital should stop is a call for increasing or decreasing the current taxation of capital income. Indeed the models call for maximal taxation on existing capital since it is inelastically available. Taxation of existing wealth is discussed in VII.B.

[64] Analysis of aggregate uncertainty that affects all earnings possibilities proportionally is quite different. See Golosov, Tsyvinski and Werning, 2007.

[65] Using a different setup, Chamley (2001) has an example in which randomness is in the timing of future incomes, with the outcome learned ahead of time, giving an advantage to subsidizing capital income rather than taxing it.

Another source of concern about the results in existing models is that the models assume that the tax on capital income is linear. Saez (2002a) has examined a linear tax with an exemption, as opposed to a tax linear from the origin. Asymptotically no one is paying the capital income tax, as initial wealths above the exemption level decline to the exemption level – with everyone having the same utility discount rate the before-tax interest rate is driven to the highest discount rate in a steady-state, implying a lower after-tax return if there are dynasties with wealth above the exemption level and thus wealth that grows more slowly than the economy. But the tax has served to raise revenue from those with the highest wealth, reducing their wealth to the exemption level - an exemption level that is finite (as opposed to infinite which would be equivalent to no tax) is part of an optimum.

Note that in the long run of the usual models, each period is exactly the same for a dynasty. Recognizing that the dynasties are a collection of successive individuals makes all of the issues considered above for a single cohort relevant in this model as well. For example, earnings are uncertain and the average age-earnings profile is not flat. These observations raise similar issues for capital income taxation as they do in the single-cohort and OLG models. The analysis of Judd (1999) is interesting for addressing this issue. Judd allows greater generality in the evolution of the economy and obtains the result that the average capital income tax tends to zero even if it is not zero in any period.[66] When the model is interpreted as each generation living for a single period, a tax on capital income is equivalent to a tax on bequests. Once individuals live longer than a single period, then one can distinguish between a tax on capital income and a tax on bequests. This point has been made by Chamley (1986, page 613) "If a specific tax can be implemented on the interest income of savings used for life-cycle consumption, its rate is in general different from zero." To preserve a long-run convergence to a zero average tax on capital income while distinguishing between capital income and bequest taxes, if one were taxing capital income during lifetimes, as argued for above, then one would be subsidizing bequests. Such a starting place for analysis focuses attention,

[66] For example, assume the period utility functions are the same in all even-numbered years and all odd-numbered years, but different across adjoining years. Then there will be alternating taxes that would show long run zero taxation across pairs of years (consistent with taxation being zero on average in Judd, 1999).

appropriately, on the analysis of bequest motives (and their heterogeneity). The relevance of long-run results from this class of models depends critically on the degree of realism of the underlying model of bequest behavior. Yet, as noted above how important bequest considerations are for behavior is unclear and widely varying in the population.[67]

Thus we conclude that the Chamley-Judd result that there should be no taxation of capital income in the long run is not a good basis for policy. Nevertheless the issue remains of the compounding of taxation of capital income resulting in a growing tax wedge the longer the horizon for decision-making – a point also made in models with finite lives of many periods. This is suggestive of a possible role for capital income taxation that varies with the age of the saver and/or with the time lapse between savings and later consumption (as with tax-favored retirement savings). The role of capital income taxation when earnings are uncertain particularly suggests that rules might well be different for those at ages when workers are mostly retired.

[67] People give inter vivos gifts as well as bequests. Given the tax advantage in the US for inter vivos gifts relative to bequests, the dynasty model would imply far more use of inter vivos gifts than is the case (Poterba, 1998).

Part III. Taxing consumption

Part II analyzed the extent to which capital income should be taxed in the presence of taxation of labour income. While the starting place was the Atkinson-Stiglitz theorem giving conditions under which capital income should not be taxed, realistic extensions of the model support the taxation of capital income. There was some support for marginal taxation of capital income at rates that varied with the marginal rate on labour income, as opposed to the linear taxation in the Nordic dual income tax model. Part of the case for the Nordic model is the political argument that base widening is more readily accepted along with lowering the tax rate on capital income – an important point given the efficiency costs of differential taxation of different sources of capital income. Thus, the conclusion of Part II was that there should be a wedge between the intertemporal consumption MRS and MRT. While not analyzed in detail, the models in Part II did generally also involve a wedge between the intertemporal earnings MRS and MRT.

In this part, we consider the properties of the annual taxation of consumption, rather than the annual taxation of earnings. The recommendation of the Meade Report was for annual progressive taxation of consumption, together with annual taxation of wealth, with particular attention to inheritances.[68] As in Part II, we begin with analysis in a setting of only safe investments – the same rate of return available to everyone. After comparing linear taxation of consumption and earnings, including a discussion of transition, we briefly mention the difference resulting from progressive taxation. Part IV examines issues raised by stochastic returns to investment.

A. Linear taxation

Consider a worker whose entire life is under the same linear tax on earnings. The PDV of the tax paid is then $t_z \sum_{s=1}^{S} z_s \left(1+r\right)^{1-s}$, where t_z is the tax rate on earnings, z_s is

[68] The Meade report considered savings in separate accounts with different tax treatment.

earnings in year s and earnings stop after S years. If the worker neither receives nor gives gifts or bequests[69] then lifetime consumption satisfies the lifetime budget constraint, $\sum_{s=1}^{S'} c_s (1+r)^{1-s} = (1-t_z) \sum_{s=1}^{S} z_s (1+r)^{1-s}$, where S' is the length of life. With a tax, t_c on consumption, and no tax on earnings, the lifetime budget constraint is $(1+t_c) \sum_{s=1}^{S'} c_s (1+r)^{1-s} = \sum_{s=1}^{S} z_s (1+r)^{1-s}$, and the taxes paid are $t_c \sum_{s=1}^{S'} c_s (1+r)^{1-s}$. Thus the systems are equivalent on a PDV basis for each member of such a cohort – for each linear earnings tax rate there is a linear consumption tax rate that results in the same budget sets (and so the same earnings and consumption decisions) and same PDV of tax revenues.[70] The matching tax rates satisfy $(1-t_z)(1+t_c) = 1$.[71]

In order for equilibrium to be unchanged by this matched change from an earnings tax to a consumption tax, we need government behavior to also be unchanged. Since the timing of consumption does not match the timing of earnings, the timing of tax revenue changes. While there is some borrowing that permits consumption to exceed earnings for young workers,[72] saving for retirement is the larger element, so that, with consumption taxation, on average individuals would pay taxes later in their lives and so would save more, buying bonds in anticipation of future taxes. In turn, this increased demand for bonds would permit the government to do its financing for unchanged spending as part of equilibrium without altering the interest rate.[73] Whether this is what would actually happen depends on how the government actually responds to collecting

[69] To incorporate bequests and inheritances we would also want to incorporate estate or inheritance taxes.
[70] Below we note the circumstances where equivalence holds with stochastic returns to savings.
[71] If there are binding borrowing constraints limiting consumption to what can be financed by contemporaneous earnings, the equivalence carries over nevertheless. The perfect capital market assumed in this budget constraint ignores differences between borrowing costs and lending returns, which would make the timing of taxes matter to individuals.
[72] Presumably house purchases would not be fully taxed as consumption spending, but rather converted into a flow of later taxation.
[73] Since consumption is larger than earnings because of interest income, the delay in taxes is offset by this source of consumption. In an OLG setting this is combined with differences across cohorts in both size and level of age-earnings trajectories. As long as the rate of interest exceeds the rate of aggregate earnings growth, this difference does not matter on an aggregate PDV basis for all cohorts living fully under one system or the other.

revenue later with a consumption tax rather than earlier with an earnings tax. If government spending changed, so too would the equilibrium.

To see how this plays out over time, consider a change from an earnings tax to a consumption tax in an OLG setting. Assume the transition rules kept taxes the same for cohorts taxed under the old system, so the taxes only involve the new generations and thus do not involve redistribution across generations. Then, after a period with only very young workers taxed, which we ignore, there is a period dominated by savings for retirement, implying a drop in tax revenue as consumption is less than earnings.[74] Once the new steady state is reached, which now includes consumption by retirees, tax revenue exceeds what it would have been under an earnings tax, by an amount matched by the interest cost of government borrowing because of the lower tax revenue in the initial periods. If the government is making its tax and spending decisions based on a long horizon, then the situation is unchanged. However, if the government spends its revenue each period (pay-as-you-go for the full budget), then government spending is lower in the early periods and higher in the later periods as a result of the change to consumption taxation. Adaptation of the economy to this pattern (assuming government spending is consumption, not investment) implies a rise in the aggregate capital stock from having less government consumption earlier, private consumption and output held constant. For private consumption to remain constant generally, government consumption needs to be separable from private consumption in individual preferences. We are also ignoring any change induced by changes in the wage, interest rate and relative prices of consumer goods.

How does this difference in timing of government consumption matter for evaluation of the tax change? If one were to look only at the new steady state, one would find higher capital with consumption taxation, and so higher output and one might conclude (by erroneous logic) that the change was beneficial, whether it was or not. Proper policy evaluation should look at the entire path of an economy and not just the steady state. Doing that, one would need to evaluate the change in the pattern of

[74] Since workers may borrow early in their careers, this is really referring to a time period with positive savings for retirement consumption. An uncomplicated picture can be seen in a two-period OLG model, with one period of work and two of consumption.

government consumption spending (more earlier, less later) as the primary basis for evaluation. The increase in capital from changed timing of government consumption and tax revenue is merely an efficient equilibrium adaptation to the change in the government consumption pattern, not an appropriate source of a positive evaluation.

The political economy of how much borrowing a government does is important and controversial, making it unlikely that some specific model of political outcomes implicit in a particular budget balance constraint will match actual behavior. Governments generally do not follow such a simple behavioral rule as annual budget balance on average or on the margin. Until we have a better, empirically-based understanding of government budgetary practices, an adjustment for government spending behavior is somewhat speculative. For countries like the UK, the ability to borrow, to reduce the public debt, and to save is real. Debt to GDP ratios have varied greatly over time. Examining policy in a setting with a single PDV government budget constraint is in keeping with looking at how governments ought to consider policy.

Note that commenting positively on government policy on the basis of an induced delay in government spending involves saying to the government that since it will otherwise spend relatively too much in the short run (and too little in the long run), the government should choose one tax over another because the choice will lead the government itself to do less spending in the short run (and the reverse later). Legislative process rules that affect political outcomes seem very important. And adjustment of economic advice based on a perception of actual government behavior, given the advice, also seems important. Yet we are reluctant to base too much on an oversimplified model of the influence of the timing of revenues on spending. Note that this is not a setting of permanently lower revenues but of lower revenues followed by higher revenues. While governments are slow to adapt to perceptions of such a future, anticipatory adjustments in public pension systems that we have observed over the last two or three decades suggest that some degree of forward looking planning does indeed happen.

A tax on consumption can be collected as a tax directly on consumption, as with a VAT, or by taxing earnings less net active savings.[75] The latter permits progressive tax rates, for example by use of annual exemptions.[76] The equivalence for new cohorts between taxing earnings and taxing consumption does not extend from a linear setting to a nonlinear annual tax since neither earnings nor consumption are generally constant over time.[77] That is, variations in earnings and in consumption might move above and below break points between marginal rates (for example above and below the exempt amounts) in different ways. This can happen in certainty models unless the utility discount rate matches the rate of return to savings and can happen with uncertain earnings opportunities.

Note that there is no intertemporal consumption tax wedge and no intertemporal earnings tax wedge with linear taxation of either earnings or consumption. With progressive annual consumption taxes there is still no intertemporal earnings wedge. If the age-consumption profile with optimal taxes is rising more often than falling among workers (as is empirically the case with existing taxes), then they would more often generate a positive intertemporal consumption tax wedge. How these two patterns of wedges (on consumption and on earnings) might relate to a desirable pattern has not appeared in the literature we have seen.

B. Transition

[75] Active savings are savings from earnings. Thus earnings minus net avtice savings equals income minus net savings.

[76] This point is drawn out in Hall and Rabushka, 2007, which proposes collection through a VAT combined with adminstrative shift of payment responsibility to the employee. It's a VAT with a rebate equal to earnings up to a ceiling. Thus it is a VAT with a zero effective marginal rate below the ceiling.

[77] The equivalence for new cohorts between taxing earnings and taxing consumption extends from a linear setting to a nonlinear setting provided that taxation is based on lifetime earnings and lifetime consumption. That is, lifetime taxes might be $T_c\left[\sum_{s=1}^{S'} c_s \left(1+r\right)^{1-s}\right]$, or $T_z\left[\sum_{s=1}^{S} z_s \left(1+r\right)^{1-s}\right]$, with annual taxes being withheld toward lifetime taxes. It is not clear how those with different realized lifetimes should be taxed relative to each other. Extending this equivalence to include recognition of bequests and inheritances is complicated by the nonlinearity in the tax structure which requires some integration between estate/inheritance and earnings/consumption taxes. We continue to ignore this issue, leaving it for another chapter.

Vickrey (1947) was concerned with the relative treatment by progressive annual taxes of those with constant incomes and those with fluctuating incomes

There is no impact on a generation fully under a new system from a change from a linear earnings tax to the linear consumption tax with the equivalent rate analyzed above. However, a change between the two linear systems may matter for older cohorts who live partially under one system and then under the other, depending on the tax treatment of wealth existing at the initiation of the tax regime.[78] Going from an earnings tax to a VAT will increase taxes on people holding wealth (for later consumption) at the time of change, unless there is an offsetting transition adjustment for the implied taxation of consumption from initial wealth. Thus, without a transition adjustment, this change in tax system represents a tax on initial wealth, which is then a nondistorting tax. Indeed, analyses of change to consumption taxation find that a large part of the reported efficiency gain is from the lump sum nature of the taxation of existing wealth (See, e. g., Auerbach, Kotlikoff and Skinner, 1983, Altig, Auerbach, Kotlikoff, Smetters, and Walliser, 2001). Distributionally, the change hurts those with wealth relative to those without at the time of the change. If the tax rates hold the PDV of revenue across all generations constant, then a primary pattern is a higher lifetime tax on those who are older at the time of the tax change, and a lower tax on others, particularly those not yet born. Normative consideration of such a change requires evaluation of this distribution of tax changes as well as consideration of a change from a system that people were relying on and analysis of whether an unanticipated change results in a behavioral response in light of changed expectations of possible future changes. We touch briefly on this issue below in VII.B.

A different transition issue may arise if the implementation of the tax is through taxing earnings less net active savings. If net active savings are accurately measured then earnings taxation with a savings deduction is equivalent to VAT. However, if net active savings are measured by net deposits into special savings accounts, then accurate measurement of consumption requires measuring net decreases in wealth held outside the accounts insofar as they are used to finance the deposits. With no tracking of outside wealth, transferring initial wealth into the accounts would look like net active savings, resulting in less taxation at the time. Later, withdrawals from the accounts are taxed as

[78] Also relevant is what happens to asset prices, an issue we do not discuss. See, for example, Judd, 2001.

consumption (assuming bequests are treated as consumption). Thus consumption from initial wealth is not taxed in PDV terms, preserving the equivalence with earnings taxation and breaking the equivalence with a VAT.

Part IV. Stochastic rates of return

Many models of optimal taxation assume safe returns to savings. Yet real returns to savings are stochastic. The randomness may be modeled as perfectly correlated across individuals – as would be the case with the risk coming from access to a capital market with stocks and bonds and the same risky portfolio holdings for everyone. However, portfolios vary widely across households. Different people have different beliefs about returns on different assets and access to different information sources and different investment opportunities. And a large fraction of the public holds no stocks at all. Also, not all investments are in market-traded assets.

A. Marketed risks

Taxing consumption rather than taxing total income has been described as exempting the safe rate of return from taxation, but taxing the difference between the realized risky and the safe rates of return the same (e. g., Gentry and Hubbard, 1997, Weisbach, 2005).[79] Similarly, the equivalence between taxing consumption and taxing earnings has been questioned in terms of the taxation of the difference between risky and safe returns (see, e. g., Zodrow, 1995). Evaluation of these issues requires examination of equilibria with different tax structures. Such an evaluation needs to recognize heterogeneity in the population and the behavior of the government, as noted above.

Lying behind the two equivalence views are the analyses of Gordon (1985) and Kaplow (1994) that linear taxation of the difference between risky and safe returns (with full loss offset) has no effects, with the uses of the revenue that they describe. Before turning to their analyses, let us note the lack of direct impact on an individual with a diversified portfolio and access to market transactions on fixed terms. Without taxation of returns, individual would realize a return on his portfolio of $\alpha r + (1-\alpha)\rho = r + (1-\alpha)(\rho - r)$, where α is the fraction of the portfolio invested in a

[79] The bulk of the analysis allows full loss offset, which is not generally the case with income taxes. For discussion of this issue, see Weisbach (2005).

safe asset paying return r and $1-\alpha$ is the fraction of the portfolio invested in a risky asset paying return ρ. With a tax, t, on the difference between risky and safe returns (with full loss offset), the realized after-tax return becomes $r+(1-\alpha')(\rho-r)(1-t)$. By adjusting the portfolio, assuming no binding limit on borrowing or short selling, the investment in risky assets can be increased so that the after-tax returns from the portfolio match the pre-tax returns when there are no taxes. Thus, the investor can obtain exactly the same returns with and without the tax - $r+(1-\alpha')(\rho-r)(1-t)=r+(1-\alpha)(\rho-r)$ when $(1-\alpha')(1-t)=(1-\alpha)$. In order to analyze equilibria with all investors responding in this way, we need to consider the supply of assets and how the government reacts to the (stochastic) revenue it receives from this taxation.

In showing no effect from a tax on the difference between risky and safe returns, Gordon assumes that the tax revenue from each person is returned to that person in a (stochastic) lump sum way.[80] Kaplow's assumptions are equivalent to having the government sell the stochastic tax yields in the market.[81] In both cases, the imposition of the tax and the government's portfolio or lump sum transfer policy has no effect on equilibrium. That is, the consumers do not change their consumption and earnings plans and the government does not change its real expenditures. When taxing the difference between risky and safe returns has no effect at all, then the tax treatment of this source of income is the same for an income tax, an expenditure tax and an earnings tax.

Above, we saw that with only safe investments, taxing consumption (linearly) is equivalent to taxing earnings (linearly), provided there is a perfect capital market with only a safe asset and that government behavior depends on the PDV of tax revenues, not the timing of revenues. There was equivalence in household behavior for tax rates satisfying, $(1-t_z)(1+t_c)=1$. Going from equivalence in household behavior to equivalence in equilibrium required the government to adjust public debt outstanding to

[80] In this case, the investor does not want to change his portfolio since he is also receiving the risky tax revenues.

[81] Thus, when the investor adjusts his portfolio as above, he purchases precisely the portfolio offered by the government as a consequence of the taxes he is paying. Thus the sale of the government portfolio yields no return. If the investor is indifferent at the margin between stocks and bonds, then the marginal value of the difference between stock and bond returns is zero. The marginal valuation equals the price in equilibrium.

offset the change in the timing of tax revenues. If that is done, then there is no change in equilibrium consumptions and earnings from a change to an equivalent tax (for cohorts fully under the new system; that is, assuming adjustment for transition cohorts).

Examining the household choice problem with safe and risky investment opportunities shows the same equivalence as with only safe investments. In order to have equivalence of equilibrium, the government must adjust in response to the change in the timing of revenues and to the presence of a stochastic pattern of government revenues. As with the safe return case, the government needs to adjust its debt and as with the Gordon and Kaplow analyses, it needs to shift the risk to households in a way that matches the risk they held before the taxation of risky returns. If these are done, then there is equivalence of consumption and earnings taxation, because the taxation of the difference between safe and risky returns has no effect on equilibrium. Similarly taxing total income and taxing earnings differ in the taxation of safe returns, not the taxation of the difference between risky and safe returns.

Key to this result is how the government responds to the change in tax revenues from the taxation of the difference between risky and safe returns. The Gordon and Kaplow assumptions, while informative of the workings of the economic mechanisms, are not similar to actual government practice. That suggests modeling a change in taxes, borrowing, and spending that follows practice more closely, along with a change that makes the workings of the model clear. For example, this suggests a comparison of consumption and earnings taxes without accompanying lump sum transfers or marketing of the risks in future tax revenues. Such modeling would involve two complexities – the description of the menu of risky and safe investments available to the economy and the description of how the government does adapt to a change in the risk characteristics of tax revenues. Discussion of this in the literature has contrasted interpretations with different discount rates for the equivalence in government revenue. But the "right" discount rate to use for analysis cannot be assumed but needs to be derived from a model of how the government behaves and what the investment options in the economy are. Presumably this can be done along the lines of analysis of the choice of portfolio for a public pension system (see, e. g., Abel, 2001, Diamond and Geanakoplos, 2003) and the

adjustment of a defined benefit system for different cohorts (see Gollier, 2005). But such analyses have not been done as far as we know. Our presumption is that neither equivalence holds once one recognizes heterogeneity in individual portfolios and government actions that are restricted to issuing safe bonds and adjusting tax rates (on earnings or consumption).[82]

B. Non-marketed risks

With marketed investments, all those making use of the stock market can share in bearing the risks in return and valuation, and modeling assumes that each investor is small relative to the market. While the government spreads risks from tax revenues differently than the market would, particularly over time, a comparison of market and government risk allocations involves the entire economy in both cases. Not all investments are marketed through stock markets. Taxation of the returns to non-marketed investments will matter because of the shift in risk from the single investor (or small number of investors) to the economy as a whole through the government's tax and spending policies. Also, non-marketed risks are not likely to be constant returns to scale. Thus the presence of taxation affects the inframarginal opportunities available to entrepreneurs as well as sharing the risks of those opportunities. This has some similarity to the general equilibrium impact of risk sharing through taxation with marketed risks if

[82] This framing of the issue is different from that in Gentry and Hubbard, 1997. They consider consumption taxation implemented by a wage tax combined with a business cash flow tax. Although they purportedly are addressing distributional implications, their focus is on evaluating the difference in taxation from the perspective of a firm's investment decisions, as opposed to a household's life-cycle labour supply and savings choices. As a consequence, they focus on the marginal value of immediate depreciation of investment to a firm, which they value using the safe rate of interest, supporting the view that consumption taxation exempts the safe rate of interest but not the return to bearing risk or pure rents. Modeling household choice as a base for examining the impact on the distribution of utilities of giving the deferral advantage is more complicated. While stocks and bonds have the same marginal value with portfolio optimization, the impact of deferral on the inframarginal gains from the availability of stocks is relevant for distributional analysis. As a quick example of this issue, for given wealth and Cobb-Douglas preferences the higher the distribution of risky returns, the greater the gain from deferral for a given portfolio mix. Since the optimized portfolios may well be different, a full analysis is more complicated. But this seems the appropriate way to approach the distributional impact.

the government does not return the risks to the economy in an offsetting way. Again, the returns to scale, now on the aggregate level, matters for the impact of taxation.

Part V. Age-dependent earnings taxes

From the perspective of optimal contract theory, any costlessly observable variable correlated with unobserved characteristics or behavior should influence payoffs, even if it is poorly measured and the correlation is limited. Applying this perspective to optimal taxes in an extended Mirrlees model, labour income taxes should depend on all variables correlated with the ability to earn, even those measured poorly. While tax systems have stupefying complexity, it is not from incorporating many such variables.

Primarily, the approach to optimal tax theory in this chapter has been to take as given a set of allowable tax tools (while ignoring the cost of administration), chosen to reflect actual (or plausible potential) use and chosen to enable the inferences from a model to be useful for policy discussions. Some analysts have considered it significant to replace this approach of designated tax tools by assuming that the choice of tax tools is an endogenous part of the optimization, subject only to observability constraints. A common assumption in these formal models is that taxation is based on costlessly, perfectly observed variables while all other variables are not observable at any cost. But this description of observability is not accurate on either side – earnings are costly to measure and are not perfectly observed and there are other (costly, imperfectly) observable variables that could increase social welfare if used optimally. Thus standard assertions about observability, commonly used to "derive" a tax base rather than assuming it, are not an adequate guide to the choice of a tax base for direct taxation. Complexity of the tax base matters, as do both public reactions and the political economy of a more complex structure, both related in part to views on horizontal equity. We are lacking in analyses that take us very far in considering when additional complexity is a good or bad idea, since issues raised by complexity are not part of the formal modeling. In the absence of extended analyses on which to draw, using complexity concerns to influence policy inferences from formal models is subjective, but seems important. We simply refer to variables as taxable and non-taxable, rather than observable and non-observable, reflecting an ex ante judgment call reflecting these multiple dimensions of relevance for choosing a tax base.

To explore the extent to which further complications should enter taxation, we consider three examples of variables that might be used to influence the taxation of earnings - hours worked (and so earnings per hour), height, and age. Only the third is recommended. Two issues are raised by the consideration of additional variables – the ability of (and cost to) governments and taxpayers to deal with greater complexity and perceptions of equity, both by analysts and the public.

Income taxes are based on earnings without an attempt to measure hours worked and so average earnings per hour. Minimum wage rules and requirements for paying higher wages for overtime both require some measurement of hours worked. And the Working and Childcare Tax Credit programmes in the UK base transfers on doing at least a minimum number of hours of work. In the cases of minimum wages and overtime pay rates, the employer and the employee have conflicting interests in the measurement of hours. This makes enforcement easier than enforcement of a tax that depended on hours worked would be, since neither the employer nor the employee has an interest in higher taxation of earnings. While this conflict of interests also does not exist in the tax credit programmes, they follow the common practice of programs being more intrusive and more measurement focused when applied to poorer people than when applied to the general public. An attempt to incorporate a measure of hours worked into the tax base would plausibly bear considerable correlation with actual hours. For many workers in large firms or government employment, existing financial records would form a good basis for estimating hours worked with reasonable accuracy. Moreover a requirement for self-declaration of hours, subject to some form of random monitoring, would fit the theoretical category of a correlated, poorly measured, but nevertheless useful basis for further tax distinctions. And it is not as if earnings were measured perfectly either.

Thus, if it did not recognize factors other than observability, optimal tax theory would call for basing taxes in part on estimated earnings per hour. We do not think that using an hours measure in determining taxes would be a good idea, however, and it is useful to consider why not. Basing taxation on inaccurately measured variables leaves more scope for administrative discretion and encourages cynicism about the fairness of the tax system. Both features are likely to add to the difficulty of encouraging voluntary

accuracy in reporting and support for the politics of better taxation. This is already a problem resulting from the inaccurate measurement of income. But income (or consumption) is central to distributional concerns and it is hard to see how to have satisfactory taxation without it. Adding to concerns about inaccurate measurement should not be done lightly. The theory of how to use poorly measured variables would not be intuitive to either legislators or the public, again making good tax politics more difficult. In sum, basing taxes in part on hours worked does not seem to be a good idea, although that intuition is not supported by formal analysis as far as we know.[83] As with the Meade report, concern about multiple aspects of taxation leads to this conclusion, whereas the opposite conclusion would follow from taking optimal tax theory literally and ignoring aspects of taxation not included in the formal modeling.

As another example, this one where accuracy of measurement is not at issue, consider the findings of Persico, Postlewaite, and Silverman, (2004) and Case and Paxson (2006) that there is a correlation between height and earnings abilities.[84] With standard modeling and different tax structures for adults of different heights (possibly distinguished by gender), one can then have higher social welfare than without such multiple tax structures. While it would be somewhat complicated for tax authorities to have multiple tax structures, there is not much complication for the taxpayer who does

[83] In the exploration of lessons from the literature, we do not explore the (small) literature using hours worked in determining taxes.

[84] Mankiw and Weinzierl (2007) also consider relating income taxation to height. They discuss the evidence on the link between height and earnings, present the argument that such an approach would increase social welfare, and do a first pass at the structure of such a tax. The authors' interpretations of the result differ. "One of us takes from this *reductio ad absurdum* the lesson that the modern approach to optimal taxation, such as the Vickrey-Mirrlees model, poorly matches people's intuitive notions of fairness in taxation and should be reconsidered or replaced. The other sees it as clarifying the scope of the framework, which nevertheless remains valuable for the most important questions it was originally designed to address." (Page 2.)

We share the second view. As this essay has argued, the insights from optimal tax theory are only part of the considerations relevant for tax policy, but an important part. Indeed, the role of fairness concerns in limiting allowable tax tools was argued by Atkinson and Stiglitz (1980). The methodological error in the "reconsider or replace" view comes from taking the answers to formal models as a literal policy recommendation. By their nature, models are a simplification of reality in order to have a sufficiently tractable basis for reaching conclusions within the model. As such, every model has inaccurate assumptions and could be used to derive silly inferences by focusing on the implications of that inaccuracy. At their best, models are good for some questions and not for others. Finding a question for which a model (or modeling approach, as in this case) gives a rejected answer is not a serious critique of the model or modeling approach.

not get to choose among tax structures.[85] And by restricting the set of tax functions to a small number of different height intervals, the complexity for legislation would not be enormous. What does seem important is that unlike the example of different tax structures for different ages discussed below, a set of tax functions based on height is a setting of consistently different structures for different (fully-grown) individuals rather than individuals passing through the different tax structures as they age.[86] This distinction seems important for political and public acceptability, and possibly for the ethical underpinnings of taxation.

Consider a sequence that starts with extensive research documenting that such differences are real and robust to alternative measurement approaches, explains to the public and tries to convince them that this is the case, and then tries to explain to the public why this is a useful basis for differences in taxation. Then picture a legislature considering a half dozen or so different tax structures on this basis.[87] Presumably the incentive for parents to stunt the growth of their children would be minimal if they also recognized that the factors correlated with height do affect earnings abilities. Does this scenario violate some sense of horizontal equity? If height were irrelevant, it would. But once height is linked to earnings ability, then people of different heights are not identical as far as the government's ability to infer ability is concerned. That is, the government's ability to raise revenue relative to income distribution and efficiency concerns differs by taxpayer height. This is similar to the view that people with different tastes for work are not identical, even if they have the same budget sets. Whether the gain in social welfare were small or large would depend on the magnitude of the correlation and the extent to which different tax structures had an impact on optimized social welfare.

We feel comfortable in rejecting this idea out of hand, as did Mankiw and Weinzierl (2007). What is harder than reaching that conclusion is sorting out its

[85] Allowing ex ante choice among tax structures may be a source of welfare gains (Luttmer and Zeckhauser, 2008). We do not explore this option –if significant, this added complexity may challenge the ability of many to figure out which tax structure to pick and could be viewed as inequitable as some workers successfully lowered taxes significantly by a good choice while others regretted poor choices.
[86] This ignores the shrinkage that occurs with aging.
[87] Think just about earnings, but it might also be the case that different heights are also correlated with different abilities to invest and so different possible rates of return and different intertemporal discount factors and thus different tendencies to save.

underlying basis. Mankiw and Weinzierl offer several reasons for rejection. One is that this might be the first step in a sequence of taxes that vary with demographics, and while one might be acceptable, the end point of such a process would be unsatisfactory for its administrative burden and invasiveness. They counter this argument with the view that some demographic variables are used already, others are widely unacceptable and this need not be a slippery slope.[88] They note the political risk element – "democratic societies may have an interest in avoiding the taxation of specific groups as a matter of course to counter the majority's temptation to tax minority groups." (Page 13.) More generally, there is always concern about politically well-connected groups skewing policy to their advantage, at the expense of some wider measures of the public good. This is an issue here, in part, because height is not the only demographic variable that could be used in this way. We would not like to see an exploration of which variables would be most attractive to the politically more powerful. Mankiw and Weinzierl recognize a possibility of stigma, but do not see that as important. Mankiw and Weinzierl offer two critiques of utilitarianism – coming from libertarianism and horizontal equity. Unlike libertarians, we are not "skeptical of the redistribution of income or wealth because they believe that individuals are entitled to the returns on their justly-acquired endowments." (page 15.)[89] But we do not pursue this issue here. We do share Mankiw and Weinzierl's concern with horizontal equity issues, pursued further in VII.D. An additional point is that, contrary to the hypothetical above, the public may not be convinced of the equity of such an approach since there is only a stochastic relation between height and earnings abilities. The public's sense of equity, largely formed without deep thought, nevertheless has some relevance in a democratic society. Also relevant is the public's reaction to its sense of equity. This issue is discussed further in

[88] A similar optimal tax argument could be made with regard to gender, given gender-differences in life-expectancies and the shapes of life-cycle earnings profiles. As with age, gender is not used extensively in tax systems although, again, it has played a large role in public pension system rules in some countries, such as the UK (at present).

[89] Individuals do have entitlements, but the strength of entitlements and the bases of entitlements do not lead us to skepticism of the appropriateness of redistribution, but to limits in taxation.

VII.D. Our exploration of this example is to permit distinctions with age-dependent taxes, which involve different issues.[90]

In contrast to height, age is used by actual tax structures, but very little apart from retirement-related rules. In the US there are distinctions for children (who can be dependents and so provide additional deductions) and those over 65, who may receive an additional deduction. In France tax rates depend directly on the number of children through the *quotient familial*. Whilst there are no deductions for dependent children in the UK, the system does include an additional allowance for those aged over 65 and a further additional allowance for those over 75, although for higher income individuals these are both tapered away back to the level of the under-65s allowance.[91] These examples do not provide much variation in taxes across ages, nor do they provide a systematic variation in marginal tax rates. In contrast, age does play a large role in the rules for both public and private pension systems and in some countries in tax-favored retirement savings opportunities.[92] Eligibility for receiving pension benefits is commonly age-based. Benefits typically increase with the age at which they start and the rates of increase commonly vary with age – for example by only being available for a range of ages, as in the UK, and also by having different percentage calculations at different ages, as in the US. In countries that use some form of retirement test, benefit eligibility rules relative to earnings also commonly vary with age. Further complexity often comes with pension reform, with age-related rules being different for people of different birth years. And we note that in Switzerland, the mandatory occupational pension has contribution rates that

[90] An appropriate question to ask is how complicated a tax structure a legislature can use well. Historically legislatures have relied more on their own decision-making in the realm of taxation (and other topics in economics) than in other areas – legislatures vote money for bridges, they don't vote blueprints. Perhaps further addressing of complexity (beyond what is already left to staff) could be allocated to some expert group, as Breyer (1993) has proposed for dealing with health risks. And perhaps the public would accept both the underlying idea and the use of experts.

[91] In addition those over 65 in April 2000 still receive the married couples allowance which was abolished for individuals younger than 65 on that date (i.e. born after April 1935). This allowance is also tapered away as income rises.

[92] In the UK, apart from the tax favoring of partial annuitization and the requirement to annuitize three-quarters of private pension assets by age 75, tax favoured assets are available for withdrawal with no restrictions on age or holding periods and as such are simply tax favoured general savings vehicles, unlike in the US where such assets are retirement saving vehicles (i. e., subject to extra taxation if withdrawn at a younger age).

vary with the age of the worker. Thus, it is natural to explore reducing the large difference in the use of age between pension rules and tax rules.

In the context of a one-period model of income taxation, and with a focus particularly on younger workers, Kremer (2001) called for different tax structures for different ages. Applying the Mirrlees model separately to different age groups, he argues that the distributions of earnings and the labor supply elasticities are so different across ages that the implied pattern of optimal tax rates would vary greatly by age. Borrowing constraints that are prevalent among younger workers may be a further basis for different tax structures.[93]

Let us consider a political process if such an approach were taken. The first step might be to allocate each age to one of a small set of ages, in order to limit the number of tax schedules.[94] Perhaps the set might be under-30, 30-50, 50-65 (or the state pension age), and over 65. For simplicity, there might be a given set of marginal tax rates with only the break points varying as a function of age. This doesn't sound too hard for a legislature to do.[95] And plausibly it could be worked out without undue pressure by the politically better-connected ages. With suitable transition rules, this does not violate horizontal equity concerns that are lifetime based, and presumably would be as publicly acceptable as are age-related pension rules.

As discussed above, formal models do show advantages to age-dependent earnings taxes. Beyond theoretical observations, Weinzierl (2007) has done an optimization calculation to find the advantage from age-varying rules. He compares a single tax regime with a system with three tax regimes for ages 30-39, 40-49, and 50-59. He uses data from the PSID to calibrate a model of wage rates for five representative workers representing different quintiles of lifetime earnings. He uses the mechanism design approach referred to above. With 5 agents and 3 periods, the government sets up

[93] Recent analyses of age-dependent taxes include Blomquist and Micheletto (2003), Erosa and Gervais (2002), Gervais (2003), Fennell and Stark (2005), Lozachmeur (2006), and Weinzierl (2007).
[94] If there are joint returns for couples based on a couple's total incomes, labour income might be taxed on the basis of the age of the earner while capital income might be taxed as if each received half. Or all taxable income could be treated as if half were taxed on the basis of the age schedule of each of the couple.
[95] This assertion may be undercut by the common practice of adjusting public benefit formulae for the age at which they start with a linear formula, when multiplicative or more complex formulae seem to make more sense. Supporting the thought of delegation is the automatic adjustment in Sweden, done on a roughly actuarial basis, although one with rules for the actuaries set by legislation.

to 15 earnings/net-of-tax earnings pairs. Without age-dependent taxes, each period each agent chooses one out of the full 15 pairs for each period, using the capital market to optimize lifetime utility. With age-dependent taxes, each period each agent chooses one of the (up to 5) offerings available at that age, again using the capital market to optimize lifetime utility. Compared with the optimum with a single tax function (15 choices for each period), he finds that average taxes are lower on young workers and higher on older ones with age-specific taxes (and so only 5 choices each period). He also finds a large welfare gain from the optimal three-age tax function compared with a single tax function, the same for all ages, equivalent to 2 percent of aggregate consumption. This is two-thirds of the gain from going to the full mechanism design optimum (where individuals are restricted to (up to) 5 lifetime plans, rather than being free to piece together separate plans each period). While interesting, this is clearly just a start on exploring this issue, so this is really a call for research on an issue that seems to have a good probability of leading to significant policy improvements.

A different approach to taxing earnings over a lifetime looks at current earnings in the context of previous earnings. This could be done in a variety of ways, including a moving average over a fixed number of years or basing lifetime taxes on lifetime earnings, with annual taxes viewed as withholding toward the eventual determination of lifetime taxes.[96] In the discussion of two-period models above, we noted how this might serve social welfare maximization. Now we consider the ability to implement. This certainly is doable, with the government providing historic information along with tax forms. Indeed, we can consider this as parallel to rules that determine public pensions. Defined benefit pensions are based on the history of earnings, possibly a full history (as in Sweden) or a long history (as in the US). In a wage-indexed system for initial benefits (that are then price-indexed), as in the US, the benefit formula relating benefits to earnings varies with date of birth through automatic indexing. Indeed legislated future ages for receiving full benefits vary with date of birth in the US. In the UK, such a change is already underway with the movement of the state pension age for women from 60 to 65 over the period 2010 to 2020, and further increases in the state pension age for

[96] These would be similar to the approach in Vickrey (1947), who cumulated annual income, not annual earnings and who considered various lengths of time for the cumulation.

both men and women will follow (from 65 to 66 in 2024, from 66 to 67 in 2034 and from 67 to 68 in 2044), although this can also be viewed as different age-dependent rules year-by-year. And Sweden has automatic adjustments that apply to each birth cohort different determinants of initial benefits (for a given earnings history) and of the growth of benefits from a delayed start.

Thus a key question is whether variation in annual tax rates as a function of age is a bad idea because of complexity or a case of theory being ahead of policy, with research on tax design needed, but reform called for. We are inclined to take the latter view for countries that have a good legislative process.

Part VI. Diverse savings behavior

The models explored above assumed life-cycle savings. Yet it is clear that this is not a highly accurate model of behavior for everyone.[97] Alternative modeling of savings behavior, seemingly relevant for significant portions of the population, include precautionary savings, time-inconsistent behavioral models consistent with too little savings, and utility-of-wealth models which appear to make more sense for those with very high wealth.[98] Moreover, behavioral models and experiments have explored how individuals respond to alternative ways of encouraging additional savings.[99] Behavioral analysis of savings behavior is highly relevant for the choice of tax base. It is also important for evaluating the role of mandatory programs that require contributions when working and provide benefits when retired. And these two institutions need to be considered together. A key tax design issue is how to combine concern that some fraction of the population saves too little for an adequate replacement rate in retirement while another fraction saves too much, resulting in their retiring too soon from the perspective of social welfare optimization, as played a role in the models in Part II. Behavioral diversity as well as heterogeneity in life expectancy, intertemporal preferences, and consumption history (in light of realistic links between past consumptions and later marginal utilities of consumption) all call for diversity in individual saving rates, which also played a role in the models in Part II. And alternative modeling of those accumulating very large wealth is relevant for choosing the tax base in light of the great inequality of wealth holdings. This diversity in savings behavior has not received much attention in tax modeling and would appear to be an important issue for

[97] As Bernheim (1997) has written: "While it would be rash to dismiss the many empirical successes of the LCH [Life Cycle Hypothesis] and discard it unconditionally, it is equally rash (in light of its empirical failures and well-founded skepticism about its underlying premises) to employ this theory as the sole organizing principle for understanding savings incentives."

[98] On the diversity of savings behavior, see Dynan, Skinner and Zeldes (2004) and VIII.A.

[99] Behavioral economics has become a major research area for many economists and some of the findings are very exciting (for a survey relative to public finance, see Bernheim and Rangel, 2007). Indeed, analyses of the difference in outcomes with opt-in and opt-out rules for retirement savings plans are already influencing policy makers in both the US and the UK - the introduction of Personal Accounts in the UK, whereby individuals are automatically enrolled in private pensions by their employer unless they choose to opt out was announced in 2007, is being legislated in 2008 and will be introduced in 2012.

future research. The following conjectures are highly speculative, but seem worth exploring.

The behavior of those with very large wealth appears to require modeling utility for some people as coming directly from wealth holding, not indirectly from later consumption (Carroll, 2000). This suggests an inelasticity in consumption behavior that would seem to justify very high taxes on capital income on those with very high wealth.

Concern about too little savings for retirement suggests a program of tax-favored retirement savings (to supplement mandatory provision of retirement income if that program is not extremely large). Recognition of diversity of savings and the advantage of discouraging too early retirement suggest limiting the extent of access to tax-favored retirement savings accounts, as well as preserving their character as retirement accounts. But recognition of diversity in the savings behavior in the population does not appear to call for rejection of the basic conclusions reached above. Instead it suggests modifications of the policy (e. g., tax favored retirement savings). And behavioral issues (both mental accounting and self-aware self-control) suggest it may be useful to have additional reform as opposed to just exempting from taxation some level of income from capital. Examples are some form of autoenrolment (see Beshears et al (2007)) or else some active roles for third parties (e. g., employers and financial institutions) as noted in Bernheim (1997). But this is primarily a call for research and a conjecture about outcomes of such research, not a firm basis for policy.

More research is also warranted on the optimality properties of the different ways of structuring tax-favoring for retirement savings. Options in use for tax treatment of deposits, of accumulations, and of withdrawals include: (1) exempt-exempt-taxable (EET), as in Personal Pensions in the UK or IRAs in the US, (2) taxable-exempt-exempt (TEE), as in Tax Exempt Special Savings Accounts or their successor, Individual Savings Accounts in the UK, or Roth IRAs in the US, (3) having both available, and (4) having partial taxation of accumulation income (as was in Australia). Further research is also warranted relative to proposals and practices that allow tax-favored savings for other purposes, such as house purchase, medical expenses, and unemployment.

The impact of earnings uncertainty on the desirability of taxing capital income suggests that taxation of capital income might well be different at ages when much of the working population is expected to be retired than at earlier ages. Combining this with the role of tax-favored treatment of retirement savings and the presence of precautionary balances at all ages suggests there may be an advantage (unexplored in the literature as far as we know) from age varying capital income taxation for capital income outside the retirement accounts. This could be done, for example, by capital income tax exempt amounts that varied with age.

Part VII. Further issues

Part VII touches on a number of issues including a further discussion of the use of a social welfare function (VII.A.), government commitment (VII.B.), some modeling assumptions (VII.C.), and horizontal equity (VII.D.). These sections examine the underpinnings of the approaches to taxation discussed above.

A. Social welfare function

Based on its use of a social welfare function, the optimal tax approach is often accused of assuming a benevolent government. This criticism has both right and wrong elements. Calculation of what a benevolent government should do is not the same as assuming that there is a benevolent government. Rather it is asking a key question - what policies would one want to see a benevolent government follow. The answer to such a question can help inform a democratic debate about government policies, which is all that academic economic research can hope to accomplish by itself. Moreover, it is hard to see how one gives policy advice without knowing the link between good design of policies and the accomplishment of social ends.[100]

The relevant part of the accusation is that the political tendencies of actual governments are highly relevant for good policy recommendations. Awareness of political tendencies can readily take two separate forms. One is to extend optimal tax theory to incorporate additional constraints reflecting what governments are likely to do, either in response to current recommendations or in future policies that may be influenced by current legislation. This is a richer, and possibly more relevant, environment than considering a constitutional approach to limits on taxability. The literature on tax policy without government commitment is a form of such analysis,

[100] As Musgrave wrote: "Just as homo economicus or a competitive Walrasian system are useful fictions to model an ideal market, so it is helpful to visualize how a correctly functioning public sector would perform. … Unless "correct" solutions are established to serve as standards, defects and failures of actual performance cannot even be identified." Buchanan and Musgrave, 1999, page 35.

although one that typically does not have a rich, empirically supported theory of government behavior in a democracy. A second form that awareness of political tendencies can take is through judicious use of the insights from optimal tax results when moving from basic theory to policy recommendations. Recommendations can reflect beliefs about the workings of the political process, based on the current state of politics and political science and projections of political evolution.

The optimal tax literature works simply with a social welfare function. With individual utility depending on both consumption and the disutility of labor, this is not equivalent to attention focused on income distribution, particularly using a social income evaluation function as developed by Atkinson (1970). While we share a concern about income distribution, a social income evaluation function is no substitute for a social welfare function in thinking about tax policy.[101] This approach appears to give too much weight to encouraging work, particularly by low earners, and we do not think that maximizing a social income evaluation function is a useful variant on social welfare function maximization. Nevertheless, one might consider limiting income variation (perhaps because of political implications), which would also imply rejecting possible Pareto gains.

B. Time frame, commitment and transition

Support for total annual income as the ideal tax base appears to rely on using a year as the time frame for thinking about individuals when doing normative analysis. In contrast, the optimal tax models that are the basis for this chapter rely on lifetimes (or beyond) as the time frame for normative analysis. Exclusive focus on either of these two time frames seems incomplete. On the one hand, the current position of individuals is a result, in part, of their own past decisions. It does not seem adequate to frame the basis for policy choice in a way that ignores intertemporal aspects of incentives, a normative dimension of responsibility for future consequences of one's current actions, and a normative response to the consequences of one's past actions. On the other hand, a

[101] Nor do we see a case for an objective function that combines both a social welfare function and a social utility function.

lifetime perspective does not adequately allow for individual time-inconsistency and does not contain a normative adjustment for the consequences of decision mistakes. For example, previous high levels of saving do seem to provide some normative support for higher current consumption, while previous low levels do not seem to be sufficient warrant for enforcing some very low levels of consumption. And such concerns need to be tempered by their incentive effects.

In democratic (and nondemocratic) societies, further complicating consideration of government policy at a particular time are the inevitable changes in normative evaluations from the bases for past government policies as governments change. Also relevant is the inevitable incompleteness of both government plans for future policies and government understanding of the consequences of chosen policies. That is, normative analysis needs to consider the degree of adjustment that should be made for the implications of past policies. That different models use different time dimensions is part of the reason why it is inappropriate to rely too heavily on any single model's implications.

1. Commitment

Although tax legislation can have an open-ended horizon, it is expected that taxes will change as circumstances develop and governments change. Moreover, governments do not commit to a complete (contingent) set of future policies. Individuals making decisions that affect their future tax liabilities (such as investments and education) are faced with uncertainty about future circumstances, future governments and their possible tax reforms, and any transition rules the government may include in tax legislation. The Meade report call for "a certain stability in taxation in order that persons may be in a position to make reasonably far-sighted plans" (Page 21) also suggests seeking tax instruments that are relatively simple and transparent to aid the formation of appropriate tax expectations by individuals.

In the ongoing process of the adaptation of tax policies to economic and demographic developments as well as to changing normative perceptions and political

balance, a set of rules/guidelines for transition issues is important both economically and politically. From this perspective we can appreciate the Meade Report's concern for flexibility and stability: "A good tax structure must be flexible for two rather distinct purposes. … there must be recognition of the need to be able to adjust total tax burdens reasonably rapidly and frequently in the interests of demand management. … In a healthy democratic society there must be broad political consensus – or at least willingness to compromise – over certain basic matters; but there must at the same time be the possibility of changes of emphasis in economic policy as one government succeeds another. … But at the same time there is a clear need for a certain stability in taxation in order that persons may be in a position to make reasonably far-sighted plans. Fundamental uncertainty breeds lack of confidence and is a serious impediment to production and prosperity." (Page 21.)

Beyond any possibility of short-run demand management, there are changes in long-run fiscal needs that are likely to occur from trend developments in economic and demographic circumstances, as well as the spreading over the future of short-run changes in fiscal needs (e. g., after a war).[102] A research program that addresses the need for both adjustment and stability would seek a tax structure that has enough political acceptability to relegate tax changes primarily to parameter changes in a class of parameters anticipated to adjust to circumstances. The tax design would need to recognize that individual expectations about future taxes are endogenous to the policy framework being created. Such modeling would examine a balance between the different effects of changing policies.[103]

In addition, given the difficulty of radical change, the existing basic structure of taxation influences the political process. Indeed, links between the form of public pension design and anticipated future legislation has been part of the debate in the United

[102] Currently discretionary fiscal policy, while pursued by governments, is not in high favor among academic economists (Auerbach, 2002). But built-in stabilizers, while not getting much active attention, are still viewed positively (Auerbach and Feenberg, 2000). It is odd that there was not discussion of built-in stabilizers in the Meade Report.

[103] Such analysis might parallel for an economy the analysis for individuals in Amador, Werning and Angeletos, 2005.

States between defined benefit and defined contribution mandatory public systems.[104] Similarly, implicit in our focus on the tax base, separate from tax rates, is an assumption that tax rates are being optimized for given tax bases, thereby ignoring the political linkage that may well be present between tax base and tax rates. It is incomplete to say that a suitable choice of tax rates can make a different tax base have comparable overall progressivity if that suitable choice will not happen. Recognition of the link between the form of tax institutions and the perceptions and salience that then influence policy making is important.

In light of the expectation of repeated adjustments of taxes, how should we use the findings of the models analyzed above, which considered government policy being set for a lifetime or an infinite future? A start of an answer is to say that in thinking about policy, one would like to know what policies would be good if they could be set for a long time. And drawing inferences from a model with committed taxes would recognize the decreased relevance of those parts of the optimization that relied on unrealistic elements of the modeled commitment.

For example, the Chamley and Judd papers have two results. The first, discussed above, is to have no taxation of capital income, either after a finite date or asymptotically (that is taxation can be positive indefinitely, but with a steadily shrinking tax rate). The second is to tax initial wealth as heavily as possible, at least in the representative agent version. In the context of these models with infinitely-lived agents, the second finding has had little direct influence on policy recommendations drawing on the literature. Nevertheless, the same perspective, clearly stated, lies behind arguments in OLG models for switching from income taxation to consumption taxation particularly as a way to transfer wealth from older cohorts at the time of tax implementation with little in the way of distorting incentives.[105]

It is appropriate that these two Chamley-Judd results have been viewed so differently. Taxing initial wealth as much as the available tax tools allow (whether as a

[104] For example, see Diamond, 1999, chapter 3.
[105] This basis for a change in taxation is very sensitive to implementation. It works for taxing consumption directly and for taxing consumption as income less savings provided initial wealth is measured, but may not work for taxing consumption as income less savings if initial wealth is not measured.

wealth tax or a capital income tax) strains the relevance of the assumption that the government can commit to a policy that this taxation of wealth will end. Without a genuine commitment technology, confiscatory wealth taxation would adversely affect savings behavior and have serious efficiency costs (even if the government saves the revenue) because of concern that such taxation will return. A switch from income to consumption taxation (with limited grandfathering of existing wealth) could be interpreted as a move against wealth which has limited implications for future taxation of wealth since the set of politically plausible tax policies has not changed very much – increases in the taxation of consumption are limited because they fall on everyone. On the other hand, some people may recognize that the underlying principle of the efficiency advantage of taxing existing wealth would continue to be present, even if it required a different tax change to implement.

These assertions raise the critical question of how to model the link between tax legislation and expectations about future taxes. One approach in the literature is to model a consistent game-theoretic equilibrium between tax setters, potential alternative tax setters, and taxpayers, with the threatened reactions by the taxpayers limiting the setting of taxes. This literature seems to rely too heavily on a game-theoretic equilibrium drawn from oligopoly theory with a limited number of sophisticated players for use in a setting of vast numbers of players, many of whom are ill-informed. The literature, now in its early stages, may well develop into something useful, but does not yet seem very informative. Nevertheless, the literature is interesting in making clear the effects of expectations about taxes on economic incentives.[106]

An alternative way to view 'commitment' is in the realm of precedents, paralleling their role in legal decisions (see, e. g., Kaplow 2006b). Assume the government announces a one-time capital levy. That is a precedent for doing the same again, and so lacks credibility that it really is one time. Perhaps there are special

[106] We note that the Chamley-Judd finding of asymptotically vanishing taxation of capital income with full commitment has been extended to a setting without commitment (Dominguez, 2007, Reis, 2006). These papers assume a single infinite horizon budget constraint. Zero asymptotic taxation of capital is not optimal when the government faces period-by-period budget constraints. For recent modeling of tax equilibrium with potentially competing governments, rather than a single government, see Acemoglu, Golosov and Tsyvinski (forthcoming).

circumstances, such as a war or meteorite impact that is unlikely to recur. Then the precedential cost may be much lower, although there remains the effect of a possible perception of an increased risk of a widening of the precedent. Just as individuals set rules for themselves, with bright-line rules easier to adhere to,[107] so too the government process recognizes that crossing a bright-line rule runs the risk of major backlash – whether it is losing elections, with possible reversals of policies, or street demonstrations, or political backlashes in other realms. Thus one might prefer a small annual wealth tax rather than a large one-time tax, on the grounds that expectations of continuing and possibly slow growth of the annual tax has less of a deterrent effect on savings through perceptions of future policies. Switching from an income tax to a consumption tax has the effect of taxing existing wealth, with possible future increases in the tax rate as then a risk discouraging savings. Again, we would expect less of an impact. This way of approaching the issue of commitment, or its lack, differs from a common game-theoretic approach using trigger strategies in not assuming widespread sophisticated understanding of equilibrium, and in recognizing the limited awareness of politics of some and the multiple motivations affecting voting.

2. Transition

Transition issues arise in two ways in a discussion of the tax base. First, analysis of the tax base needs to recognize that there will be future tax changes, and those changes will involve transition issues. Second is the set of one-time transition issues if the contents of this essay (or some other) were to be accepted as the basis for current and future taxation. These issues differ in that current and past expectations are given when considering today's changes, but expectations about future changes are endogenous to the policy framework created today. Both settings can call for giving some degree of respect to legitimate expectations for both incentive and fairness reasons.[108]

[107] It appears easier to comply with a no cookies or no cigarettes rule than trying to allow oneself only a few.

[108] For discussion of ongoing changes, see Graetz (1985) and the sources cited there. For discussion of an initial change, see Auerbach (2006), which presents many issues and highlights the importance of transition by contrasting simulations that have the same long-run tax incentive properties but very different transition

Today's changes can influence expectations (and the normative pull of the expectations) by including adjustments for transition reasons and by legislation and statements about future tax changes. Adjustments for transition reasons include grandfathering, delaying implementation, and explicit transition relief.[109] Given the frequency with which taxes change, no one should expect that taxes will never change. Taxes change because circumstances change and because governments change. And sometimes tax legislation has a time limit (a sunset), which gives a date by which taxes are more likely to change again (rather than a commitment to a return to the tax law that would take effect with no further legislation). In their own self-interest people should recognize the possibility of a tax response to changing circumstances. And such recognition can improve social welfare. Modeling with changing taxes (and expectations of changing taxes) in response to changing circumstances is common in the tax literature coming from macroeconomists, (e. g., Golosov, Tsyvinski and Werning, 2007). It may well be useful to take this approach in more complex economic environments (e. g., with human and physical investments of different effective lifetimes) and with explicit transition rules. And it would be good to explore how the basic tax structure may affect tax setting with endogenously changing governments, although it is not clear how to set up a suitable social welfare function.[110]

Beyond standard social welfare analysis in terms of lifetime expected utilities, there may be a further normative concern for limiting the deviations from appropriately held expectations about policies.[111] The presence or absence of an ongoing political discussion should affect the appropriate degree of respect for actions based on expectations. And one would need an evaluation of the political process to allow different normative treatments of changing "loopholes" that come from less satisfactory

---

impacts. Whether ending the taxation of capital income raises or lowers social welfare varies with the transition impact in some simulations.

[109] Use of these tools was raised in Feldstein (1976b).

[110] As noted above, the type of pension system is thought to influence the changes in a pension system in response to changed circumstances (Diamond, 1999). For an example of equilibrium dividend taxation with changing governments, see Korinek and Stiglitz (2008).

[111] This might parallel the same issue in the legal analyses of contracts, where courts attempt to interpret contracts in the light of the expectations of the contract parties. The endogeneity of legitimate expectations to court processes that try to decide in terms of the expectations of the parties has not always received adequate attention.

aspects of politics and changes of "appropriate" political outcomes. That is, the degree of respect to past taxes and the expectation of their continuation need to recognize a widely held view that the tax structure is not satisfactory and ought to be reformed (a view that underlies the commissioning of this report).

C. Modeling assumptions

The optimal tax literature analyzes real taxes dependent on real labour and capital incomes. We do not think there is any significant disagreement among economists that to the extent feasible, the relevant basis for taxation should be real capital income, not nominal capital income. A literature has examined how and to what extent this can be done (Aaron, 1976). We have not considered how optimal tax insights should be adapted to the common practice of taxing nominal incomes. Other than pointing out that (with positive inflation) taxing nominal interest and dividends results in taxes on real interest and dividends at rates higher than the stated marginal tax rate, we do not explore the real-nominal distinction. We also do not explore issues related to the realization of income, but note that for equal treatment with other capital income, taxation of deferred realization of incomes, as with capital gains, calls for heavier taxation than non-deferred capital income, not lighter taxation as is common practice (Helliwell, 1969, Auerbach, 1991, Bradford, 1995). Heavier taxation for longer holding periods can limit the lock-in effect.

Overwhelmingly, optimal tax models assume competitive behavior by firms. While this is not a genuinely satisfactory assumption, we have not explored the limited literature that considers other market structures.

Typically, the labour market is modeled as if workers can choose the number of hours to work at the wage available to them. Such a simple linear before-tax budget constraint is not realistic for many people, given rules on overtime pay and possibly different earnings per hour on primary and secondary jobs. Also many jobs come with a standard number of hours, although the standard number of hours at an employer is a choice variable that plausibly reflects to some degree the hours that workers would like to

work. Some of the literature recognizes the discontinuity in disutility of work at zero hours (e. g., from commuting) that makes withdrawal from the labour force a possible next-best alternative to work with a significant number of hours. The distinction between extensive (labour force participation) and intensive (hours worked) labour supply margins is very important for considerations of tax rates and acknowledging both margins can lead to a greater role for the average tax rate in policy analysis (Saez, 2002c). This issue is particularly important for programs aimed at encouraging work by low earners. Moreover, since the relative importance of intensive and extensive margins varies widely by age, this is relevant for the case for age-dependent taxes. Since it is most common in the literature, we focus on models with adjustable hours, although the retirement literature often makes use of a zero-one model of employment opportunities.

D. Horizontal equity

We rejected starting the discussion of tax policy with an ideal tax base based on equity considerations. But we do recognize a role for considerations of horizontal equity, mentioned briefly above. In this section, we elaborate on the reasons for rejecting the centrality of an ideal tax base and then consider some of the literature about horizontal equity.

1. Ideal tax base

To consider horizontal equity in a simplified setting, let us consider a basic one-period, two-good model. With no savings, consumption and earnings are the same. As indicated in the Meade Report, there is tension between the idea that ability to pay should be based on actual outcomes or on budget sets (potential outcomes).[112] If everyone really does have the same preferences over work and leisure, and preferences have plausible

[112] Reflecting the acknowledged difficulty in defining taxable capacity, the Report asks: "Is it similarity of opportunity or similarity of outcome which is relevant?" and "Should differences in needs or tastes be considered in comparing taxable capacities?"

properties,[113] then there is no tension between the actual and potential measures since those with higher potential earnings have higher actual earnings. This convergence of different competing measures of ability to pay could strengthen the case for paying attention to horizontal equity. However, with identical preferences in this two-good model, there is no conflict between this horizontal equity concept and the standard optimal tax calculation since individuals with the same productivities pay the same taxes in equilibrium.[114]

In modeling preferences in an optimal tax problem, it is common to use $u[x]-v[z/n]$, where x is consumption, $u[x]$ is the utility of consumption, z is earnings, n represents what varies in the population, and $v[z/n]$ is the disutility of labour. The variable n is normally interpreted as skill. With these preferences, those with higher skill (higher n) earn more. Interpreted as skill, there is no tension between optimal taxation and a horizontal equity measure based on actual or potential earnings. But, the optimal tax structure is exactly the same if n reflects the extent of dislike of work rather than skill. In this case everyone has the same potential earnings, yet those with less dislike of work earn more and are taxed more heavily.[115] If hours of work were observable, the two cases could be distinguished. If hours are not used in tax determination, does the distinction between interpretations of the variable n matter for the appeal of the calculation? Is there really a good ethical basis for treating ability to earn per hour differently from genuine dislike of working per hour?

Dislike of working may have a variety of sources, involving both physical and mental tolls from working. Reactions to chosen levels of earnings vary with the cause of the difference in earnings. Viewing a worker as lazy (liking leisure) is very different from viewing a worker as having difficulty working, perhaps for physical reasons. And some people choose lower paying jobs because of the characteristics of the jobs, which might reflect simply standard preferences (such as aversion to job stress) or might reflect

[113] It is plausible that preferences are such that those with higher wage rates have higher earnings.
[114] If all workers at each skill level have the same preferences, differences in preferences across skill levels may or may not be a problem for horizontal equity, although the degree of progressivity of an optimal tax is likely to be affected.
[115] Potential earnings are normally interpreted in terms of a budget constraint in hours-consumption space.

other concerns, such as a desire to "do good works" by working in the nonprofit sector, or perhaps pursuing a religious calling. That is, the realized relationship between earnings and earnings potential does not seem to be a sufficient statistic for a normative judgment. Should those choosing poverty for religious reasons be taxed on their abilities to earn in the commercial world? Admittedly, the presence of characteristics of jobs that are not subject to taxation (fringe benefits such as the quality of an office) along with taxation based on actual earnings implies a distortion in the choice of jobs. Perhaps these considerations would become less important if the tax code were accompanied by subsidies of certain activities – those viewed as generating externalities or particularly socially worthy in a way not captured by a standard social welfare function.[116] But then we would be choosing a complex solution, not only in taxation but also in government spending, a complexity that may be beyond the capability of the legislature.

A related issue is the time horizon to be used for considering taxable capacity – annual or lifetime or something in-between? If a lifetime perspective is taken, then the present discounted value of earnings becomes a (partial) measure both of income and consumption on a lifetime basis.[117]

In sum, given the key role played by the definition of ability to pay as the traditional starting place for discussing taxes, we do not find a convincing basis for accepting the budget set (potential outcomes) as an adequate proxy for desired taxation. Nor do we find realized earnings an adequate proxy, for pretty much the same reason viewed in reverse – sometimes the budget set is a better measure. We conclude that we can not see a good argument for adjusting taxes away from an optimal tax calculation (optimizing an evaluation of individual utilities in economic equilibrium) based on concerns drawn from budget sets, which recognize skill differences but not preferences. Nor do we see a strong case for deviating from an optimal tax calculation based on realized income or consumption. As the Meade Report put it: "But on examination 'taxable capacity' always turns out to be very difficult to define and to be a matter on which opinions will differ rather widely."

[116] One example is the forgiving of student loans for graduates taking particular jobs.
[117] This discussion ignores inheritances, which need to be considered as well, and are generally taxed separately from the income tax. Inheritance taxes are discussed in a separate chapter.

Similarly, with many skill levels and diverse preferences at each skill level, different earnings levels are reached by different workers with the same skill but different disutilities, thereby violating a measure of horizontal equity that is based on the workers' budget sets rather than the workers earnings or consumption levels. In other words, satisfying horizontal equity defined as workers with the same budget set should pay the same taxes is impossible in a sensible setting.[118] It is hard to see how to start policy analysis with a measure that is impossible to satisfy. This stance is enhanced by the difficulty of finding a good measure of how much to care about different size deviations from a measure of horizontal equity (Kaplow, 1989).

There may be tensions between tax bases thought to be ideal for horizontal equity definitions and tax bases that optimize social welfare. What if one thinks that the best measure of ability to pay is Haig-Simons income and one also accepts the empirical validity of the conditions under which the social welfare optimum involves no taxation of capital income? What if one thinks that the best measure of ability to pay is consumption expenditures and one also accepts the empirical validity of the conditions under which the social welfare optimum involves positive taxation of capital income? The weight that should be given to a chosen measure of horizontal equity in offsetting the conclusions from social welfare optimization depends on the strength of conviction that one really does have a good (usable, widely accepted) measure of horizontal equity (and sufficient strength in the belief that this consideration matters).[119] Since we do not see a really good usable measure, we do not see a good reason to lower social welfare by using horizontal equity as the starting place for policy analysis.

The end of this discussion is that we reject the Meade Report view, quoted in Part I, that taxes "should" relate monotonically to some measure of taxable capacity. In addition to finding taxable capacity not well-enough measurable and not sufficiently

[118] We focus on earnings since it makes the same point as the one with different discount rates and so different savings rates, which is the more common setting for calling for taxation that does not vary with savings levels since the budget sets are the same. We see no good basis for distinguishing between these cases.

[119] Another concern is that the choice of tax base will influence the degree of progressivity because of political behavioral effects – it is one thing to envision a consistent optimization across interacting dimensions of tax policy and another to recognize that the political process has some sequential elements.

uniformly evaluated to be usable for this purpose, we also do not see an underlying normative basis for reaching the conclusion that taxes should be related to taxable capacity without full consideration of the equilibrium consequences of following such an approach.[120] That is, we accept the view that the starting place for thinking about taxation should be the impact of taxes on the utilities of people in the economy.

2. Additional normative concerns

We begin our discussion of additional concerns by recognizing the core argument for concerns beyond a standard social welfare maximization, as stated by Musgrave in Buchanan and Musgrave (1999).

> The state and its public sector thus form an integral part of a multifaceted socioeconomic order. …
>
> That order, I hasten to add, includes not only the Pareto efficient use of resources, important though that is but also other and no less vital dimensions of social coexistence-distributive justice and the balance of individual rights and obligations upon which a meaningful concept of liberty has to be built. A view of fiscal economics, which holds that all is well if only Pareto optimality prevails, bypasses these essential components of social coexistence and fails on both normative and positive grounds. Without allowing for a sense of social justice the good society cannot be defined, and without it democratic society cannot function. Page 31-32.

It seems useful to distinguish three elements in the "fair" taxation of individuals. One, reflecting the role of individuals as ends in themselves, and not merely means to increase social welfare, calls for fair treatment of individuals in terms of some ethical basis for fairness. Following Atkinson and Stiglitz (1980), we saw this issue as influencing the allowable tax tools to be used in tax optimization. Second is the extent to which a concept of fair taxation used in tax analyses can influence government behavior, encouraging both the design of tax institutions and the implementation of policies that

[120] This conclusion is similar to that reached by some earlier economists – that equal marginal sacrifice (minimized sacrifice – equivalent to optimized social welfare) was the appropriate criterion, not equal absolute or equal proportional sacrifices. "Edgeworth, and later Pigou, held that there was no logical or intuitive choice between the equity principles of equal absolute and equal proportional sacrifice. Arguing on welfare grounds, they considered equal marginal sacrifice the only proper rule, not as a matter of equity, but because it met the welfare objective of least aggregate sacrifice." Musgrave 1959, page 98.

better satisfy social objectives. And third is the citizens' perceptions of fairness, which may or may not coincide with some philosophical concept, and which matter for both the political process and individual compliance.

Let us consider these issues in the somewhat analogous, but much starker, setting of punishment for criminal activity. First, severe punishments as deterrents, particularly in the presence of limited apprehensions of those committing crimes, may go too far, violating a sense of the proper treatment of individuals. Indeed, Amendment VIII of the US Bill of Rights states: "Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted." Similarly, taxes should not be defined differently for different people in ways that would violate the concept, somewhat slippery in this context, of "equal protection of the laws."

Second, reliance on selective enforcement and severe punishments might leave too much power to the discretion of officials deciding which alleged criminal acts are pursued in court. In the tax setting, Adam Smith argued: "The tax which each individual is bound to pay ought to be certain, and not arbitrary. … Where it is otherwise, every person subject to the tax is put more or less in the power of the tax-gatherer." (Page 778.)

And third, the perception of excessive punishment may not only violate the extent to which actions of the state should reflect the views of the citizens, but also may be self-defeating if juries are not willing to convict when they view the punishment as too severe. Similarly, taxation perceived as unfair may encourage evasion.

Tax assessments do not affect individuals as sharply as some criminal punishments, as long as tax collections are not too large relative to an individual's ability to pay. Nevertheless the same three elements are present. Consider the situation analyzed by Atkinson and Stiglitz (1976) and Stiglitz (1982b), where social welfare maximization calls for different tax treatment of two identical individuals.[121] Total reliance on social welfare function maximization would not be directly concerned by this difference in tax treatment. However, a concern for fairness would strictly prefer a truly

[121] As Atkinson and Stiglitz, 1976, Page 355 note: "If tastes are identical, the equal treatment of equals is still not necessarily implied by welfare maximization … where the feasible set is non-convex, treating otherwise identical individuals differently may increase social welfare."

random, ex ante equal probability mechanism for deciding which individual gets which tax assessment (Diamond, 1967).

But there are several concerns about such an approach. Will the implementation mechanism ensure that the randomization is done properly, avoiding improper assessments? And will individual citizens accept this approach to fairness? These issues arise even if there is sufficient information to conclude that unequal treatment is the right approach, as may or may not be the case, and even if the legislature is sufficiently sophisticated to be willing to accept and vote a suitable implementation. Randomization, as was done for the US military draft during part of the Vietnam war, might be safe from manipulation. But given the complexity and empirical uncertainty of an argument for differential treatment, we have doubts that the citizens would ever accept the underlying argument that it is better than simply levying the same taxes on those in the same circumstances. This is particularly an issue if the tax rate differences are to be long-lasting. Such a concern, assuming it is correct (without any underlying polls or focus groups) lends itself to the idea that some aspects of horizontal equity may best be addressed by viewing them as a limitation on allowable tax tools, as has been argued by Atkinson and Stiglitz (1980). We accept the view that tax tools should be limited by such considerations and that policies should be restricted to ones which are uniform over their stated tax base. And concepts and discussion of horizontal equity may help improve the political process.

3. Horizontal equity based on hypothetical alternatives[122]

A small literature addressing horizontal equity has followed from Feldstein (1976a and b), which based horizontal equity on utility rankings with and without taxes. [123] This approach is based on comparing outcomes in an existing equilibrium with

[122] This section draws particularly on Atkinson, 1980 and Kaplow, 1989.

[123] "The principle of horizontal equity in tax reform thus requires that any tax change should preserve the utility order, and should imply that if two individuals would have the same utility level if the tax remained unchanged, they should have the same utility level if the tax is altered." (Feldstein, 1976b P 124.) Feldstein recognizes that satisfying this definition of horizontal equity is not possible and thus calls for a balance between the degree of horizontal inequity and social welfare maximization.

outcomes in a hypothetical alternative. The hypothetical alternative may consider changed behavior by individuals one-at-a-time or by everyone at once, thereby incorporating general equilibrium responses.[124] The one-at-a-time approach considers what a single individual would do if that individual were exempted from taxation, with prices in equilibrium unchanged. A general equilibrium approach, including changing prices, seems particularly relevant for transition issues. Either way, horizontal equity is approached in terms of the vector of utility levels in the hypothetical alternative and the vector of utility levels in equilibrium.

As an example of this literature, Rosen (1978) considers the pattern of utilities if each person were allowed to maximize utility at equilibrium prices but without taxes. This resembles the measurement of sacrifice in sacrifice-based theories of optimal taxation (Musgrave, 1959). Rosen then looks for utility reversals between this vector of utilities and the vector in the actual equilibrium. We see no reason to give normative consequence to this particular hypothetical alternative, nor have we seen one offered.[125] And we see no reason to be particularly concerned with utility reversals in this comparison or more generally. That is, the hypothetical alternatives depend on the behavior of both the government (through expenditures) and other individuals (in determining prices). Thus it is not clear why an individual has a particular claim to protection measured from such a position, since the position depends on everyone's behavior – individuals cannot generally achieve comparable incomes on their own in a world without government expenditures and without trade with others. Indeed the taxes themselves play a role in the determination of relative prices. Moreover, there are likely to be other hypothetical alternatives that appear as normatively plausible as this one, for example the world with no taxes and no government spending - no police, no regulation of markets, etc. This would take us back to the benefit approach to taxation, which has

[124] This distinction is not as clear as appears. For example, when considering tax exempt bonds, one can recognize that the bonds would pay higher interest if taxable, relying on an arbitrage interpretation of current equilibrium prices without considering the interest rate changes that would occur in an equilibrium response to removal of the tax exemption (as, for example, in Diamond, 1965).

[125] In referring to Feldstein and the literature pursuing measures of inequity following his approach, Kaplow writes: "HE [horizontal equity] is now frequently measured and applied even though there has been virtually no exploration of why one should care about the principle in the contexts and in the manner in which it is now being used." Page 139.

suffered from an inability to make useful distributional inferences. And why those best capable of looking after themselves in some such hypothetical setting should be tax protected is not apparent.

As to giving great importance to rankings – we agree with Kaplow's (1989) criticism of such measures: "Minute movements leading to order reversals count as full violations of [horizontal equity] while substantial disturbances in the initial distribution that result in no order reversals are ignored." [Footnote omitted] (Page 141.) More generally, there is no obvious reason why rankings matter at all normatively.

Part VIII: Some Empirical Underpinnings

The discussions of the previous sections have been predominantly theoretical in nature but they have made clear that theory alone is insufficient for tax policy design. Indeed, in many cases the qualitative policy insights of the dynamic optimal tax approach outlined above depend crucially on the particular nature of some key empirical relationships. In this section we briefly consider the relevant econometric evidence on two of these relationships that crop up as recurring themes throughout our analysis. These are the nature of differences in tastes for saving across types defined by high and low earnings abilities and the degree to which different types face different earnings growth and earnings uncertainty over their lifetimes.[126] Both are areas in which recent econometric evidence, often based on data or methods that have only recently become available, means that substantially more is known about the key empirical relationships than was available to the Meade Committee. This section summarises some key findings.

To gain insights theoretical models leave out many aspects of reality. When turning to empirical evidence on the assumptions of such studies, there are two complications. One is that the empirical work can readily incorporate more elements than in the theoretical structure, indeed must do so for plausible results. But, second, the empirical work is also limited, by data availability and complexity, as to the factors that can be included. This section reviews the literatures on differences in savings rates and earnings trajectories and the extent to which one can draw conclusions from the empirical studies. Here we briefly summarize our conclusions.

There is considerable evidence across multiple countries that on average those with higher earnings potentials and those with higher earnings levels save more and accumulate more wealth during their careers, supporting the relevance of a key theoretical reason for taxing capital income. There is also considerable evidence that those males on higher earnings trajectories have steeper age-earnings profiles that peak at higher ages and after more periods in the labour market. When considering the amount of

[126] Additional empirical evidence might inform not the optimal tax structure itself, but understanding of the nature of gains and losses that would result from movements towards such tax structures given current circumstances. Examples of this might be the life-cycle evolution of the fraction of wealth held in assets with different tax treatments, which is an issue left to others in this review.

uncertainty about future earnings, a key issue is the nature of information individuals have and how it relates to the information available to the econometricians when estimating earnings models. On a strict cross section basis, there is considerable variation in earnings in each year and that variation grows with age. Some of this variation is certainly associated with different anticipated earnings tracks, anticipated from an early age, e. g., at the time education decisions are made. Indeed a considerable amount is explainable in this way. But there appears to remain a considerable degree of individual uncertainty beyond this.[127]

A. Differences in saving propensities across earnings types

Whilst the empirical evidence on differences in savings propensities across individuals of high and low earnings capacities is far from complete there are nevertheless a number of empirical studies that suggest such differences do exist and hence should be taken into account in tax design. But concrete empirical identification of differences in propensity for saving across types from economic data alone is often hindered by one (or both) of two factors. Firstly, we do not typically observe preferences directly but instead need to make inferences about preferences from data on savings or wealth outcomes. Second, the true separation of types is not known and must typically be assumed to be proxied by other observed characteristics (such as education group or social class or sometimes current or life-time income). Typically, caution is therefore required in the interpretation of evidence relating to differences across groups since these proxy characteristics are only imperfect measures of *ex-ante* earnings capacity and may indeed be partly dependent on the same intertemporal preference parameters that are under investigation. Nevertheless, in some situations the resulting biases in results can be characterised and qualitative findings may be robust to such biases.[128] Given these issues, one useful starting point is to turn to the evidence from cognitive psychology in which

[127] In addition, there is macroeconomic uncertainty about future earnings, which is not fully addressed in the literature exploring individual differences in (past) experiences, and was also not addressed in the theoretical discussion above.

[128] One pertinent example would be if more impatient individuals were less likely to choose to stay in education to older ages and if lower skill groups were on average more impatient. In this case the effects would work in the same direction and qualitative inferences regarding earnings capabilities and saving rates could be made from data on education and saving. Other situations may not be as clear cut.

recent papers have used experimental methods to examine the relationship between ability, time preference and willingness to take risks. Such studies typically use experimental designs to reveal preference measures on small groups of subjects in a laboratory environment. Some recent studies have also exploited cognitive load manipulation in the experimental design (essentially distracting subjects whilst they are taking their choices) in order to exploit within-subject variation in 'ability'.[129] Within this literature there seems to be wide acceptance that higher ability individuals are more patient (see, for example, Parker and Fischhoff (2005), Bettinger and Slonim (2005) and Kirby, Winston and Santiesteban (2005)). The relationship between risk aversion and cognitive ability is less widely studied, although what evidence there is suggests that higher ability individuals are in fact less risk averse than those of lower ability (e.g. Frederick (2005) and Benjamin, Brown and Shapiro (2006)).

The reason why higher ability may lead to lower risk aversion or more patience is not fully understood, but it seems that cognitive resources are required to make patient, risk-neutral decisions. Frederick (2005) argues that it is not just the ability to calculate expected returns correctly that leads the more intelligent to take a gamble more often. Again, using experimental data he finds that those with higher cognitive ability were more likely to take a gamble than those with lower ability even when the expected return on the gamble was lower than the safe bet.

Consideration of the issue of the extent of cognitive resources employed in decision making, however, reveals the shortcomings of such empirical evidence for our purposes since the time, effort and information deployed in making savings decisions in 'real life' situations is itself a choice variable. In contrast, such factors are strictly controlled in a laboratory experiment. As an example, individuals with lower cognitive abilities may spend more (or less) time on their savings and pensions decisions than those with higher ability, or be more likely to use various forms of advice or information in their savings decisions.[130]

[129] By increasing the cognitive load the "working memory" capacity of the brain is decreased. Since working memory capacity is almost perfectly correlated with general cognitive function, this manipulation is argued to effectively reduce cognitive ability.

[130] Lusardi (1999) and Ameriks, Caplin and Leahy (2003) both show an association between financial planning and higher financial wealth but neither study looks at differences by ability.

Conversely, higher ability (and, particularly, more numerate) individuals may be more able to process information and make complex "optimal" decisions in a less costly manner. A series of studies has explored how ability to understand and transform probabilities relates to performance on judgment and decision tasks. Peters et al (2005) summarise their evidence as showing that more numerate individuals were 'more likely to retrieve and use appropriate numerical principles, thus making themselves less susceptible to framing effects'[131] and 'tended to draw different (generally stronger or more precise) affective meaning from numbers and numerical comparisons, and their affective responses were more precise'. Numerical ability appears to matter to complex judgements and decisions in important ways although the extent to which this evidence is relevant depends on the extent to which individuals know their abilities and change their investment planning behaviour accordingly.

Given the complexity of savings and portfolio choices facing individuals in modern financial markets it is not clear that simple preference measures established in somewhat abstract experiments can adequately describe the differences in saving propensities across types that are of interest to economists. Therefore there is still considerable merit in looking at economic data on the distribution of savings outcomes across abilities, even bearing in mind the empirical difficulties discussed above. Data combining information on economic outcomes and cognitive abilities are now becoming available with which such hypotheses can be investigated. Benjamin, Brown and Shapiro (2006) use the US National Longitudinal Survey of Youth (NLSY) to look at the relationship between cognitive ability and a very crude measure of asset accumulation and find low cognitive function to be associated with low asset accumulation and financial market participation. Using more detailed data on cognitive abilities and on all components of savings of a large sample of older adults (aged 50-74) in England, Banks and Oldfield (2007) show significant correlations between the level of financial wealth and both a broad measure of cognitive functioning and a narrow measure of numerical ability based on performance in a series of simple calculations. These associations hold when both measures are used simultaneously in a model that also includes measures of

[131] A framing effect is where the interpretation of a number depends on the way in which it is presented. For example, if meat is presented as being "25% fat" or "75% fat-free".

education as well as gender and age dummies. Of course, higher cognitive abilities typically result in higher earnings and some of the literature relating to this will be discussed in section VIII.B below. What is striking, however, is the role of numeracy over and above other dimensions of cognitive abilities. To the extent that human capital is sufficiently controlled for by general measures of cognitive functioning and memory in these estimates, the role of numeracy may be thought to be indicating a separate mechanism relating to preferences for saving out of lifetime income. Finally, when it comes to portfolio decisions, Banks and Oldfield show that cognitive ability and numeracy are both associated with a higher likelihood of holding stocks and of having a private pension, even when controlling for the level of financial wealth in addition to the factors mentioned above.[132]

A variety of further evidence is beginning to emerge that relates savings choices and outcomes to the psychology of decision making, and much of that research is motivated by the view that simple preference heterogeneity in the context of a standard intertemporal economic model is not sufficient to explain certain features of observed behaviour or other outcomes. Most important, perhaps, is a rapidly expanding literature broadly relating to people's ability to exercise self-control when choosing between present and future options. Variants of this include experimental evidence on the dynamic inconsistency of choices (e.g. Ainslie (2001)), exploration of the economic implications of quasi-hyperbolic discounting models (e.g Laibson (1997)), or the modification of the underlying axioms of individuals economic preferences to allow for temptation (Gul and Pesendorfer (2004)). In each case, important implications for savings, portfolio and consumption behaviour have been demonstrated and ideally such implications would need to be considered in designing a dynamic optimal tax policy. Empirical evidence suggests that levels of self-control vary substantially within the population and are affected by cognitive load (Shiv and Fedorikhin (1999)). Additionally, those demonstrating higher self-control in early childhood (measured by experimental evaluations of young children's ability to delay gratification) have been shown to have better outcomes in a variety of economic and social dimensions in adolescence and early

[132] Lusardi and Mitchell (2007) show similar results for a broader measure of financial literacy using data from the US Health and Retirement Study.

adulthood (see Eistig et al (2006) in particular, or Borghans et al (2008) for a brief overview of the evidence). This is an area where much more needs to be known, both in terms of theoretical public finance models and relevant empirical evidence, before the full policy prescriptions with regard to the optimal taxation of capital income over the life-cycle can be assessed.[133] As such, it represents an important area for future research.

The final possibility when looking for evidence in this area is to examine studies looking at direct relationships between economic outcomes, i.e. the correlation between levels or rates of saving and levels of education, permanent income or financial wealth. As discussed in section II.A, Dynan, Skinner and Zeldes (2004) show that in a complex economic environment containing income and health uncertainty and means-tested benefits it is still the case that those with higher lifetime incomes save more than those with lower lifetime income. Carroll (2000) shows that differences in saving between the (very) rich and the poor cannot be explained by income differences alone and goes on to argue that if one rules out preference heterogeneity, the observed saving differences cannot be explained by models in which the only purpose of wealth accumulation is to finance future consumption. Evidence relevant to differences further down the wealth distribution can be obtained by looking at differences by education. Lawrence (1991) documents differences in saving rates between education groups that she argues are unexplained by differences in demographic profiles and incomes between groups and suggest a lower savings propensity in the lower education groups.[134] In all these studies, however, the rich are seen to save more than the poor, which is consistent with the preference differences between types identified above. [135]

[133] Bernheim (1997) discusses the particular problem of implications for tax incentives for retirement saving and Bernheim and Rangel (2007) provide a fine overview of the key issues for broader policy analysis.

[134] Of course, these differences may be partly due to the education itself in which case they cannot be taken as direct evidence on differences between types, although the different types will have different educations, sustaining an indirect link that may also matter for optimal taxation. Bernheim, Garrett and Maki (2001) show that high school financial curriculum mandates have long term effects on asset accumulation in adulthood.

[135] Patient households will clearly accumulate more wealth than the less patient. Furthermore, reasonable specifications for intertemporal preferences, coupled with the rates of return on risky assets that have been observed in recent years, would lead one to expect individuals with lower degrees of risk aversion to have accumulated more assets over their lifetime.

When it comes to the life-cycle profiles for saving, extensive descriptive evidence on saving profiles by age (and, where possible, age profiles within education and income or wealth groups) is available for the US, UK, Canada, Germany, Japan and Italy, in a comparative study undertaken as part of an NBER project on comparisons of household saving (see Poterba (1994)). While data limitations are substantial and the studies are far from able to identify all forms of saving, the overall messages that emerge are remarkably consistent across countries. In cross-section, saving rates are higher for those with higher income and education consistent with the studies identified above. Saving rates rise from young to middle age, often by more for high education or high income groups. Following middle age, the data show very little, if any, decline in saving rates which is on the surface somewhat puzzling.[136] Finally, median saving and financial asset holding is relatively low in all countries, indicating the importance of social security and housing for life-cycle consumption smoothing outcomes for the large majority of individuals.

At any one age, and across ages, saving propensities will ultimately depend on more than pure preference parameters alone and it would be naïve to attribute the age or education variation observed in the studies discussed above solely to differences in preferences with age. Additional determinants of saving over the life-time will be the nature of consumption needs relative to income over the life-cycle, life expectancy, access to capital markets and any possible dependency of the marginal benefit from consumption in one period on factors such as leisure or consumption in other periods, particularly if this dependency change with age. At the household level, consumption needs show a distinct hump shape over the life-cycle due to household formation, marriage and the presence of children. Other things equal, this will result in the marginal propensity to save out of current income changing systematically with age. Differences in the shape of these demographic profiles also exist for education groups – with less educated groups having more children and having them, on average, earlier in the life-cycle. Such differences, if assumed to be known in advance, lead to differences in the

[136] The exact interpretation of this in the context of life-cycle accumulation and decumulation depends on the stance one takes on the treatment of pension income and age-related decline in the present discounted value of future pension income schemes which is not explicitly addressed in the Poterba (1994) study.

shape of optimal consumption profiles over the lifetime (see, for example, Attanasio, Banks, Meghir and Weber (1999)) and hence the degree of borrowing and saving for a given income trajectory. These predicted differences are in accordance with the descriptive evidence for the UK by Banks and Blundell (1994) in the previously discussed NBER comparative study, which show that within age groups saving rates decline with family size.

At younger ages, the possibility for consumption smoothing is also determined by individual's ability to borrow. Zeldes (1989) shows that, contrary to the predictions of the consumption-smoothing model with no liquidity constraints, consumption paths track predictable changes in income for low wealth groups.

Turning to the other end of the life-cycle, substantial empirical evidence is now available on how expenditure changes with age at and after retirement, even if the connection from these results to statements about changing 'needs' is not always totally straightforward. Banks Blundell and Tanner (1998) and Bernheim, Skinner and Weinberg (2001) show falls in consumption expenditures around the time of retirement and, as briefly discussed above, data from many countries show that saving rates (defined as a ratio of total household income including pensions) remain positive, and often increase, as individuals retire and then move through their retirement. Analysis of expenditure changes for older households have also led to initial investigations into the relationship between consumption expenditures and leisure and how this might change as individuals leave paid work and as they become less healthy. Aguiar and Hurst (2005) show that individuals spend more time shopping for and preparing food after retirement, with the result that consumption of food is smoothed even though expenditure falls. Börsch-Supan and Stahl (1991) argue that a dependency on health of the utility of consumption expenditures can be shown to rationalise the fall in expenditures that is observed as households age post-retirement. Both effects would have implications for tax design to the extent that the dependencies between consumption, health and leisure, are different to those occurring at earlier ages.

One final factor relating to consumption needs is life-expectancy, as discussed earlier in section V. Ideally, for tax-design purposes we would like empirical evidence on how life-expectancy (and uncertainty in life-expectancy) varies across types defined

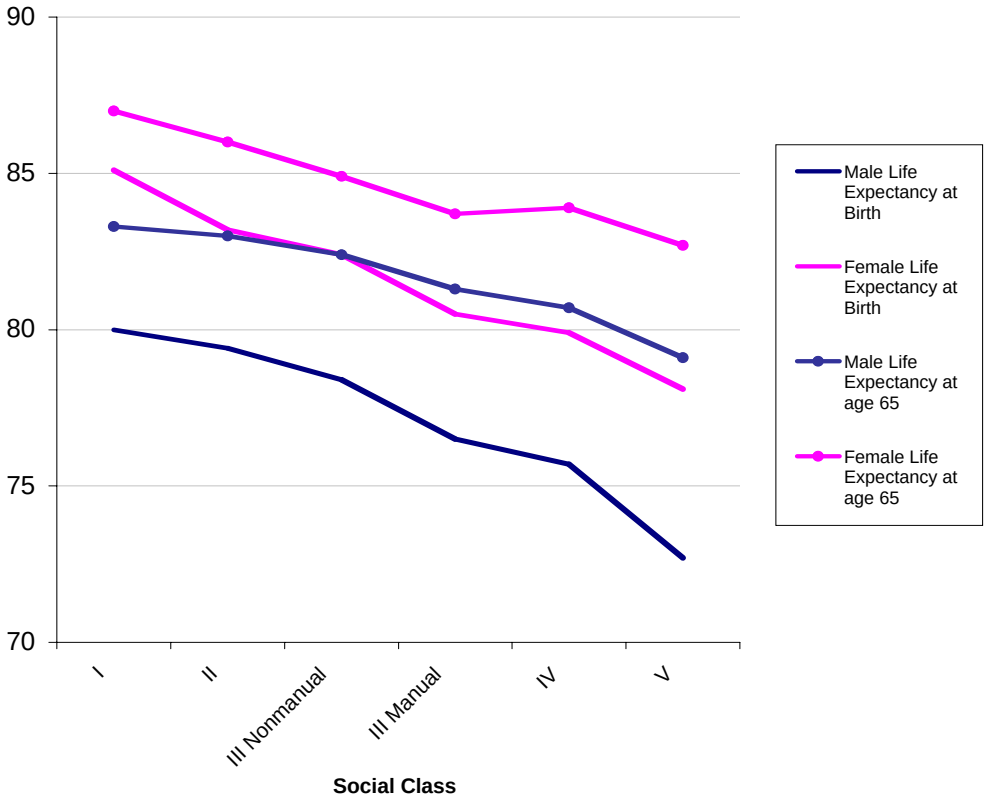
by high and low earnings capacities. Much like the debate on preferences above, we can only get an approximate understanding of this from the available data. The UK produces life tables by social class that give some indication of the extent of these effects.[137] While the variation in earnings capacity across individuals will be undoubtedly much greater than that approximated by simple social class differences, the latter will still be strongly correlated with earnings capacity, at least within cohort.

Figure 1 shows data on life-expectancy by social class in England and Wales in 2004 and displays considerable variation across groups, with the males in the lowest groups having seven years lower life expectancy at birth, and four years at age 65 than those in the highest groups. Differences of similar magnitude are observed for females. If anything, these differences have been increasing over time. Analysis of the same data as that in Figure 1 shows that between 1972–76 and 2002–05, both males and females classified to non-manual occupations had a greater increase in life expectancy at birth and at age 65 than those classified to manual occupations, although there was some narrowing of the gap in the most recent years from 1997-2001 to 2002-5 (ONS, 2007),

The reduction in life-expectancy differences between types as age increases is presumably due to a healthy-survivor effect whereby those from lower income groups that do live to older ages are a non-randomly selected set with some combination of particularly high resilience, low mortality risk factors and/or relatively good health behaviours. In contrast, for a given age, such selection is not so acute in the richer groups. The gradual erosion of life-expectancy differentials with age is important for policy design since life-expectancy at older ages, not at birth, will determine the consumption and saving behaviour of middle age and older individuals.

[137] Unfortunately similar analyses broken down by either education or wealth are unavailable in UK although a considerable body of evidence exists in the US (see, for example, Pappas et al (1993) or Preston and Elo (1995). To the extent it has value for our purposes, the use of social class as an indicator of an individual types is probably more appropriate for men than for women given it's definitional dependence on occupation. However, microdata linked to mortality records are becoming available so that analyses by education or life-time wealth could be computed in the future, at least for the case of late life life-expectancy.

Figure 1: Life-expectancy by social class in England and Wales, 2004



Source: ONS Longitudinal Study (2005)

Such socio-economic differences in length of life are also apparent when looking at mortality probabilities, where it is possible to look at outcomes by factors other than class. Attanasio and Hoynes (2000) show a strong correlation between mortality and wealth in US data and use their estimates coupled with further assumptions on wealth mobility, to correct age-saving profiles in cross-sectional data. Examination of the English Longitudinal Study of Ageing also reveals sharp differences in two-year mortality probabilities across the wealth and education distribution for those older than 50. These differences also lessen with age, at least when expressed in relative terms (see Banks et al 2006).

Considerable debate exists over the relative importance of the causal mechanisms that might be thought to underlie such differential mortality. In addition to the differences

across groups (and differences in any uncertainty surrounding these life-expectancies) tax design will also presumably depend on the precise nature of the causal processes underlying these differences. The implications for (age-related) tax systems would be different if we thought that wealth was causally driving longevity outcomes as opposed to being merely a symptom of other omitted factors (such as underlying type or ability, early life factors or even parental income and beginning of life circumstances). There is also the likely possibility that health behaviours leading to subsequent mortality risk are driven by exactly the same underlying variation in intertemporal preferences as the savings outcomes discussed earlier. Whilst much more empirical work needs to be done on the issue, at present what evidence there is suggests that increments to wealth at or after middle-age age have relatively weak effects on subsequent health and mortality once one controls for initial differences between individuals (see McFadden et al. (2003) for a test based on those aged 70 and above, and Smith (2005) for a similar test on those over age fifty. In contrast the studies investigating the effects of early life factors on subsequent mortality and morbidity seem to find much stronger results on subsequent trajectories (see, for example, Lleras-Muney (2005) for the effects of education and van den Berg, Lindeboom and Portrait (2006) for the effects of early life economic circumstances.

B. Life-cycle income profiles and permanent income uncertainty

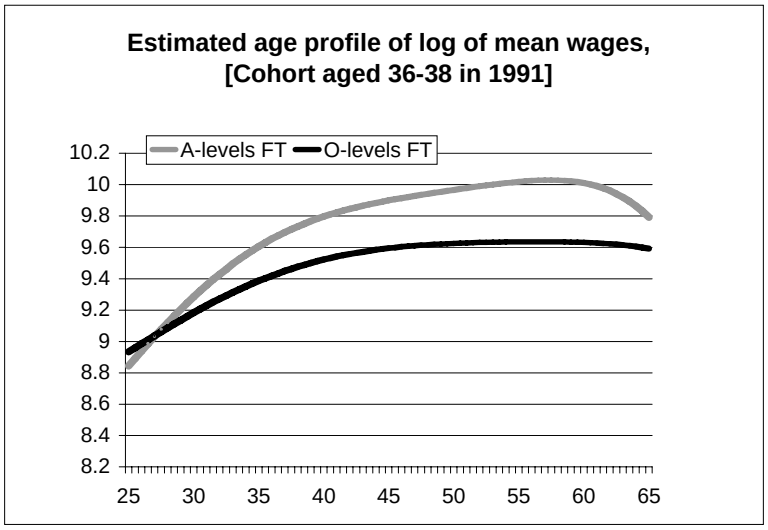
We have argued above that a second key set of empirical issues in determining optimal tax schedules are those surrounding the nature of differences in life-time earnings profiles within the population, and the degree to which such differences are correlated with skills and preferences. For our purposes three key features of the data are important. These are: the extent to which the shape of earnings or income profiles over the life-time differ by types, the extent to which uncertainty about the level of life-time earnings differs by types, and the extent to which there are systematic age-related patterns in the evolution of earnings uncertainty over the life-cycle (and, if there are, whether these age-patterns differ by type). Once again, unraveling the key lessons for the purposes of tax-design from the empirical evidence is somewhat difficult, particularly if one wants to move beyond qualitative statements. In addition to the issue, discussed above, that one

has to make assumptions to deduce the nature of underlying differences by earnings capacities from data on proxy variables such as education there are two further problems when looking at earnings profiles. Firstly, the majority of the literature has typically limited its focus to understanding the dynamics of earnings profiles for prime-age males as opposed to for all ages and both sexes. Second, when looking to understand the nature of age-profiles, investigators cannot avoid encountering the identification problem that prevents the separation of true age effects from a combination of time and generational effects without further assumptions.[138] Both of these issues need to be borne in mind when considering the available empirical evidence, and each will be referred to below.

With the increasing availability of longitudinal data on individual earnings a gradually growing body of empirical work, using the Panel Study of Income Dynamics in the US but also the British Household Panel Study and various other data sources in the UK, has begun to document earnings processes in some detail. At the crudest level, and in accordance with simple intuition, earnings for more educated households in the US have been shown to rise more steeply in early life and peak at later ages than those for less-educated households (see, for one of many examples, Attanasio, Banks, Meghir and Weber (1999)). Similar calculations from the BHPS data over the period 1991-2004, shown in Figure 2, suggest the same is true in Britain, with earnings of full-time workers basically flat for the low education group from age 40 but continuing to rise until age 58 for their high education counterparts.[139] Note that the differences between these two earnings profiles is most pronounced in early and late working life, whereas throughout mid-life (from the late thirties to the mid fifties) the growth rate of (log) earnings is only slightly steeper for the more educated group than for their less educated counterparts. This is a theme that will be returned to in our reading of the empirical evidence on earnings dynamics below.

[138] Since an individual's age can always be written as the current year minus their date-of-birth this is a fundamental problem that cannot be solved without assuming that the variation observed in data due to (at least) one of these dimensions is either zero, or at least a known function of known factors.
[139] The figure plots wage profiles for full time workers split by whether they have education up to and including O levels or equivalent – the level of schooling that is compulsory in the UK – and whether they have any more advanced educational qualifications – A levels or their equivalent and above.

Figure 2



Source: Calculations from 1991-2004 BHPS micro-data

Given that rather substantial differences emerge even when looking at two very broad skill groups, one might expect the issue to be even more striking if education or skill groups could be disaggregated even further. Ideally, one would need analysis split by a much more diverse set of skills and/or abilities, particularly at the top end where the earnings profiles of successful college graduates will likely differ quite substantially from that of the average profile for those with A levels or equivalent, both in terms of levels, growth and, potentially, variance. Lillard and Weiss (1978) provide evidence on the earnings profiles of American scientists that show considerable heterogeneity within the high skilled group and the same kind of effects appear within this group - the higher earning individuals have profiles that rise more steeply and peak later than the less high earning individuals in the group. In addition, most developed countries have displayed an increasing dispersion of incomes across skill types over the last thirty years. This widening of the returns to education (measured in terms of contemporaneous incomes) has been more acute for younger cohorts than for their older cohorts (see for example Card and Lemieux (2001)), suggesting that life-time income differences across skill groups may well increase further in the future.

It is not just the shape of earnings profiles, but also the uncertainty associated with life-time earnings, that may differ across abilities. But the empirical understanding

of the nature of such uncertainty is considerably more complicated, and depends crucially on what is assumed to be known by individuals about their life-time earnings profiles and indeed the structure assumed for the nature of 'shocks' to earnings at each age or time-period. In one important strand of the literature, the time-series of data observed on log earnings for each individual is typically thought of as being generated by a combination of three components: a known component that evolves with certainty and depends on observable covariates such as education, location and age, a random component where shocks have relatively long-lasting effects and a random component where shocks have short-term or transitory effects. Given data on a particular date-of-birth cohort the evolution of variation in each of these random components across time is then documented. As mentioned above, to assert that this variation is due to the effects of time alone would require the absence of a dependence on age, and vice versa. The key early findings come from MaCurdy (1982) and Abowd and Card (1989) who show that the above structure can indeed fit observed data on earnings over the life-cycle.

Carroll and Samwick (1997) recover levels of the variance of permanent shocks to earnings of around 0.02-0.03, but do not attempt to draw out life-cycle or temporal changes. Hubbard, Skinner and Zeldes (1994) also report similar numbers and both studies decompose the variance across education groups. In general they find a higher variance for both permanent and transitory shocks among those men without college education than for those with college education. Using the longer time-series of data available now in the US, Meghir and Pistaferri (2004) attempt on a more detailed investigation of the role of individual differences, both observed and unobserved, in the deterministic earnings growth components and in shocks to earnings. Whilst their estimation does not pin down particularly precise estimates of how the variance of shocks to either earnings or income varies over age, their point estimates at least suggest that the conditional variance of shocks to earnings is U-shaped in age, with a more pronounced pattern for the less educated groups.

Two issues of interpretation arise when considering the results from these and other related studies. The first is that results have predominantly focused on the evolution of uncertainty over time rather than over individual's life-cycles. Were one to instead focus on age-profiles (as in, for example, Deaton and Paxson (1994)), then the

dependence of such profiles on the changes happening in the macro-economy would have to be controlled for. In particular, there was a particularly strong rise in the variance of permanent shocks observed in the 1980's, documented in Moffitt and Gottshalk (1994) for the US and Dickens (2000) for the UK, that seemed to hit all cohorts whilst being most pronounced for the young. Through the 1990's this variance seems to have declined and the variance of short-term shocks to earnings has risen. Thus to ensure that MaCurdy/Abowd and Card type models continue to fit earnings data over this longer period requires allowing the variances of shocks to change over time, a fact which is confirmed by the studies cited below that exploit data on the joint evolution of consumption and earnings. But these secular changes can lead to biases in estimated age-profiles for each cohort. Heathcote, Storesletten and Violante (1994) show that the variance of wages is found to grow considerably less slowly over age if one chooses to control for year effects as opposed to cohort effects.

The key issue of interpretation in these studies of earnings dynamics relates to how differences across individuals are allowed to enter the calculations. Lillard and Weiss (1979) pointed out that if individuals faced differential trends that were not modeled in analysis then measures of the permanent uncertainty faced by individuals would overstate the true level of uncertainty faced. This has been investigated further by Baker (1997), Baker and Solon (2003) and Haider and Solon (2006), where the latter two studies exploit longitudinal income tax records to provide detailed information on the entire life-time of earnings of large samples of individuals. All three studies point to significant heterogeneity in growth rates which suggests that estimates of the importance of permanent uncertainty and its increase with age may be overstated. In addition, Haider and Solon (2006) show individual differences in trends to be most important in early and late working life which may also suggest that the finding of U-shaped permanent uncertainty may be partially due to the effect of omitted individual differences. Indeed, the nature of earnings profiles in early working life and late working life warrants further investigation more generally, since most studies of earnings dynamics focus on annual earnings of prime age males, precisely to remove any dependence of findings on issues such as the date of leaving higher education, and the timing of retirement (or other labour market withdrawal, such as that due to poor health or disability). Such issues, however,

are surely key determinants of individuals life-time resources, and will also be characterised by having an element of uncertainty. Hence, for our purposes, we would want to include their effects in an analysis of earnings uncertainty over the life-time.

Of course, the nature of such assumptions regarding what is known *ex-ante* about income processes is much more than a matter of econometric convenience. When assessing life-time uncertainty one is essentially having to make assumptions about what is known by individuals (of different types) at different stages of the life-cycle. With regard to our analysis of previous sections, whether individuals know their type is a key issue. But the nature of uncertainty about the way in which future labour markets will reward the labour supply of different types would also be a constituent factor of uncertainty even if types were perfectly known.[140] When a deterministic component of earnings and/or an average individual effect is assumed to be part of the earnings process, then econometric estimation of that component will typically rely on data across all time-periods and ages of an individual's life-time. Uncertainty, subsequently measured as deviations around this 'deterministic' component, will be understated to the extent that some of these outcomes were not anticipated by the individual at the time they were making their early-life decisions.

Consideration of this aspect brings in a second broad literature on life-time earnings processes, which addresses the question of expectations of future life-time earnings at the time schooling decisions are taken, and looks to estimate the fraction of the returns to education that can be considered known in the sense that it relates permanent and known differences between individuals (i.e. heterogeneity) and the fraction that will ultimately be due to uncertainty or luck. In an early paper on schooling decisions, Keane and Wolpin (1997) estimate that around 90% of the life-time returns to education are predictable at age 16. Cunha and Heckman (2007a) develop a different approach using test scores to identify types and then look at data on college participation decisions and subsequent earnings profiles to form estimates of the amount of life-time earnings variance that is forecastable. Their calculations for the US come up with a

[140] Taking a different modeling approach, Guvenen (2007) chooses to model a process whereby individuals gradually learn about their type and update their expectations as they move through early working life. He finds that learning is slow, and thus initial uncertainty is important throughout the life-cycle.

similar number, suggesting that around 80% of the life-time variability in returns to schooling can be viewed as forecastable by agents at age 17. Applying this methodology to changes over time, Cunha and Heckman (2007b) calculate that much of the increase in inequality for low skill groups has been due to increases in uncertainty, whereas the vast majority of the increase in inequality for high skill groups has been due to increased variation in the predictable component of earnings. In addition, around one-quarter of the increase in returns to education is calculated to be due to increases in the uncertainty component.[141]

Taken together, compared with viewing individuals as randomly drawing from the distribution of annual earnings, this latter group of studies suggest that much of the subsequent evolution of life-time earnings profiles is known by individuals at the beginning of life and there is a relatively smaller role for uncertainty than that suggested by those studies using the Permanent-Transitory methodology described above. By exactly the same argument as above, however, conclusions are inevitably highly dependent on assumptions of the nature of shocks to earnings. In this case, these studies have only studied environments where shocks are independent and identically distributed across time which rules out the existence of shocks that have persistent effects or the possibility of earnings processes where the variance of uncertainty changes with age. In both situations, were such factors to be controlled for the relative importance attributed to uncertainty would increase and the relative importance of known differences across types would decrease.

In short, the empirical literature is at a very early stage in these dimensions and as longer time series of data on larger samples of individuals become available then some of these issues should be resolved. In this respect, further research on tax record data is particularly promising. As an example, whilst the findings of Kopczuk, Saez, and Song (2007) do not directly address the issue of heterogeneity versus persistent uncertainty

[141] Finally, this literature serves to remind us that schooling decisions are themselves taken in the context of future life-time income expectations and hence education levels may only be imperfect proxies of *ex-post* earnings capabilities. Cunha, Heckman and Navarro (2005) for example, use similar calculations to show that, were the future evolution of earnings to be known in advance, one quarter of high school graduates would have chosen college education and over 30% of college graduates would have left education after high school.

described above, their related calculations on short, medium and long-run mobility in US earnings processes from 1937 onwards illustrate the potential power of such tax-record data to provide new evidence on these issues.

What is certain, however, is that the outcome of this debate will be important in generating an understanding of individual decision-making over the life-cycle, which in turn is at the heart of potential dynamic tax calculations. Some idea of the potential magnitude of the difference between the two alternative scenarios can be seen in the calculations in Scholz, Seshadri and Khitatrakun (2006) who look at the extent to which a particular and somewhat restricted form of the life-cycle model can explain the observed distribution of retirement saving in the US. Under the assumption that the life-time average of their subsequent income growth rates is known to individuals at the beginning of their life their simulations suggest that the life-cycle model can explain 86% of the variation observed in wealth data in the US. When this assumption is modified, such that individuals are assumed to only know the average of future income growth for people of their broad characteristics (defined by marital status, education and the number of earners in their household) then the fraction of saving explained by the model falls to 43%.

These latter calculations suggest consideration of an alternative approach to the understanding of life-time earnings profiles, namely to make indirect inferences about the nature of such profiles from additional data rather than study earnings data in isolation.[142] In particular, since under the standard model of economic decision making over the life-time individuals' expectations of their permanent income should be determinants of their consumption choices, data on income and expenditure have been combined to investigate the importance of permanent and transitory earnings risk. In this case, more sophisticated controls for other factors need to be introduced since consumption will typically depend on many factors other than earnings alone, such as other sources of future household income, expected taxes and transfers, and expected future household circumstances.

[142] A third alternative would be to measure individuals' income expectations using survey methods. Such measures have been pioneered in a number of dimensions in recent years and have now shown to be feasible and reliable. See, for example, Dominitz and Manski (1996, 1997, 2001) and Guiso, Jappelli and Terlizzese (1992)) for short-run income expectations and uncertainty measures and Betts (1996) or Smith and Powell (1990) for measures of longer run income expectations. The continued collection and analysis of data on long-run expectations of earnings, or more generally living standards, and in particular on uncertainty surrounding such expectations is an interesting and important avenue for future research.

Deaton and Paxson (1996) document the increasing variance of consumption with age across a wide range of countries and Blundell and Preston (1998) use data on the joint evolution of the variance of consumption and income in the UK to argue that increases in income inequality in the 1980s were predominantly due to increases in permanent uncertainty and Storesletten, Telmer and Yaron (2007) show that the increasing consumption and income dispersion is consistent with a standard life-cycle model provided that a substantial fraction (roughly half) of variability in life-time earnings is accounted for by uncertainty. Finally, for all but low wealth households, Blundell, Pistaferri and Preston (2006) find such permanent components to be the dominant factor in the evolution of the variance of consumption growth, once demographic change is allowed for. However, accounting for family labour supply behaviour, taxation and transfers, they find only around 50% of the variance in male earnings growth transmits through to variation in consumption.[143]

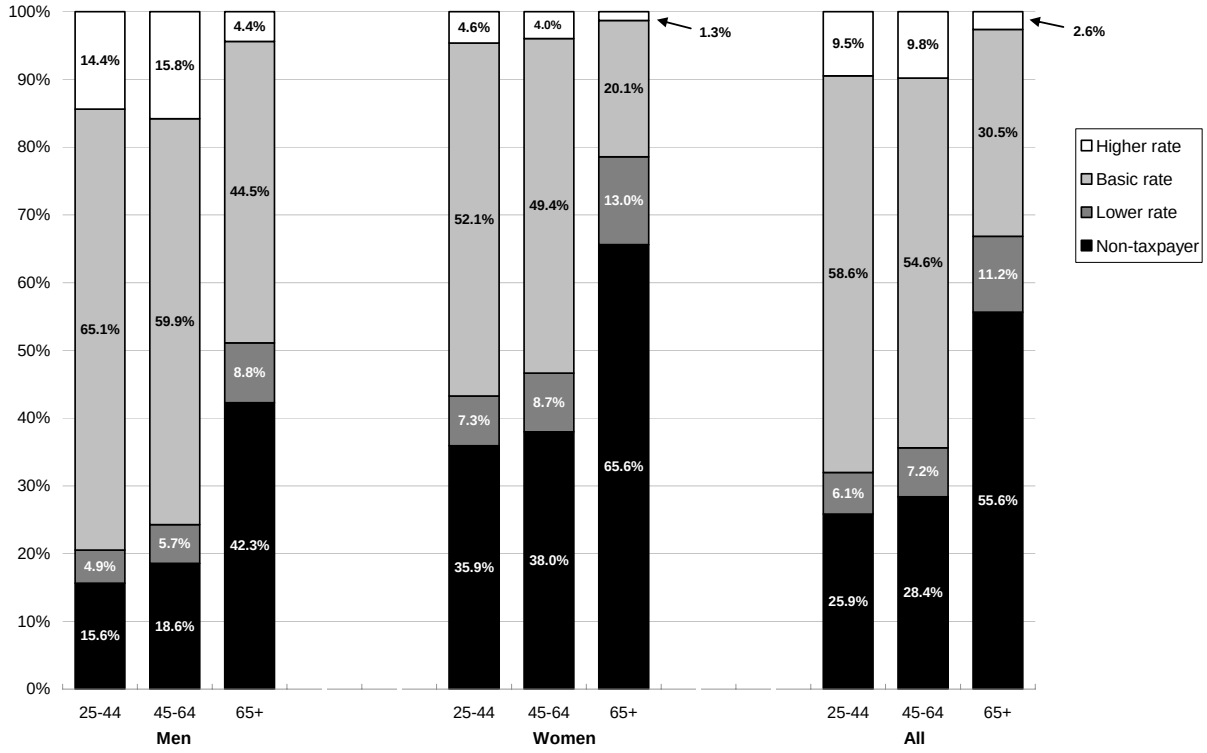
Finally, both short and long run income mobility, whether anticipated or otherwise, can create substantial movement across marginal rate tax brackets within the population and such mobility is also relevant for our discussions in previous sections. Blundell, Emmerson and Wakefield (2006) look at such tax rate mobility using BHPS data and show that, for example, 17.3% of non-higher rate income tax payers aged 30 to 39 in 1991 became higher rate income taxpayers at some point in the following 12 years and this proportion was almost one in three (32.5%) if one looked at basic rate taxpayers aged 30 to 39 in 1991. Our own calculations from the Survey of Personal Incomes (the dataset derived from tax return in the UK) show cross sectional age variation in the distribution of marginal tax rates (see Figure 3). The figure shows that whilst 15.8% of men aged 45-64 pay higher rates of income tax, only 4.4% of men aged 65 and over pay that rate.[144] Similarly only 1.3% of women aged 65 and over pay higher rates of tax and there are large fractions of the population moving from basic rates in middle age to lower

[143] Once again, such models have predominantly focused on documenting the time-series evolution of uncertainty and any such time-effects, coupled with any changes in the nature of credit markets (as argued by Krueger and Perri (2004)) would need to be accounted for when looking at age profiles.

[144] In reality, due to tapering away of tax allowances and the Pension Credit, the 'true' marginal rates may be higher than those presented in this figure for some income ranges. The marginal tax rates presented in this figure are simply statutory tax rates on income alone.

or non-taxpayer status in old age. Whilst the true cohort profiles are not captured by this age cross-section, the cohort effects in life-time incomes are unlikely to be sufficient to distort this pattern. And indeed, differential mortality along the lines discussed earlier – whereby the life-time rich are more to survive to old ages than the life-time poor – will tend to work in the opposite direction. Consequently, the opportunity for tax-rate smoothing, and the relative preference for individuals for an EET as opposed to a TEE treatment is immediately apparent.

Figure 3: Distribution of marginal income tax brackets by age and sex, UK 2004/5



Source: Authors calculations from 2004/5 Survey of Personal Incomes Microdata
Additional information on 2004 population by age and sex from ONS Population trends

C. Where do we stand?

We do not have the full empirical picture required to make precise quantitative statements about optimal tax schedules. Even for qualitative statements about the broad sign of tax wedges there is much more we could usefully know, and with the data now

available in both the US and the UK, there are many possibilities for future empirical research that addresses itself to providing estimates of the key empirical relationships required for dynamic optimal tax design. Such research would be hugely valuable and is to be encouraged. At present, what empirical findings there are come from studies whose primary focus is not the set of issues raised here and, as such, are not always specific enough to our key questions.

Nevertheless, some tentative conclusions can be drawn. In short, what matters for the design of dynamic tax policy in the models described in previous sections is the degree to which individuals are able, and willing, to smooth out any variation that they face in 'net' life-time resources over the life-time, where by 'net' resources, we mean life-time earnings adjusted for life-time needs. To the extent that individuals of higher abilities can be shown to have both higher propensity for saving (lower discount rates and lower risk aversion) and stronger earnings growth over the life cycle, and to the extent that there exist considerable uncertainties in long-run net resources (regardless of differences across ability groups), this combination of factors would lead to a role for an optimal wedge and some taxation of the normal rate of return on capital income. The role of the potential dependency of the benefit of consumption in one period on consumption, leisure and health in other period is more complex and we do not know enough about broad empirical patterns to be able to speculate on how such additional considerations would affect optimal wedges.

Part IX. Concluding remarks

The long-standing debate on the best base for nonlinear (progressive) annual individual taxation has been between total income and total consumption expenditures (with recognition of special treatment for some incomes and/or some expenditures). A more informative debate may be about the relative taxation of different sources of income and, relatedly, the implications for progressive taxation of different uses of income, with the focus here on savings, but plausibly also on medical expenses, education expenses, housing expenses, and taxation by other levels of government. We have proceeded as in the quote from Alfred Marshall at the start of this essay, "it [is] necessary for man with his limited powers to go step by step; breaking up a complex question, studying one bit at a time, and at last combining his partial solutions into a more or less complete solution of the whole riddle." (Marshall, 1948, page 366.) We have seen the implications of a wide variety of individual analyses and asked about policy inferences that seemed appropriate to draw. We do not think we have "a more or less complete solution of the whole riddle." But policy making, and so policy recommendations, cannot wait for a complete solution.

As noted at the start, the Meade Report recommends a three-part structure made up "of a new Beveridge scheme, … of a progressive expenditure tax regime, … and of a system of progressive taxation on wealth with some discrimination against inherited wealth." We have not considered issues being addressed in other chapters, particularly the role of labour force participation (the extensive margin) which is important for policy for those with very low or no earnings and limited wealth. Also, we have not explored models that might shed light on the relative advantages of annual taxation of wealth relative to taxation of capital income, as the models we have examined have mostly been restricted to a single safe asset, available on the same terms to all, leaving the two sources of taxation the same. We have had little discussion of uncertain returns to assets and none to issues related to the realization of income or the value of illiquid assets. And we have not considered bequests.

The Meade Report discussed measuring the ability to pay taxes as part of tax design. It concluded that: "on examination 'taxable capacity' always turns out to be very difficult to define and to be a matter on which opinions will differ rather widely." (Page 14.) We see no reason to reach a different conclusion from that in the Report – indeed, we have gone further in dismissing taxable capacity from a central place in tax design.

In considering the Meade Report recommendations in light of thirty years of additional research, experience, and economic development, we explored two questions that shed some light on the Meade Report recommendations –

- If there is an annual earnings tax, how should capital income be taxed?

- If there is an annual earnings tax, should there be a deduction for net savings, resulting in a tax on consumption?

In addition, we explored an issue not addressed in the Meade Report, the potential advantages, despite increased complexity, of having age-dependent income tax rates. Each of these three issues has both a design dimension and a transition dimension, but we concentrated on the former.

A. Taxation of capital income with an annual earnings tax

Support by economists and tax lawyers for exempting capital income from direct taxation has been influenced by the well-known Atkinson-Stiglitz and Chamley-Judd analyses. However, we conclude that the policy relevance of the sharp finding of the optimality of no taxation of capital income is thoroughly undercut by the implications of large uncertainty about future earnings and the growing disparity in earnings as a cohort ages. Adding such uncertainty and disparity to the frameworks employed by Atkinson-Stiglitz or Chamley-Judd results in the conclusion that taxation of capital income or of wealth is indeed part of optimal taxation. Furthermore, the full thrust of the Chamley-Judd result depends critically on bequest behavior, but behavior assumed in the model is not widespread in the population. In addition, in light of the widely varying individual savings rates in the economy, there is a natural presumption that during working years there is a positive correlation between the tendency to save and earnings potential

(although the empirical underpinning is not so clear). This is another reason for taxing capital income as a means of more efficiently taxing those with higher earnings potentials. A further case comes from the difficulties in distinguishing between labour and capital incomes, which gives an advantage to reducing the difference in taxes between them. While we have not explored the literature incorporating human capital investment into tax considerations, with a progressive earnings tax (particularly one that is not age-dependent), the presumption that human capital investment steepens the age-earnings trajectory may call for some taxation of capital income to get closer to even treatment of these two forms of investment.

Should capital income be taxed more or less heavily than labour income? With a thought process that starts with the conditions for zero taxation and then adds some taxation for elements not in the models that imply zero taxation, there is the danger of anchoring towards zero, resulting in a conclusion that capital income taxation should be lighter, without a good basis for reaching that conclusion. There is probably no substitute for extensive calculations using calibrated models, with models that incorporate the elements thought to be most important in determining relative taxation. Some existing calculations show heavier taxation while others show lighter taxation. We did not attempt to evaluate the relevance of different calculations, but point to the need for lots more.

A second issue is the appropriate relationship between the marginal taxation of capital income and the marginal taxation of labour income. The Nordic dual tax has linear taxation of capital income. The tax rate can be set at the highest or lowest positive tax rates or something in between. In the US, recent lower tax rates on dividends do relate that tax rate to the rate on labour income. The old US system that had inclusion of one-half of capital gains in taxable income (for those in lower tax brackets) also had a clear relationship. Apart from the point that trying to discourage conversion of labour income into capital income seems to call for marginal tax rates on the two types of income that relate positively to each other, it is not clear without extensive calibrated calculations how strong the relationship should be. And the choice of tax rate on capital income is plausibly related to the extent of use of tax-favored retirement savings

opportunities. To explore the normative properties of different relationships among marginal tax rates, one would again need extensive calculations. We think such calculations are called for and do not see a way to draw a firm conclusion from the evidence we have examined.

B. A deduction for active savings with an annual earnings tax

One way to have a consumption tax base is to deduct from earnings the net increase in active savings.[145] In countries such as the United Kingdom that already have EET tax-favored retirement savings accounts, this corresponds to removing limits on deposits in such accounts along with removing limits on withdrawals. Thus, compared with an accrual-based income tax (or an approximation from taxing realized capital gains to adjust for deferral), a consumption tax gives the advantage of deferral on all savings for future consumption. As Judd (1999) has pointed out, this approach does not get incentives right for human capital.

It is worth noting that there are significant differences between exempting capital income from taxation and a consumption tax base. In a model with a single safe rate of interest, the two are the same apart from differences needed in transition rules to match them. However, both different rates of return for different investors and uncertain rates of return can make the two approaches different.

C. Age-dependent taxes

Public pension systems commonly have age-dependent rules for eligibility for claiming benefits, for determination of the size of benefits, and for the implicit taxation of earnings. And Switzerland has contribution rates to the mandatory occupational pension that vary with the age of the worker. Pension systems generally have rules that have a strong reliance on individual histories over a long period in determining benefits. Income taxes make little use of such structures (apart from what is inherent in measuring capital

[145] Active savings are savings from earnings. Thus earnings minus net avtice savings equals income minus net savings.

gains). An implicit exception, similar to pension calculations, is tax-favoring of retirement savings, which incorporates explicit tax rules based on age when withdrawing funds as well as different implicit degrees of tax-favoring depending on the age at which funds are put into an account.

Is it worth the administrative complexity and the added political process to extend tax structures to include age-related features? Their presence in existing national pension rules suggests it is feasible, and analyses of optimal pension systems suggests it has value. Support for age-dependent tax rates comes from two separate arguments: differences in the distributions of circumstances across different ages and individual forward-looking calculations when making decisions. Both arguments matter, but the former may be more persuasive than the latter because of ease of measurement and the substantial diversity in individual decision-making.

Because age-dependent taxes can address both of these arguments, we think it is useful for governments to contemplate introducing them in some form and for analysts to explore them in more detail than has happened so far. We reviewed some of the support for age-dependent taxation of labour income, possibly based on setting different break points among marginal tax rates for workers in four age groups - under-30, 30-50, 50-65, and over 65. Analysis of the break points would reflect the distribution of earnings possibilities by age and the intertemporal incentives inherent in facing different break points over time. The latter might reflect uncertainties about future earnings, human capital accumulation, and borrowing constraints. This doesn't sound too hard to model and analyze, nor too hard for a legislature to incorporate in the tax structure. And plausibly this could be legislated without undue pressure by the politically better-connected ages. Obviously, any optimal tax analysis will find a higher-valued optimum from using more policy tools. The literature suggests that the gains from age-dependent labour income taxes may not be trivial and detailed analysis could explore how substantial the gains might be. There may be a case for age-varying exempt amounts of capital income as well.

Any real policy recommendation must address issues of transition. Some transition issues are lost when equity analyses look only at lifetimes of cohorts living

under a new system. Others are lost with consideration of the properties of the best steady state rather than the steady state that arises from a full intertemporal optimization.

D. Concluding remarks

The Meade Report wanted to tax both consumption and wealth annually. We share the view that capital income (or wealth) should be part of the tax base. We do not find any support in optimal tax considerations for the argument that annual capital income should be taxed exactly as annual labour income is taxed – a tax base of Haig-Simons income. We suspect that positively relating marginal tax rates on labour and capital incomes is better than having separate taxation of the two sources of income. We have also argued for the advantages of explicit variation of taxation with age. We have noted repeatedly issues that warrant further research. Pointing out the obvious need for further research is not meant to undercut the relevance of research developments to date for improving tax policy debates, and possibly tax policy.

References

Aaron, Henry J., ed. 1976. *Inflation and the Income Tax*. Washington, D.C.: Brookings Institution.

Aaron, Henry J., Len Burman and Eugene C. Steuerle (eds.). 2007. *Taxing Capital Income*. Urban Institute Press.

Abel, Andrew B. 2001. "The Effects of Investing Social Security Funds, in the Stock Market when Fixed Costs Prevent Some Households from Holding Stocks." *American Economic Review*, 91(1) March: 128-48.

Abowd, J., and D. Card (1989), "On the covariance structure of earnings and hours changes", Econometrica, 57, 411-45.

Acemoglu, Daron, Mikhail Golosov and Aleh Tsyvinski. 2007. "Political Economy of Mechanisms." http://econ-www.mit.edu/files/860. Forthcoming in *Econometrica*.

Adams, P., M. Hurd, D. McFadden, A. Merrill and T. Ribeiro, Healthy, Wealthy, and Wise?, Tests for Direct Causal Paths between Health and Socioeconomic Status, Journal of Econometrics (2003), 112, 3-56 [pdf]

Aguiar, Mark and Erik Hurst, "Consumption Versus Expenditure", Journal of Political Economy, 2005, 113(5), pp. 919-48.

Ainslie, George, 2001, *The Breakdown of Will*, Cambridge University Press

Aiyagari, S. Rao. 1995. "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting." *Journal of Political Economy*, 103(6): 1158-75.

Albanesi, Stefania and Christopher Sleet. 2006. "Dynamic Optimal Taxation with Private Information." *Review of Economic Studies*, 73(1): 1-30.

Allingham, Michael G. and Agnar Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics*, 1(3-4): 323-338.

Altig, David, Alan J. Auerbach, Lawrence J. Kotlikoff, Kent A. Smetters, and Jan Walliser. "Simulating Fundamental Tax Reform in the U.S." *American Economic Review* 91 (June 2001): 574-95.

Amador, Manuel, Werning, Iván and, George-Marios Angeletos. 2005. "Commitment Versus Flexibility." *Econometrica*, 74(2): 365-96.

Ameriks, John, Andrew Caplin and John Leahy. "Wealth Accumulation And The Propensity To Plan," Quarterly Journal of Economics, 2003, 118 (3), August, 1007-1048.

Atkeson, Andrew, V. V. Chari, and Patrick J. Kehoe. 1999. “Taxing Capital Income: A Bad Idea.” *Federal Reserve Bank of Minneapolis Quarterly Review*, 23(3): 3-17.

Atkinson, Anthony B. 1970. “On the Measurement of Inequality.” *The Journal of Economic Theory*, 2(3): 244-63.

Atkinson, Anthony B. 1980. “Horizontal Equity and the Distribution of the Tax Burden.” In *The Economics of Taxation*, ed. Henry J. Aaron and Michael J. Boskin, 3-18. Washington, D.C.: Brookings Institution.

Atkinson, Anthony B. and Agnar Sandmo. 1980. “Welfare Implications of the Taxation of Savings.” *Economic Journal*, 90 (September) 529-549.

Atkinson, Anthony B. and Joseph E. Stiglitz. 1976. “The Design of Tax Structure: Direct Versus Indirect Taxation.” *Journal of Public Economics*, 6(1-2): 55-75.

Atkinson, Anthony B. and Joseph E. Stiglitz. 1980. *Lectures on Public Economics*. London: McGraw Hill.

Attanasio, Orazio P, James Banks, Costas Meghir and Guglielmo Weber, 1999. "Humps and Bumps in Lifetime Consumption," Journal of Business & Economic Statistics, American Statistical Association, vol. 17(1), pages 22-35, January.

Attanasio, Orazio P., and Hilary W. Hoynes. 2000. "Differential Mortality and Wealth Accumulation." Journal of Human Resources 35 (1), 1-29.

Auerbach, Alan J., Laurence J. Kotlikoff and Jonathan Skinner. 1983. "The Efficiency Gains from Dynamic Tax Reform." *International Economic Review*, Vol. 24 , No. 1, pp. 81-100.

Auerbach, Alan, and Daniel Feenberg. 2000. "The Significance of Federal Taxes as Automatic Stabilizers*." Journal of Economic Perspectives*, 14(3): 37-56.

Auerbach, Alan. 1991. “Retrospective Capital Gains Taxation.” *American Economic Review*, 81(1): 167-78.

Auerbach, Alan. 2002. “Is There a Role for Discretionary Fiscal Policy?” In *Rethinking Stabilization Policy: Proceedings of a Symposium Sponsored by the Federal Reserve Bank of Kansas City*, 109-50. Kansas City: Federal Reserve Bank of Kansas City.

Auerbach, Alan. 2006. “Choice between Income and Consumption Taxes: A Primer." NBER Working Papers No. 12307.

Baker, M. (1997), "Growth-rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings", Journal of Labor Economics, 15, 338-375.

Baker, Michael and Gary Solon, Earnings Dynamics and Inequality among Canadian Men, 1976–1992: Evidence from Longitudinal Income TaRecords, Journal of Labor Economics, 2003, vol. 21, no. 2

Bankman, Joseph and David Weisbach. 2006. "The Superiority of an Ideal Consumption Tax over an Ideal Income Tax," 58 *Stanford L. Rev.* 1413.

Bankman, Joseph and Michael L. Schler. 2007. "Tax Planning under the Flat Tax," in *Taxing Capital Income*, edited by Henry J. Aaron, Leonard E. Burman, and C. Eugene Steuerle, Urban Institute Press, pp. 245-284.

Banks, James & Blundell, Richard & Tanner, Sarah, 1998. "Is There a Retirement-Savings Puzzle?," American Economic Review, American Economic Association, vol. 88(4), pages 769-88, September.

Banks, James and R. Blundell, "Household saving in the UK" in Poterba (1994), op cit

Banks, James and Zoe Oldfield, 2007, 'Understanding pensions: cognitive function, numerical ability and retirement saving', Fiscal Studies 28 (2) , 143–170

Banks, James, Elizabeth Breeze, Carli Lessof and James Nazroo, *Retirement, health and relationships of the older population in England: The 2004 English Longitudinal Study Of Ageing (Wave 2)*, Institute for Fiscal Studies, July 2006

Benjamin, D., Brown, S.A. and Shapiro, J.M., (2006) "Who is Behavioral? Cognitive Ability and Anomalous Preferences" Available at SSRN: http://ssrn.com/abstract=675264

Bernheim, B. Douglas and Garrett, Daniel M. & Maki, Dean M., 2001. "Education and saving:: The long-term effects of high school financial curriculum mandates," Journal of Public Economics, Elsevier, vol. 80(3), pages 435-465, June.

Bernheim, B. Douglas and Jonathan Skinner & Steven Weinberg, 2001. "What Accounts for the Variation in Retirement Wealth among U.S. Households?," American Economic Review, American Economic Association, vol. 91(4), pages 832-857, September.

Bernheim, B. Douglas and Antonio Rangel. 2007. "Behavioral Public Economics: Welfare and Policy Analysis with Non-Standard Decision-Makers." in *Behavioral Economics and Its Applications*, ed. Peter Diamond and Hannu Vartiainen, 7-84. NJ: Princeton University Press.

Bernheim, B. Douglas and Kyle Bagwell. 1988. "Is Everything Neutral?" *Journal of Political Economy*, 96(2): 308-38.

Bernheim, B. Douglas. 1997. "Rethinking Savings Incentives." In Alan Auerbach (ed.) *Fiscal Policy: Lessons from Economic Research*. Cambridge: MIT Press.

Beshears, John, James Choi, David Laibson and Brigitte Madrian, 2007, "The importance of default options for retirement savings outcomes: Evidence from the United States", NBER working paper 12009

Bettinger. E., and Slonim, R., (2005) "Patience among children: Evidence from a field experiment" Case Western Reserve University Working Paper, May 2005

Betts (1996), "What do students know about wages?", *Journal of Human Resources*, 31, 27-56

Blomquist, Soren, and Luca Micheletto. 2003. Age Related Optimal Income Taxation. Unpublished.

Blundell, R., and I. Preston (1998), "Consumption inequality and income uncertainty", Quarterly Journal of Economics 113, 603-640.

Blundell, R., L. Pistaferri and I. Preston (2005), "Consumption Inequality and Partial Insurance ", UCL WP04/06, forthcoming American Economic Review.

Boadway, Robin, Maurice Marchand, and Pierre Pestieau. 2000. "Redistribution with Unobservable Bequests: A Case for Taxing Capital Income." *Scandinavian Journal of Economics*, 102(2): 253-67.

Boersch-Supan, Axel and Konrad Stahl, 1991, Life-cycle savings and consumption constraints: Theory, empirical evidence and fiscal implications, Journal of Population Economics, Volume 4, 233-255.

Borghans, Lex, Angela Duckworth, James Heckman and Bas ter Weel, 'The Economics and Psychology of Personality Traits', *Journal of Human Resources*, forthcoming

Bovenberg, A. Lans, and Bas Jacobs. 2005a. "Redistribution and Education Subsidies are Siamese Twins." *Journal of Public Economics*, 89(11-12): 2005-2036.

Bovenberg, A. Lans, and Bas Jacobs. 2005b. "Human Capital and Optimal Positive Taxation of Capital Income." Tinbergen Institute Discussion Paper TI 2005-035/3

Bradford, David F. 1986. *Untangling the Income Tax*. Cambridge: Harvard University Press.

Bradford, David F. 1995. "Fixing Realization Accounting: Symmetry, Consistency and Correctness in the Taxation of Financial Instruments." *Tax Law Review*. 50,731-785.

Breyer, Stephen. 1993. *Breaking the Vicious Circle: Toward Effective Risk Regulation.* Cambridge, MA: Harvard University Press.

Buchanan, James M., and Richard A. Musgrave. 1999. *Public Finance and Public Choice*. Cambridge, MA: MIT Press.

Card, David & Thomas Lemieux, 2001. "Can Falling Supply Explain The Rising Return To College For Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, MIT Press, vol. 116(2), 705-746, May.

Carroll, Chrisopher D., 2000, "Why Do the Rich Save So Much?," in Does Atlas Shrug? The Economic Consequences of Taxing the Rich, ed. by Joel B. Slemrod. Harvard University Press,

Carroll, Christopher D. 2000. "Why Do the Rich Save So Much?" In *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, ed. Joel B. Slemrod, 463-485. Cambridge, MA: Harvard University Press.

Carroll, Christopher, and Andrew Samwick (1997) "The Nature of Precautionary Wealth", Journal of Monetary Economics, 40, No. 1, pp. 41-72.

Case, Anne and Christina Paxson. 2006. "Stature and Status: Height, Ability, and Labour Market Outcomes." NBER Working Papers No. 12466.

Chamley, Christophe. 1986. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica*, 54(3): 607-22.

Chamley, Christophe. 2001. "Capital Income Taxation, Wealth Distribution and Borrowing Constraints." *Journal of Public Economics*, 79, 55-69.

Choné, P. and G. Laroque, 2001, Optimal incentives for labour force participation, INSEE working paper No 2001-25

Choné, P. and G. Laroque, 2006, Should low skilled work be subsidized?, INSEE working paper No 2006-08

Christiansen, Vidar, and Matti Tuomala. 2007. "On taxing capital income with income shifting." Unpublished.

Conesa, Juan Carlos, Sagiri Kitao and Dirk Krueger. 2007. "Taxing Capital? Not a Bad Idea After All!" NBER Working Paper 12880.

Corlett, W. J. and D. C. Hague. 1953-54. "Complementarity and the Excess Burden of Taxation." *Review of Economic Studies*, 21(1): 21-30.

Correia, Isabel H. 1996. Should capital income be taxed in a steady state? *Journal of Public Economics*, 60, 147-51.

Cremer, Helmuth, and Firouz Gahvari. 1995. "Uncertainty, Optimal Taxation and the Direct Versus Indirect Tax Controversy." *Economic Journal*, 105, 1165-1179.

Cremer, Helmuth, and Pierre Pestieau. 2003. "Wealth Transfer Taxation: A Survey**."** CESifo working paper 1061.

Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet. 2001. "Direct Versus Indirect Taxation: The Design of the Tax Structure Revisited*." International Economic Review*, 42(3): 781-99.

Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet. 2003. "Capital Income Taxation when Inherited Wealth is Not Observable." *Journal of Public Economics*, 87(11): 2475-90.

Cunha, Flavio and James Heckman, 2008, A New Framework for the Analysis of Inequality, forthcoming, Macroeconomic Dynamics

Cunha, Flavio and James Heckman, 2008, The Evolution of Inequality, Heterogeneity, and Uncertainty in Labor Earnings in the U.S. Economy

Cunha, Flavio, James Heckman and Salvador Navarro, 2008,Separating Heterogeneity from Uncertainty in Life Cycle Earnings, Oxford Economic Papers, Vol 57(2), p. 191-261.

Deaton, A., and C. Paxson (1994), "Intertemporal choice and inequality", Journal of Political Economy, 102, 384-94.

Deaton, Angus, and Christina Paxson. 1994. "Intertemporal Choice and Inequality." *Journal of Political Economy*, 102(3): 437-67.

Deaton, Angus. 1979. "Optimally Uniform Commodity Taxes." *Economic Letters*, 2(4): 357-61.

Devereux, M. and R. Griffith, 'Taxes and the location of production: Evidence from a panel of US multinationals, *Journal of Public Economics, 1998*

Diamond, Peter and John Geanakoplos. 2003. "Social Security Investment in Equities," *American Economic Review*, 93(4) September: 1047-1074.

Diamond, Peter, editor. 1999. *Issues in Privatizing Social Security, Report of an Expert Panel of the National Academy of Social Insurance*, Cambridge: MIT Press.

Diamond, Peter. 1973. "Taxation and Public Production in a Growth Setting," in J.A. Mirrlees and N. H. Stern, eds., *Models of Economic Growth.* London: MacMillan.

Diamond, Peter. 1965. "On the Cost of Tax Exempt Bonds." *Journal of Political Economy*, 73(4): 399-403.

Diamond, Peter. 1967. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment." *Journal of Political Economy*, 75(5): 765-66.

Diamond, Peter. 1980a. "An Alternative to Steady State Comparisons," Economic Letters 5, 7-9.

Diamond, Peter. 1980b. "Income Taxation with Fixed Hours of Work," *Journal of Public Economics* 13, February, 101-110.

Diamond, Peter. 2006. "Optimal Tax Treatment of Private Contributions for Public Goods with and without Warm Glow Preferences." *Journal of Public Economics*, 90(4-5): 897-919.

Diamond, Peter. 2007. Comment on Mikhail Golosov, Aleh Tsyvinski, Iván Werning, New Dynamic Public Finance: A User's Guide, NBER Macroeconomics Annual 2006, 365-379.

Dickens, R. (2000), "The Evolution of Individual Male Earnings in Great Britain", Economic Journal, 110, pp27-49.

Dominguez, Begona. 2007. "Public Debt and Optimal Taxes Without Commitment." *Journal of Economic Theory*, 135(1) July: 159-170.

Dominitz, Jeff and Charles F. Manski, 1996, "Eliciting Student Expectations of the Returns to Schooling," Journal of Human Resources, 31, 1-26

Dominitz, Jeff and Charles F. Manski, 1997, "Using Expectations Data to Study Subjective Income Expectations," Journal of the American Statistical Association, 92, 855-867, 1997.

Dominitz, Jeff and Charles F. Manski, 2001, "Estimation of Income Expectations Models Using Expectations and Realizations Data," Journal of Econometrics, 102, 165-195, 2001.

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes. 2004. "Do the Rich Save More?" *Journal of Political Economy*. 112, 2, 397-444.

Eigsti, I., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M. B., Davidson, M. C., Aber, J. L., & Casey, B. J. (2006). Predictive cognitive control from preschool to late adolescence and young adulthood. *Psychological Science, 17,* 478-484.

Elo and Sam Preston, 1996, Educational differentials in mortality: United States, 1979-1985. Social Science and Medicine, 42:47-57.

Erosa, Andres, and Martin Gervais. 2002. "Optimal Taxation in Life-Cycle Economies." *Journal of Economic Theory*, 105(2): 338-69.

Farhi, Emmanuel and Ivan Werning. 2006. "Inequality and Social Discounting." Working paper.

Feldstein, Martin. 1976a. “On the Theory of Tax Reform.” *Journal of Public Economics*, 6(1): 77-104.

Feldstein, Martin. 1976b. “Compensation in Tax Reform.” *National Tax Journal*, 29(2): 123-30.

Fennell, Lee Anne, and Kirk J. Stark. 2005. “Taxation over Time.” UCLA School of Law: Law-Econ Research Paper 05-24.

Frederick, S., (2005), “Cognitive Reflection and Decision Making”, Journal of Economic Perspectives, Volume 19, Number 4, Fall 2005, Pages 25–42

Gaube, Thomas. 2007. “Optimum Taxation of Each Year's Income.” *Journal of Public Economic Theory*. 9(1), 127-150.

Gentry, William M. and R. Glenn Hubbard. 1997. “Distributional Implications of Introducing a Broad-Based Consumption Tax.” In *Tax Policy and the Economy,* vol. 11, ed. James M. Poterba, 1-48. Cambridge, MA: The MIT Press.

Gervais, Martin. 2003. “On the Optimality of Age-Dependent Taxes and the Progressive US Tax System.” Unpublished. http://aix1.uottawa.ca/~vbarham/PT03.pdf

Gilbert, Daniel. 2006. *Stumbling on Happiness*. Alfred A. Knopf.

Gollier, Christian. 2005. “Intergenerational Risk Sharing and Risk Taking of a Pension Fund.” No. 391, IDEI Working Papers from Institut d'Économie Industrielle, Toulouse.

Golosov, Mikhail and Aleh Tsyvinski. 2006. “Designing Optimal Disability Insurance: A Case for Asset Testing.” *Journal of Political Economy*, 114(2): 257-69.

Golosov, Mikhail, Aleh Tsyvinski, and Iván Werning. 2007. "New Dynamic Public Finance: A User's Guide." NBER Macroeconomics Annual 2006, 317-363.

Golosov, Mikhail, Narayana Kocherlakota and Aleh Tsyvinski. 2003. "Optimal Indirect and Capital Taxation." *Review of Economic Studies*, 70(3): 569-87.

Gordon, Roger H. 1985. Taxation of Corporate Capital Income: Tax Revenues vs. Tax Distortions. Quarterly Journal of Economics, Feb. 1-27.

Gordon, Roger H., and Jeffrey K. MacKie-Mason. 1995. “Why is There Corporate Taxation in a Small Open Economy?” In *The Effects of Taxation on Multinational Corporations*, ed. Martin Feldstein, James R. Hines, Jr., and R. Glenn Hubbard. Chicago: University of Chicago Press.

Gordon, Roger H., and Joel Slemrod. 1998. “Are ‘real’ responses to taxes simply income shifting between corporate and personal tax bases?” NBER Working Paper W6576.

Graetz, Michael J. 1985. "Retroactivity Revisited." *Harvard Law Review*, 98, 1820-1841.

Gul, F., and W. Pesendorfer, 2004, 'Self-control and the theory of consumption', *Econometrica*, 72:119-158

Guiso, Luigi & Jappelli, Tullio & Terlizzese, Daniele, 1992. "Earnings uncertainty and precautionary saving," Journal of Monetary Economics, Elsevier, vol. 30(2), pages 307-337, November.

Guvenen, F. (2007), "An Empirical Investigation of Labor Income Processes", mimeo, University of Minnesota, August.

Guvenen, F. (2007): "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?", American Economic Review, Vol. 97, No. 3, pp. 687-712.

Haider, S. J. (2001), "Earnings Instability and Earnings Inequality of Males in the Unisted Sates: 1971-1991", Journal of Labor Economics, 19(4), pp 799-836.

Haider, S.J. and G. Solon (2006), "Life-Cycle Variation in the Association between Current and Lifetime Earnings", American Economic Review.

Haig, Robert M. 1921. *The Federal Income Tax.* New York: Columbia University Press.

Hall, Robert E., and Alvin Rabushka. 2007. *The Flat Tax (2nd edition).* Hoover Institution Press.

Hamilton, Jonathan. 1987. "Optimal Wage and Income Taxation with Wage Uncertainty." International Economic Review 28: 373 388 (June).

Hausman, Jerry A. 1979. "Individual Discount Rates and the Purchase and Utilization of Energy-Using Durables." *The Bell Journal of Economics,* 10(1): 33-54.

Heathcote, J., K. Storesletten and G. Violante (2004), "The Macroeconomic Implications of Rising Wage Inequality in the United States", Working Paper, New York University, August.

Helliwell, John. 1969. "The Taxation of Capital Gains." *Canadian Journal of Economics,* 2(2): 314-18.

Hubbard, R Glenn, Jonathan Skinner and Steve P. Zeldes (1995) "Precautionary Saving and Social Insurance" Journal of Political Economy, Vol. 103, pp. 360-399.

Hubbard, R. Glenn, and Kenneth L. Judd. 1986. "Liquidity Constraints, Fiscal Policy, and Consumption." *Brookings Papers on Economic Activity*. (no. 1.) 1-59.

Hurd, Michael D. 1987. "Savings of the Elderly and Desired Bequests." *American Economic Review,* 77(3): 298-312.

Jones, Larry E., Rodolfo E. Manuelli and Peter E. Rossi. 1997. "On the Optimal Taxation of Capital Income." *Journal of Economic Theory*. 73 (1) 93-117.

Judd, Kenneth L. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics*, 28(1): 59-83.

Judd, Kenneth L. 1999. "Optimal taxation and spending in general competitive growth models," *Journal of Public Economics*, 71: 1-26.

Judd, Kenneth L. 2001. "The Impact of Tax Reform on Modern Dynamic Economies." In Kevin A. Hassett and R. Glenn Hubbard. *Transition Costs of Fundamental Tax Reform*. Washington: The AEI Press.

Judd, Kenneth L., and Che-Lin Su. 2005. "Optimal Income Taxation with Multidimensional Taxpayer Types." Working Paper.

Kaldor, Nicholas. 1955. *An Expenditure Tax*. London: Allen and Unwin. 1955.

Kaplow, Louis. 1989. "Horizontal Equity - Measures in Search of a Principle." *National Tax Journal*, 42(2): 139-54.

Kaplow, Louis. 1994. Taxation and Risk Taking: A General Equilibrium Perspective. *National Tax Journal*. 47, 879-798.

Kaplow, Louis. 1995. "A Note on Subsidizing Gifts." *Journal of Public Economics*, 58(3): 469-78.

Kaplow, Louis. 2006a. "On the Undesirability of Commodity Taxation Even When Income Taxation Is Not Optimal." *Journal of Public Economics*, 90(6-7): 1235-50.

Kaplow, Louis. 2006b. "Capital levies and transition to a consumption tax." NBER WP 12259

Keane, M. P. and K. I. Wolpin (1997, June). The Career Decisions of Young Men. Journal of Political Economy 105 (3), 473—522.

Kirby, K., Winston, G.C. and Santiesteban, M., (2005) "Impatience and grades: Delay discount rates correlate negatively with college gpa" Learning and Individual Differences, 15(3):213—222

Kleven, Henrik Jacobsen, Wolfram F. Richter, and Peter Birch Sørenson. 2000. "Optimal Taxation with Household Production." *Oxford Economic Papers*, 52(3): 584-94.

Kocherlakota, Narayana. 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation." *Econometrica*, 73(5): 1587-621.

Konishi, Hideo. 1995. "A Pareto-improving commodity tax reform under a smooth nonlinear income tax." *Journal of Public Economics*. 56, 413-446.

Kopczuk, Wojciech, Emmanuel Saez and Jae Song, 2007, 'Uncovering the American Dream: Inequality and Mobility in Social Security Earnings Data since 1937', NBER Working Paper 13345

Korinek, Anton, and Joseph E. Stiglitz. 2008. "Dividend taxation and intertemporal tax arbitrage." Unpublished.

Kremer, Michael. 2001. "Should Taxes be Independent of Age?" http://www.economics.harvard.edu/faculty/kremer/papers.html.

Krueger, D., and F. Perri (2006), "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory", Review of Economic Studies, 73, 163-193.

Krusell, P., L. E. Ohanian, J. V. Rios-Rull, G. L. Violante. 2000. Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*.

Laibson, David, 1997, 'Golden Eggs and Hyperbolic Discounting', *Quarterly Journal of Economics*, 112:443-77.

Laroque, Guy R. 2005. "Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof." *Economics Letters*, 87(1): 141-4.

Lawrence, Emily C., 1991, "Poverty and the Rate of Time Preference: Evidence from Panel Data," Journal of Political Economy 99 (February 1991): 54-77

Lillard, L. A. and R. J. Willis (1978, September). Dynamic Aspects of Earning Mobility. Econometrica 46(5), 985—1012.

Lillard, L.A., and Weiss, Y.A. (1979), "Components of Variation in Panel Earnings Data: American Scientists, 1960-70", Econometrica, 47, 437-454.

Lleras Muney, A. (2005), 'The Relationship Between Education and Adult Mortality in the United States', Review of Economic Studies, 72(1), 189–221. [pdf]

Lozachmeur, Jean-Marie. 2006. "Optimal Age-Specific Income Taxation." *Journal of Public Economic Theory*, 8 (4) 697-711.

Lusardi, Anna-Maria and Mitchell, O., 2007, "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth," joint with Olivia Mitchell (Wharton School, University of Pennsylvania), Journal of Monetary Economics, vol. 54, January, pp. 205-224.

Lusardi, Anna-Maria, 1999, Information, expectations and savings for retirement, in Behavioural dimensions of retirement economics, ed Henry J Aaron, Brookings Institution Press,

Luttmer, Erzo F. P., and Richard J. Zeckhauser. 2008. Schedule Selection by Agents: from Price Plans to Tax Tables." HKS Faculty Research Working Paper Series RWP08-008, (February).

MaCurdy, T. (1982), "The use of time series processes to model the error structure of earnings in a longitudinal data analysis", Journal of Econometrics, 18, 82-114.

Mankiw, N. Gregory, and Matthew Weinzierl. 2007. "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution." Unpublished. http://www.economics.harvard.edu/faculty/mankiw/files/Optimal_Taxation_of_Height.pdf

Marshall, Alfred. 1948. *Principles of Economics*, 8th ed. New York: Macmillan.

Mayshar, Joram. 1977. "Should Government Subsidize Risky Private Projects? *American Economic Review*. 67, 2, 20-28.

Meade, James Edward. 1978. *The Structure and Reform of Direct Taxation, Report of a Committee chaired by Professor J. E. Meade*. London: George Allen & Unwin.

Meghir, C., and L. Pistaferri (2004), "Income variance dynamics and heterogeneity", Econometrica, 72(1), 1-32.

Mirrlees, James. 1971. "Exploration in the Theory of Optimal Income Taxation." *Review of Economic Studies*. 38: 205-229.

Moffitt, R., and P. Gottschalk (1994), "Trends in the autocovariance structure of earnings in the US: 1969-1987", Brookings Papers on Economic Activity, 2.

Moffitt, R., and P. Gottschalk (2002), "Trends in the Transitory Variance of Earnings in the United States", Economic Journal, vol. 112(127), pages C68-C73, March.

Musgrave, Robert A.1959. *The Theory of Public Finance: A Study in Political Economy*. New York: Mcgraw-Hill.

Myles, Gareth D. 1995. *Public Economics*. UK: Cambridge University Press.

Naito, H. 1999. "Re-examination of Uniform Commodity Taxes under a Non-Linear Income Tax System and its Implications for Production Efficiency." */Journal of Public Economics/*, 71(2): 165-88.

Nielsen, Søren Bo and Peter Birch Sørensen. 1997. "On the Optimality of the Nordic System of Dual Income Taxation." *Journal of Public Economics*, 63(3): 311-29.

Office for National Statistics (2007), *Health Statistics Quarterly, 36, Winter 2007*

Pappas, Gregory, Susan Queen, Wilber Hadden, Gail Fisher (1993): The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986. New England Journal of Medicine, 329(2):103–109.

Parker, A.M., and Fischhoff, B., (2005) "Decision-making competence: External validation through an individual-differences approach" Journal of Behavioral Decision Making, 18:1—27

Pechman, Joseph A. (ed.). 1980. *What Should Be Taxed: Income of Expenditure?* Washington: Brookings Institution.

Persico, Nicola, Andrew Postlewaite, and Dan Silverman. 2004. "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height." *Journal of Political Economy*. vol. 112(5), pages 1019-1053, October.

Peters, E. , Vastfjall, D., Slovie, P., Mertz, C.K., Mazzocco and K., Dickert, S., (2005) "Numeracy and Decision Making", Psychological Science, Vol 17, Number 5

Pirttilä, Jukka, and Håkan Selin. 2007. Income Shifting within a Dual Income Tax System: Evidence from the Finnish Tax Reform of 1993. unpublished.

Poterba, James (1994) ed., "International Comparisons of Household Saving", Chicago University Press

Poterba, James. 1998. "Estate Tax Avoidance by High Net Worth Households: Why Are There So Few Tax Free Gifts?" *Journal of Private Portfolio Management*. 1. 1-9.

Preston, Sam and Elo, 1995, Are educational differentials in adult mortality increasing in the United States? Journal of Aging and Health, 7:476-496.

Reis, Catarina. 2006. "Taxation Without Commitment." Working Paper.

Reis, Catarina. 2007. "Entrepreneurial Labour Taxation." Dissertation chapter, MIT.

Rosen, Harvey S. 1978. "An Approach to the Study Of Income, Utility, and Horizontal Equity." *Quarterly Journal of Economics*, 92(2): 307-22.

Saez, Emmanuel. 2002a. "Optimal Progressive Capital Income Taxes in the Infinite Horizon Model." NBER Working Papers No. 9046.

Saez, Emmanuel. 2002b. "The Desirability of Commodity Taxation under Non-Linear Income Taxation and Heterogeneous Tastes." *Journal of Public Economics*, 83(2): 217-30.

Saez, Emmanuel. 2002c. "Optimal Income Transfer Programs: Intensive Versus Extensive Labour Supply Responses." *Quarterly Journal of Economics*, 117(3): 1039-73.

Salanie, Bernard. 2003. *The Economics of Taxation*. Cambridge, MA: The MIT Press.

Samwick, Andrew. 2006. "Saving for Retirement: Understanding the Importance of Heterogeneity." *Business Economics*. 41 (January) 21-27.

Sandmo, Agnar. 1981. "Tax evasion, labour supply and the equity-efficiency tradeoff." *Journal of Public Economics*, 16, 265-288.

Sandmo, Agnar. 2005. "The theory of tax evasion: A retrospective view," *National Tax Journal* 58, no. 4, December, 643-663.

Scholz, Karl, Seshadri and Khitatrakun, 'Are Americans saving optimally for their retirement?, Journal of Political Economy, 2006, vol. 114, No 4

Shaviro, Daniel. 2002. "Endowment and Inequality." In Joseph Thorndike and Dennis Ventry (eds.) *Tax Justice Reconsidered: The Moral and Ethical Bases of Taxation*. Washington, D.C.: Urban Institute Press.

Shiv, B., and A. Fedorikhin, 1999, 'Heart and mind in conflict: The interplay of affect and cognition in consumer decision making', *Journal of Consumer Research*, 26: 72-89.

Simons, Henry. 1938. *Personal Income Taxation.* Chicago: University of Chicago Press.

Slemrod, Joel and Shlomo Yitzhaki. 2002. Tax Avoidance, Evasion, and Administration, in A. J. Auerbach and M. Feldstein (eds.) *Handbook of Public Economics*. Volume 3. 1425-1470.

Smith, Adam. 1937. *Wealth of Nations*. New York: The Modern Library.

Smith, H., and B. Powell, (1990), "Great Expectations: Variations in Income Expectations among college seniors", *Sociology of Education*, 63, 193-207.

Smith, James P., "Unraveling the SES-Health Connection," Population and Development Review Supplement: Aging, Health and Public Policy: Demographic and Economic Perspectives, Volume 30, 2004, pp. 108-132.

Sørensen, Peter Birch. 2001. "The Nordic Dual Income Tax – In or Out." Invited speech delivered at the meeting of Working Party 2 on Fiscal Affairs, OECD 14 June 2001.

Sørensen, Peter Birch. 2005. "Dual Income Taxation: Why and How?" CESifo Working Paper 1551.

Stiglitz, Joseph. 1982b. "Utilitarianism and Horizontal Equity: The Case for Random Taxation." *Journal of Public Economics*, 18(1): 1-33.

Storesletten, Kjetl, Telmer and Yaron, 2004, Consumption and Risk Sharing over the Life Cycle (pdf). Journal of Monetary Economics, vol. 51, pp. 609-633

Summers, Lawrence H. 1981. "Capital Taxation and Accumulation in a Life Cycle Growth Model." *American Economic Review*, 71(4): 533-44.

Tarkiainen, Ritva and Matti Tuomala. 1999. “Optimal Non-Linear Income Taxation with a Two-Dimensional Population: A Computational Approach.” *Computational Economics*, 13(1): 1-16.

Tresch, Richard W. 2002. *Public Finance: A Normative Theory*. Second edition. Plano, Texas: Business Publications, Inc.

Tuomala, Matti. 1990. *Optimal Income Tax and Redistribution*. New York: Oxford University Press.

Van Den Berg, Lindeboom and Portrait (2006), “Economic Conditions in Early life and individual mortality”, American Economic Review, 96, 290-302

Vickrey, William. 1947. *Agenda for Progressive Taxation.* New York: The Ronald Press Company.

Weil, Philippe. 1989. “Overlapping Families of Infinitely-Lived Agents.” *Journal of Public Economics*, 38(2): 183-198.

Weinzierl, Matthew. 2007. “The Surprising Power of Age-Dependent Taxes.” Unpublished. http://www.people.fas.harvard.edu/~weinzier/Age-dependent%20taxes--120407--Weinzierl.pdf

Weisbach, David. 2005. "The (Non)Taxation of Risk," 58 *Tax L. Rev*.

Weisbach, David. 2006. "The Case for a Consumption Tax," 110 *Tax Notes* 1357.

Werning, Iván. 2005. “Tax Smoothing with Redistribution.” Federal Reserve Bank of Minneapolis, Staff Report 365.

Zodrow, George R. 1995. “Taxation, uncertainty and the choice of a consumption tax base.” *Journal of Public Economics*, 58, 257-265.