TAXES AND PENSIONS

Peter Diamond*

BC09-13A0409F

CRR WP 2009-12
Released: May 2009
Draft Submitted: April 2009

Center for Retirement Research at Boston College
Hovey House
140 Commonwealth Avenue
Chestnut Hill, MA 02467
Tel: 617-552-1762 Fax: 617-552-0191

* Peter Diamond is Institute Professor and Professor of Economics at the Massachusetts Institute of Technology. This was the Distinguished Guest Lecture, delivered at the Southern Economic Association's 78th Annual Meetings, November 21, 2008. An earlier version of this talk was given at the 64th Congress of the International Institute of Public Finance (IIPF), August 23, 2008. The author is grateful to Nick Barr, Jon Hamilton, Jim Poterba, and Johannes Spinnewijn for comments. The research reported herein was supported by the National Science Foundation (Award 0648741) and the Center for Retirement Research at Boston College pursuant to a grant from the U.S. Social Security Administration funded as part of the Retirement Research Consortium. The findings and conclusions are solely those of the author and should not be construed as representing the opinions or views of the Social Security Administration or any agency of the Federal Government; Massachusetts Institute of Technology; or the Center for Retirement Research at Boston College.

About the Center for Retirement Research

The *Center for Retirement Research at Boston College,* part of a consortium that includes parallel centers at the University of Michigan and the National Bureau of Economic Research, was established in 1998 through a grant from the Social Security Administration. The Center's mission is to produce first-class research and forge a strong link between the academic community and decision makers in the public and private sectors around an issue of critical importance to the nation's future. To achieve this mission, the Center sponsors a wide variety of research projects, transmits new findings to a broad audience, trains new scholars, and broadens access to valuable data sources.

Center for Retirement Research at Boston College
Hovey House
140 Commonwealth Avenue
Chestnut Hill, MA 02467
phone: 617-552-1762 fax: 617-552-0191
e-mail: crr@bc.edu
www.bc.edu/crr

Affiliated Institutions:
The Brookings Institution
Massachusetts Institute of Technology
Syracuse University
Urban Institute

Abstract

Pension benefit rules depend on individual history far more than taxes do, and age plays a much larger role in pension determination than in tax determination. Apart from some simulation studies, theoretical studies of optimal tax design typically contain neither a mandatory pension system nor the behavioral dimensions that lie behind justifications commonly offered for mandatory pensions. Conversely, optimizing models of pension design typically do not include annual taxation of labor and capital incomes. After spelling out this contrast and reviewing (and rejecting) zero taxation of capital income based on the Atkinson-Stiglitz and Chamley-Judd results, this article raises the issue of tax-favored retirement savings, a topic where the two subjects come together.

I. Introduction

When I was a young economist, I thought methodology was uninteresting and unnecessary, and just something old guys did when they didn't have anything better to do. I taught theory and public finance, and the applied theory I did was with an eye on relevance for policy questions, although I had had almost no experience with thinking about policy – just a summer job with the Council of Economic Advisors under President Kennedy. Now that I am at a stage where methodology is age-appropriate, I think it is important.

Some of this comes from the natural aging process, and some comes from my extended involvement in various policy processes, primarily about pensions, not taxes. In particular, I am concerned that too many economists take the findings of individual studies literally as a basis for policy thinking, rather than seeking inferences from an individual study to be combined with inferences from other studies that consider other aspects of a policy question, as well as with intuitions about aspects of policy that are not in the models. To me, taking a model literally is not taking the model seriously. It is worth remembering that models are incomplete – indeed that is what it means to be a model. We construct multiple models to highlight different aspects of an issue, so, thinking thoroughly about policy calls for thinking through multiple models, and requires recognizing issues that have not made it into any of the available models. My focus here is on the connection between basic research and policy advice, particularly basic theoretical research. The argument for using multiple models to gain insight and understanding is not new, and was stated clearly by Alfred Marshall.[1] Previous research (Banks and Diamond 2009) considered methodology more thoroughly as part of considering the taxation of capital income from the perspective of alternative theoretical models. This article draws on that essay, after contrasting tax policies and public pension rules, along with the normative modeling of the two. This contrast struck me when thinking back on some of the differences between the tax paper and the book on pensions (Barr and Diamond 2008) being written at the same time.

II. Policy

Contrasting pensions with taxes on earnings, two elements stand out – (i) pension benefit determination depends on individual history far more than taxes do and (ii) age plays a much larger role in pension determination than in tax determination. Pension benefits are typically related to a lot of an individual's history, for example, the best 35 years of indexed earnings in the United States, and sometimes a complete history is taken into account (as in Germany and Sweden, for example). This holds for earnings-related pensions, both defined benefit and defined contribution. Even non-contributory pensions typically depend on years of residence. For example, the Dutch National Old Age Pension (AOW) gives a full pension on the basis of 50 years of residence between ages 15 and 65; and is reduced proportionally for years of nonresidence. The New Zealand Superannuation is subject to 10 years' residency after age of 20 and at least five years' residency after age of 50. A full Swedish Guarantee pension is available after 40 years residence in Sweden after age 25; also with proportional reduction for fewer years of residence. The Guarantee pension is reduced based on 18/16.5 times the benefits received from Sweden's notional defined contribution (NDC) pension, the Inkomstpension, which, in the nature of NDCs, is based on lifetime covered earnings.[2] In contrast, taxation of earnings is focused on earnings within a single year, although some averaging over a few years has sometimes been allowed (and capital gains taxes depend on a cost basis from the time of acquisition).

As for the role of age, not only do pension rules vary significantly by age, but also the age-related rules often vary by date of birth. Taking the United States as an example, retirement benefits can be claimed after age 62, but not before. Retirement benefit claims are subject to an earnings test before the age for full benefits, but not after. And the monthly benefit for a given earnings history depends on the gap between the age at which the benefits start and the age for full benefits. In contrast, age plays some, but a small role in earnings taxation of adults. For example, in the United States there is an additional standard deduction amount ($1,050 in 2008) for a taxpayer over 65. In the United Kingdom, the personal allowance of £6,035 (for the 2008-09 tax year) becomes £9,030 for those 65-74 and £9,180 for those 75 and over (but subject to an income limit).

The age for full Social Security benefits in the United States is in transition from 65 to 67, varying with date of birth (see Table 1). Similarly shifting age rules by date of birth have occurred with pension reforms in other countries. This is consistent with the common expression that a good pension system should not be significantly adjusted too often (beyond its automatic indexing) and should be changed with enough lead time for workers to adjust their voluntary retirement savings. In contrast, legislated tax changes often vary by year.

Pension systems use indexing to limit the frequency of needing to adjust rules. There is widespread indexing to prices and/or wages and, in some systems, for a life expectancy measure (NDC systems as in Sweden) or for a dependency ratio (as in Germany). Moreover, the indexing might work differently for workers with different dates of birth. In the United States, wage indexing of earlier earnings up to the year of turning 60 implies that the wage indexing is done differently for workers with different birth years.[3] On the tax side there is indexing of bracket end points for prices in the United States, but no adjustment for how inflation hits capital and labor incomes differently (Diamond 1975). Table 2 identifies four aspects of differences between pension and tax policies.

Interestingly, there have been recent calls for significant variation of earnings taxes with age in contrast with the minor variations that sometimes exist.[4] An age-varying tax structure appears administratively feasible and does not add an undue complexity to compliance and enforcement in advanced countries. And it does not appear to violate intuitive fairness measures, although the transition to such a system might raise some issues of intergenerational fairness. Note that these issues, administration, complexity, and perceived fairness, are missing in the typical model of equilibrium used for tax analyses. Yet they matter for making use of the insights from those models. I favor greatly expanding analyses of how age-varying earnings taxes might be done, but that is not the subject of this article.

Vickrey (1947) on income averaging not withstanding, a considerably larger reliance on earnings histories for earnings taxation, much less lifetime reliance, as is

common with pensions, appears to go strongly against the grain of the history of discussion of income taxation.[5] For example, Adam Smith (1937) writes of basing taxation on revenue, with no mention of a longer time span.[6] And two centuries later the Meade Report (1978) viewed taxable capacity as the starting place for income taxation[7] and discussed the competition between total income (Schanz-Haig-Simons income[8]) and consumption as the better measure – again considering annual measures, although arguing that consumption reflects lifetime considerations.

As noted above, a good pension system is thought to be significantly adjusted infrequently (beyond its automatic indexing) and changed with enough lead time for workers to adjust their voluntary retirement savings. No one says anything like that about annual budget expenditures. These are expected to adjust to developments on a nearly continuous basis, for example with the outbreak of a war or risk of a recurrence of the Great Depression. And adjusting taxes along with spending is seen as important for the politics of spending and taxing, as well as part of a sensible response to changes in a country's economic, political and spending-needs environments. Yet, considerable continuity is considered good policy. The Meade Report (1978) calls for taxes that reflect a concern for both flexibility and stability:

> "A good tax structure must be flexible … 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." Meade, 1978, p. 21

An interesting question to muse on is why these policy institutions are so different – and I have not gone beyond musing. Complexity of the world and of analyses makes it natural to approach these areas separately. Whether thought of in terms of politics or in terms of policy analysis, "framing" seems to be a key issue in how these areas have

developed. How one starts thinking about an issue can affect how one finishes thinking about an issue (anchoring). Thinking about tax policy starts as thinking about revenue needs in the short term, recognizing that revision of spending and taxes is expected in the following year, and substantial revision may occur after the next election. While thinking about pensions includes concerns about the current benefit recipients, the focus is on rules that affect current workers (as both taxpayers and future benefit recipients) as well as current beneficiaries. And the political process in the United States has been designed to incorporate long-run concerns through annual reporting of 75-year projections and legislative rules that tend to separate Social Security legislation from the annual budget cycle. The link between benefits and previous earnings subject to tax affects perceptions of fairness and political legitimacy. While the annual spending and taxation process has great inertia, this comes more from the political process than from an underlying argument that the process should have great inertia. Although some of the support for transition rules, including grandfathering, argues for a legitimate role for some inertia. Pensions are focused on a single long-run concern, acquiring adequate retirement income, while stability in tax policy matters for a large and diverse set of decisions where "reasonably far-sighted plans" matter.

Empirical work on decisions such as retirement saving and retirement timing naturally include both earnings taxes and pension rules. However, in parallel with the policy differences, theoretical analyses in these two areas also differ. Apart from some simulation studies, theoretical studies of optimal tax design typically contain neither a mandatory pension system nor the behavioral dimensions that lie behind justifications commonly offered for mandatory pensions. Conversely, optimizing models of pension design typically do not include annual taxation of labor and capital incomes. Recognizing the presence of two sets of policy institutions raises the issue of whether normative analysis should be done separately or as a single overarching optimization. Or, as I believe, there should be both types of analyses as sources of insight into practical policy issues. Just as complexity in issues being addressed by legislation calls for considering different programs separately, with some concern for coordinating, so too does complexity in models call for separate and joint studies.

Consideration primarily of a shorter time horizon in tax policy, for example in Mirrlees (1971) with a single period, makes it more comfortable to work primarily in the context of consistent, rational choice.[9] Pension design addresses long time horizons and, in contrast to the discussion of taxes, mandatory pension plans are justified primarily by an apparent failure, for a significant fraction of the population, of consistent, rational choice, in the form of a life-cycle model, to be an adequate description. In addition to any possible reason from shortcomings in the life-cycle model, the focus of many policy questions on a shorter time horizon than lifetimes may help explain the focus of normative tax analyses on short periods. I think there might be interesting ideas coming from exploring implications of a reluctance to rely too strongly on standard lifetime individual models when considering annual government taxes and spending. But I have not started on such considerations.

III. Capital Income Taxation

Tax and pensions issues become intertwined when we consider taxing capital income and tax-favoring retirement savings. To touch on some connections, I want to start by briefly going over the discussion of capital income taxation in Banks and Diamond. That essay starts with the policy question of how capital income should be taxed. (Table 3 about here) The focus of the essay was the process of drawing inferences from the existing literature to help answer this question. Our bottom line was that neither zero taxation nor taxing total income were supported by the weight of theoretical analyses. We inclined toward relating marginal tax rates to each other in light of the ability of some people, particularly the self-employed and executives, to convert labor income into capital income and vice versa. Since then, Johannes Spinnewijn and I (2009) have analyzed a simple model of work and retirement where optimal taxation calls for taxing the capital income of high earners and subsidizing it for low earners, as can be done within the rules for tax-favored retirement saving.[10]

Atkinson-Stiglitz

My starting place for thinking about taxing capital income is the Atkinson-Stiglitz theorem (1976). Consider a model with two periods, with labor supply in the first period and consumption in both the first and second periods. Suppressing a role for taxing initial wealth, savings from first-period earnings are used to finance second-period consumption and so generate capital income that is taxable (in the second period).[11] With only a single period of work, the model is about the taxation of savings for retirement. The well-known Atkinson-Stiglitz theorem is that when the available tax tools include nonlinear earnings taxes, there should be zero differential taxation of first- and second-period consumption (no "wedge" between the intertemporal marginal rate of substitution [MRS] and the intertemporal marginal rate of transformation [MRT] between consumer goods in different periods) if two key conditions are satisfied: (i) all consumers have preferences that are separable between consumption and labor, and (ii) all consumers have the same sub-utility function of consumption, $u^h\left[x_1, x_2, z\right] = \tilde{u}^h\left[\phi\left[x_1, x_2\right], z\right]$, where x_1, x_2 are consumption levels and z is earnings. The first condition states that the intertemporal marginal rate of substitution of consumption does not depend on labor supply. And the second requires all consumers to be the same in their interest in smoothing consumption across their life-cycle.

The theorem extends to having many periods of consumption with a single period of labor. It also extends to multiple periods of earnings provided lifetime taxation can be a general function of the earnings in all periods. An interesting extension (Kaplow 2006; Konishi 1995; Laroque 2005) is that for *any* earned income tax function, given the same preference assumptions, moving from distortionary consumption taxes to non-distorting consumption taxes can be coupled with a change in the earnings tax in order to have a Pareto gain.

Before arguing for zero capital income taxation on the basis of the theorem, it is appropriate to consider the robustness of the result relative to our understanding of the workings of the economy (see Table 4). With non-separability between consumption and labor, from the Corlett-Hague (1953) analysis, a key issue for the sign of taxing capital

income – taxing versus subsidizing - is the pattern of the cross-elasticities between labor supply and consumption levels in the two periods. However, not much is known about these cross-elasticities and thus there is not a good reason from this argument to reject the zero tax policy implication.

With separability preserved, a second consideration would be that the subutility functions of consumption are not the same for everyone. Saez (2002) shows that the Atkinson-Stiglitz theorem does not generally hold with differences in discount rates, and, therefore, desired savings rates, across individuals with different skills. Saez argues that it is plausible that there is a positive correlation between labor skill level (wage rate) and the savings rate and cites some supporting evidence. Banks and Diamond reviews more of the evidence on individual savings. Saez provides a condition to sign the preferred direction of introduced taxation of capital income. Diamond and Spinnewijn (2009) builds on this analysis, using a model with jobs, rather than choice of hours by workers facing a given wage rate. In a four-types model (two wage rates and two discount factors) they show that starting with the optimal earnings tax, introduction of a small tax on savings of high earners raises social welfare, as does introduction of a small subsidy on savings of low earners. Both introductions ease the binding incentive compatibility constraint. The result makes no use of the correlation across types, although it does assume that at the optimum all higher skilled workers hold the higher output job. With a restriction on preferences, they also show that the optimal linear earnings-varying savings tax has the same character. And Tenhunen and Tuomala (2009) calculate the mechanism design optimum with the usual labor market and find implicit marginal taxation of savings for one high skill person and implicit marginal subsidization of savings for one low skill person for all but the highest correlations.

Uncertain future earnings

While the natural way to consider uncertain future earnings[12] is in a two-period model with both work and consumption in both periods, the basic point can be made in a model with work only in the second-period. The key assumption is that a consumption decision is made before the individual's second-period wage is known. In the Atkinson-

Stiglitz model, a worker knows full life-time income before doing any consumption. But when consumption decisions are taken before earnings uncertainties are resolved the Atkinson-Stiglitz result does not hold. With earnings occurring only in the second period, first-period consumption is chosen before the uncertainty about future earnings is resolved. In this model, second-period consumption should be taxed at the margin relative to first-period consumption. 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. Indeed, in this case we get an inverse Euler equation:

$$\frac{1}{u'^{h}[x_1]} = \int \left(\frac{1}{(\delta / p_2) u'^{h}[x_2[w]]} \right) dF[w]$$

where w is the random second-period wage and p_2 is the price of second-period consumption. This implies implicit marginal taxation of savings:

$$u'^{h}[x_1] < \int \left((\delta / p_2) u'^{h}[x_2[w]] \right) dF[w]$$

With uncertain (future) wage rates, the government would like to provide insurance by lowering after-tax earnings in the event of high wages in order to raise after-tax earnings in the event of low wages. With asymmetric information the government is inferring wage rates from earnings and is limited by the ability of someone with a high wage rate to choose low earnings nevertheless. The incentive compatibility constraint is that those with high wage rates must find it in their interest to work harder and earn a higher amount. But, a worker intending to earn a low amount despite a high wage rate has a higher valuation of savings than if the worker were planning to earn a high amount (assuming normality of consumption). Thus taxing savings eases the incentive compatibility constraint by making it less attractive to work less in the future. 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 later retirement and permits more generous pensions for those who need to retire early and so have lower accumulated lifetime earnings.

This result has appeared in the pension literature as part of design of a pension system to recognize that some workers lose good earnings opportunities while others do not. To provide lifetime earnings insurance, the encouragement for delayed retirement should be less than fully actuarial, implying an implicit tax on continued work. In the setting of providing insurance in this way, discouraging savings is part of providing insurance more efficiently. This result appears in models of pension design that have no income taxes, so it is not clear how it would carry over, if at all, in models that also have standard annual taxation of earnings, not just lifetime taxes.

I want to pass quickly through the other arguments I have identified as blocking the Atkinson-Stiglitz result. Standard modeling assumes perfect observation of capital and labor incomes. This omits the ability of some workers, particularly the self-employed, to legally transform labor income into capital income (and vice versa). Pirttilä and Selin (2007) found significant shifts of labor 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. On a more widespread basis, labor effort devoted to earning a higher return on savings also represents a shifting from labor income to capital income. Christiansen and Tuomala (2007) examine a model with costly (but legal) conversion of labor 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. Gordon and Slemrod (1998) raise the issue of shifting between corporate and personal tax bases. Even with an infinite horizon, the Chamley-Judd result of asymptotically zero capital income taxation does not hold in a model with an inability to distinguish between entrepreneurial labor income and capital income (Correia 1996; Reis 2007).

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 labor income that do not vary with age (Aiyagari 1995; Chamley 2001; Hubbard and Judd 1986).

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 because 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 overlapping generations model that incorporates the different ways that people think about bequests.[13]

Beyond these arguments, 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 generally 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 labor supplies, lifetime after-tax consumption spending can depend in a nonlinear way on both first-period and second-period earnings including an interaction term. Once one envisions modeling longer lives, this degree of interaction becomes implausible to implement in a general form. 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. The result of zero taxation of capital income does not seem to extend to this setting. [14]

Chamley-Judd

The thinking of the profession about taxing capital income has been strongly influenced by the work of Chamley (1986) and Judd (1985) showing the optimality of an asymptotically zero taxation of capital income. In these models, workers (really dynasties) have an infinite horizon for their savings decisions and, in the long run, each

period is exactly the same. Judd (1999a) allows greater generality in the evolution of the economy than in the original Chamley and Judd models. He obtains the result that the *average* capital income tax tends to zero even if it is not zero in any period. 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. Distinguishing between capital income and bequest taxes, if one were taxing capital income during lifetimes, as argued for above, then, in order to have a long-run convergence to a zero average tax on capital income, one would be subsidizing bequests. With this formulation, analysis is focused, appropriately, on the analysis of bequest motives. 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. The literature on bequests does not offer a ringing endorsement of this model. Indeed, it is unclear how important bequest considerations are for behavior, and bequest considerations appear to be widely varying in the population. Thus I 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, since it relies critically on bequest behavior that does not seem to be supported empirically.

Nevertheless the issue remains that the compounding of annual taxation of capital income results in a growing tax wedge as savings are accumulated to finance consumption at later dates in the future – a point also made in models with finite lives of many periods.[15] As has been noted in Judd (1999b), such a pattern of taxation of consumption in different years is unlikely to be optimal if a more general tax structure were available. A starting place for thinking about taxing capital income over many years is to note the relationship between MRS and MRT if there is a constant tax rate on capital income, τ. The ratio between the MRS and MRT between consumption today and consumption T periods from now is $\left\{\left(1+\left(1-\tau\right)r\right)/\left(1+r\right)\right\}^{T}$. 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 as shown in Table 5.

Comparing the table to a tax on labor earnings makes several points. A 30% tax on earnings puts a 30% wedge between contemporaneous earnings and consumption. As the right-hand column of table 5 shows, a 30% tax on capital income puts only a 3% wedge between consumption today and consumption in a year (when the rate of return is 10%). But it puts a 67% 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. Table 5 makes it 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, annual capital income taxation at a constant rate that impacts distant consumption will be inefficient.

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.[16] And it points to a potential welfare gain from tax-favoring retirement savings, since retirement saving tends to be for longer times. Also, the role of capital income taxation when future earnings are uncertain suggests that capital income tax rules might well be different for those at ages when people are mostly retired, a common feature of tax-favored retirement accounts.

Tax-favored retirement income

The focus of this article has been on the comparison and interaction of taxes and voluntary retirement savings. But we should not lose sight of the presence of and role for mandatory programs that provide retirement income. Standard arguments for that role are shown in Table 6. Across countries, such programs vary greatly in the replacement rates they provide. The tax treatment of retirement savings is important, particularly in countries with smaller programs, like the United States, and in countries, like Germany, that are reducing the replacement rate in their mandatory program and encouraging more voluntary pension savings.

In light of the arguments for taxing capital income, and the problems raised by the compounding of annual tax rates, one can see a case for tax-favoring retirement savings. While one can withdraw balances from a retirement savings account, they are subject to a penalty. Perhaps the penalty should decline with the length of time the funds were in the account. It is also the case that someone doing precautionary savings and not hitting a zero balance will have held funds for a long time. But the motivation is different than with retirement savings. It would be good to see modeling of taxes with both concerns. There is also a behavioral reason for considering tax-favored retirement income since there may be a greater impact on savings than with a general reduction in taxation of capital income.[17] Thus, we have the question of how voluntary pensions should be taxed, something on which there is little literature, with the common structures listed in Table 7. How this tax favoring should be done is an important issue that I flag as needing research rather than offering an answer.

IV. Conclusion

I conclude by repeating my call to avoid over-reliance on any single model and with the usual researchers' plea for more research. In particular, I think we have done too little study of the issues around tax-favored retirement savings accounts, studies that need to recognize uncertainty in future earnings, uncertainty in future spending needs, diversity in savings behavior and earnings opportunities, and uncertainty about future tax rates.

References

Aiyagari, S. Rao. 1995. Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy* 103(6):1158-75.

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.

Banks, James, and Peter Diamond. 2009. The tax base. In *Dimensions of tax design: The Mirrlees review*, edited by James Mirrlees et al. Oxford: Oxford University Press for Institute for Fiscal Studies.

Barr, Nicholas, and Peter Diamond. 2008. *Reforming pensions: Principles and applications*. Oxford: Oxford University Press.

Beshears, John, James Choi, David Laibson, and Brigitte Madrian. 2008. The importance of default options for retirement savings outcomes: Evidence from the United States. In *Lessons from pension reform in the Americas*, edited by Stephen J. Kay and Tapen Sinha. Oxford: Oxford University Press, pp. 59-87.

Blomquist, Soren, and Luca Micheletto. 2008. Age related optimal income taxation. *Scandinavian Journal of Economics* 110:45-71.

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.

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.

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

Corlett, W. J. and Douglas Hague. 1953. 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* 105:1165-79.

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.

Diamond, Peter. 1975. Inflation and the comprehensive tax base. *Journal of Public Economics* 4:227-44; correction *Journal of Public Economics* 1981:16(1):129.

Diamond, Peter, and James Mirrlees. 1978. A model of social insurance with variable retirement. *Journal of Public Economics* 10:295-336.

Diamond, Peter, and James Mirrlees. 1982. Social insurance with variable retirement and private saving. MIT Working Paper No. 296, and forthcoming, *Journal of Public Economics*.

Diamond, Peter, and James Mirrlees. 1986. Payroll-tax financed Social Security with variable retirement. *Scandinavian Journal of Economics* 88:25-50.

Diamond, Peter, and Johannes Spinnewijn. 2009. Capital income taxes with heterogeneous discount rates. MIT Working Paper.

Erosa, Andres, and Martin Gervais. 2002. Optimal taxation in life-cycle economies. *Journal of Economic Theory* 105(2):338-69.

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

Gaube, Thomas. 2007. Optimum taxation of each year's income. *Journal of Public Economic Theory* 9(1):127-50.

Gervais, Martin. 2003. On the optimality of age-dependent taxes and the progressive U.S. tax system. Unpublished. Available online at http://aix1.uottawa.ca/~vbarham/PT03.pdf.

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. In *NBER Macroeconomics Annual 2006*, edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford. Cambridge, MA: The MIT Press, pg. 317-63.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski. 2003. Optimal indirect and capital taxation. *Review of Economic Studies* 70(3):569-87.

Gordon, Roger H., and Joel Slemrod. 1998. Are 'real' responses to taxes simply income shifting between corporate and personal tax bases? NBER Working Paper No. W6576.

Haig, Robert M. 1921. *The federal income tax*. New York: Columbia University Press.

Hubbard, R. Glenn, and Kenneth L. Judd. 1986. Liquidity constraints, fiscal policy, and consumption. *Brookings Papers on Economic Activity* 1:1-59.

Judd, Kenneth L. 1985. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics* 28(1):59-83.

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

Judd, Kenneth L. 1999b. "The Impact of Tax Reform in Modern Dynamic Economies." In Kevin A Hassett and R. Glenn Hubbard (eds) *Transition Costs of Fundamental Tax Reform*. Washington DC: AEI Press.

Kaplow, Louis. 2006. On the undesirability of commodity taxation even when income taxation is not optimal. *Journal of Public Economics* 90:1235-50.

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* 56:413-46.

Kremer, Michael. 2001. Should taxes be independent of age? Unpublished, available online at http://www.economics.harvard.edu/faculty/kremer/papers_kremer.

Laroque, Guy R. 2005. Indirect taxation is superfluous under separability and taste homogeneity: A simple proof. *Economics Letters* 87(1):141-4.

Lozachmeur, Jean-Marie. 2006. Optimal age-specific income taxation. *Journal of Public Economic Theory* 8(4):697-711.

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

Meade, James Edward. 1978. *The structure and reform of direct taxation*. London: \Allen & Unwin.

Mirrlees, James. 1971. Exploration in the theory of optimal income taxation. *Review of Economic Studies* 38:205-29.

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.

Pirttilä, Jukka, and Håkan Selin. 2007. Income shifting within a dual income tax system: Evidence from the Finnish tax reform of 1993. Unpublished.

Reis, Catarina. 2007. Entrepreneurial labor taxation. Dissertation chapter, MIT.

Rogerson, William. 1985. Repeated moral hazard. *Econometrica* 53:69-76.

Saez, Emmanuel. 2002. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics* 83(2):217-30.

Schanz, Georg von. 1896. Der einkommensbegriff und die einkommensteuergesetze. *Finanzarchiv* 13:1-87.

Simons, Henry. 1938. *Personal income taxation*. Chicago: University of Chicago Press.

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

Tenhunen, Sanna, and Matti Tuomala. 2009. On optimal lifetime redistribution policy. Forthcoming, *Journal of Public Economic Theory.*

Vickrey, William. 1947. *Agenda for progressive taxation*. New York: The Ronald Press Company.

Weinzierl, Matthew. 2007. The Surprising Power of Age-Dependent Taxes. Unpublished.

Table 1. Age To Receive Full Social Security Benefits*

Year of Birth**	Full Retirement Age
1937 or earlier	65
1938	65 and 2 months
1939	65 and 4 months
1940	65 and 6 months
1941	65 and 8 months
1942	65 and 10 months
1943-1954	66
1955	66 and 2 months
1956	66 and 4 months
1957	66 and 6 months
1958	66 and 8 months
1959	66 and 10 months
1960 and later	67

*Also called “Full Retirement Age" or “Normal Retirement Age"
**If you were born on January 1st of any year you should refer to the previous year.

Table 2. Contrast between Pensions and Taxes

Pension Benefits	Income Taxation of Earnings
Dependent on a long history	Focus on a single year
Dependent on age	Little variation with age
Dependent on date of birth	Varies by year
Indexed for prices and/or wages, demography	Limited indexing for prices

Table 3. Approaches to Taxing Capital Income

If there is annual non-linear taxation of earnings, how should annual capital income be taxed?

- not at all
- linearly (Nordic dual income tax)
- relating the marginal tax rates of capital and labor incomes (United States)

– taxing all income the same (Schanz-Haig-Simons)

Table 4. Models where Atkinson-Stiglitz theorem does not hold

– nonseparable preferences
– nonuniform preferences
– uncertain future earnings
– hard to distinguish capital income from entrepreneurial earnings
– borrowing constraints
– different initial wealths
– limited tax tools

Table 5. Ratio of MRS to MRT - $\left\{\left(1+\left(1-\tau\right)r\right)/\left(1+r\right)\right\}^{T}$.

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

Table 6. Reasons for a Mandatory Retirement Pension System

– free riding with non-optimal taxes
– too little savings
– poor investing
– too little annuitization: individual and joint-life
– absence of age and history related tax rules

Table 7. Approaches to Tax Favoring Retirement Savings

– Exempt contributions, Exempt accumulations, Taxable withdrawals (EET): IRA

– Taxable contributions, Exempt accumulations, Exempt withdrawals (TEE): Roth IRA
– Both EET and TEE available (United States)
– Exempt contributions, Partially exempt accumulations, Taxable withdrawals (Denmark, Italy, Sweden)

[1] "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." Marshall 1948, p. 366.

[2] The ratio equals the total contribution rate (including that to the funded defined contribution account) relative to contributions to the Inkomstpension.

[3] For someone who turned 60 in 2001, earnings in 1980 are multiplied by the ratio of the wage index in 2001 to that in 1980. For someone who turned 60 in 2002, the wage index ratio used is based on 2002. Price indexing also differs, starting after they have turned 61.

[4] Recent analyses of age-dependent taxes include Blomquist and Micheletto (2008); Erosa and Gervais (2002); Fennell and Stark (2005); Gervais (2003); Kremer (2001);

Lozachmeur (2006); and Weinzierl (2007). This issue is discussed in Banks and Diamond (2009).

[5] Vickrey (1947) was concerned with the impact of progressive annual taxes on those with fluctuating incomes relative to those with constant incomes. He discussed averaging of total income, not just earnings, over different lengths of time. Using a longer period for determining taxes is likely to reduce the built-in-stabilization from the income tax and lessen the easing of borrowing constraints.

[6] "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. The expence of government to the individuals of a great nation, is like the expence of management to the joint tenants of a great estate, who are all obliged to contribute in proportion to their respective interests in the estate." Smith 1937, p. 777.

[7] "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." Meade 1978, p. 14.

[8] Schanz (1896); Haig (1921); Simons (1938).

[9] Some earnings decisions, involving career concerns, on-the-job training and education have an intertemporal aspect. But this has not altered the short focus in taxing earnings.

[10] Kocherlakota (2005) provides an argument for regressive earnings-varying wealth taxation. He analyzes a model with asymmetric information about stochastically evolving

skills, which is not present in Diamond and Spinnewijn (2009). On the other hand, see Nielsen and Sørensen (1997) on the optimality of the Nordic dual income tax.

[11] With only safe assets, this can be considered taxation of savings.

[12] Articles examining uncertain future earnings include Cremer and Gahvari (1995); Diamond and Mirrlees (1978, 1982, 1986); Golosov, Kocherlakota and Tsyvinski (2003); Golosov and Tsyvinski (2006); Golosov, Tsyvinski and Werning (2007); Rogerson (1985).

[13] For models with varying initial wealth, see Boadway, Marchand and Pestieau (2000) and Cremer, Pestieau and Rochet (2003).

[14] Erosa and Gervais (2002) have examined the most efficient taxation of a representative consumer (Ramsey taxation) with intertemporally additive preferences in an overlapping generations 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.

[15] Taxation of capital gains does not involve this compounding. In light of the absence of such compounding, it is not clear what basis there is for lower taxation of realized capital gains after a longer holding period. Among the key issues in capital income taxation are the relative treatment of dividends, interest and capital gains and the role of corporate income taxation.

[16] I do not discuss the alternative approach of progressive annual consumption taxes. Analysis of such taxes has been limited thus far.

[17] See, for example, Beshears et al (2008).