Capital Income Taxes with Heterogeneous Discount Rates*

Peter Diamond[†]   Johannes Spinnewijn[‡]

MIT   MIT

March 21, 2009

Abstract

With heterogeneity in both skills and preferences for the future, the Atkinson-Stiglitz result that savings should not be taxed with optimal taxation of earnings does not hold. Empirical evidence shows that on average people with higher skills save at higher rates. Saez (2002) suggests that with such positive correlation taxing savings can increase welfare. This paper analyzes this issue in a model with less than perfect correlation between ability and preference for the future. To have multiple types at the same earnings level, the number of types of jobs in the economy is restricted. Key to the analysis is that types who value future consumption less are more tempted to switch to the low earning job. We show that introducing both a small savings tax on the high earners and a small savings subsidy on the low earners increase welfare, regardless of the correlation between ability and preferences for the future. However, a uniform savings tax, as in the Nordic dual income tax, increases welfare only if that correlation is sufficiently high.

1 Introduction

The Atkinson-Stiglitz (1976) theorem shows that when the available tax tools include nonlinear earnings taxes, optimal taxation is inconsistent with taxing savings when two key assumptions are satisfied: (1) that all consumers have preferences that are separable between consumption and labor and (2) that all consumers have the same sub-utility function of consumption. Empirical evidence suggests that on average those with higher skills save at higher rates (Dynan, Skinner and Zeldes, 2004, Banks and Diamond, 2008). We therefore relax the second condition and analyze the taxation

of savings with heterogeneity in both skill and savings propensity. We consider both uniform and earnings-varying taxation of savings.

This paper uses a simple model in which the number of types of jobs in the economy is restricted. This sheds light on the desirability of earnings-dependent taxes on savings and the role of the positive correlation between skill and savings propensity. The paper provides an argument for making the taxation of savings progressive in earnings. We find that the savings of the high earners should be taxed, whereas the savings of the low earners should be subsidized. This result is independent of the correlation between ability and discount factors, provided that the optimum has the high skilled workers on the more productive job. A uniform savings tax, however, only increases welfare if that correlation is sufficiently high.

Our paper builds on the analysis in Saez (2002). He derives conditions on endogenous variables to sign the effect on social welfare of introducing a uniform commodity tax or a subsidy, when consumers have heterogeneous sub-utility functions of consumption. With an optimal non-linear earnings tax, a small tax on savings increases welfare if the net marginal social value is negatively correlated with savings, conditional on earnings, or if on average those who choose to earn less save less than those who choose to earn more, when restricted to the same earnings. By restricting the number of types of jobs, we analyze the importance of the (exogenous) correlation between skills and savings preferences for the taxation of savings.

Primary attention is focused on a model with four worker types - with two discount factors and two skill levels. Thus we are examining a particular example of a multidimensional screening problem. The model assumes the existence of two jobs, rather than the standard model where each worker can select the number of hours to be worked.[1] This results in a setting where workers with the same skill but different discount factors choose the same job and so have the same earnings. With the introduction of earnings-related savings tax rates, they are subject to the same tax rates. We assume that at the optimum both high-skill types work at the high-skill job and that redistribution from high earners to low earners is the important redistribution. Given these assumptions social welfare increases with the introduction of a tax on the savings of high earners and with the introduction of a subsidy on the savings of low earners. The relative frequencies of the four types in the population plays no role in the derivation of this result, conditional on the assumed structure of the optimum.

The underlying assumption is that those valuing the future more are more willing to work than those valuing the future less, conditional on the disutility of work. This means that an incentive compatibility (IC) constraint just binding on a high skill worker with low value for the future is not binding on a high skill worker with high value for the future. Earnings-dependent taxes and subsidies on savings allow an increase in

---

[1] A limited number of jobs was assumed in Diamond (2006), although here we allow the planner to set hours at each job.

redistribution by targeting types in a given job with saving preferences different than those of types who are tempted to switch jobs. In particular, introducing taxation of savings of high earners (and transferring the revenue back equally to all high earners) eases the binding IC constraint since it transfers resources from the high saver to the low saver for whom the IC constraint is binding. Introducing a subsidy on savings for low earners (financed by equal taxation on all low earners) also eases the binding IC constraint by making switching to the lower job less attractive to the high earner with low savings. We explore robustness in the two jobs setting by considering models with a continuum of types of one aspect (discount factor or skill) while preserving the presence of two types of the other aspect. In this case the results are dependent on properties of the distribution of types. The case for taxing the savings of high earners appears to be more robust than the case for subsidizing the savings of low earners.[2]

The assumption that those with less discounting of the future are more willing to work is in line with standard modeling, representing preferences by $u(x) + \delta_i u(c) - v(z/n_i)$. An alternative specification $\frac{1}{\delta_i} u(x) + u(c) - v(z/n_i)$ would imply the exact opposite. That is, types with higher δ_i prefer to save more, but to work less. We examine some empirical support for our assumption, using data from the Survey of Consumer Finances (SCF). We find that conditional on education and age, people with higher discount factors tend to work more. To proxy for the discount factor, we use reported savings and the time horizon people report having in mind when making spending and savings decision. We also use these proxies to revisit the positive correlation between skills and savings propensities.

This paper contributes to the literature on the optimal choice of the tax base and the joint taxation of labor and capital incomes in particular. Banks and Diamond (2008) review the literature on the inclusion of capital income in the tax base. Gordon (2004) and Gordon and Kopczuk (2008) argue that capital income reveals information about earnings ability and thus should be included in the tax base. Blomquist and Christiansen (2008) analyze how people with different skills and different preferences for leisure who cannot be separated with an income tax, may be separated with a commodity tax. The four-types model with hours chosen by workers has been studied by Tenhunen and Tuomala (2008), which calculates a set of examples, but explores the analytics only in two- and three-type models. They consider both welfarist and paternalist objective functions. We relate the results in their calculated examples to some of our results below. We focus on the four-types model since the result in a two-types model, while striking, does not seem relevant for policy inferences.[3]

The paper is organized as follows. Section 2 sets up the model with four types

and two jobs. Section 3 characterizes respectively the first best and the restricted first best, referring to an 'equal job, equal pay' restriction. Section 4 introduces incentive compatibility constraints and characterizes the second best with the introduction of earnings-varying savings tax rates. Section 5 considers the introduction of a uniform savings tax, rather than one varying with earnings level. Section 6 considers models with a continuum of skill levels with two discount factors and vice versa. Section 7 considers the optimal linear savings tax, rather than the introduction of a small savings tax. For comparison, Section 8 reviews a two-types model. Section 9 discusses empirical support for the assumptions and Section 10 has concluding remarks.

2   Model

We consider a model with two periods. Agents consume in both periods, but work only in the first period. Preferences are assumed to be separable over time and between consumption and work. Denoting first period consumption by x, second period consumption by c, and earnings by z, preferences satisfy

$$U(x, c, z) = u(x) + \delta u(c) - v(z/n),$$

with $u' > 0, u'' < 0$ and $v' > 0, v'' > 0$. An agent's ability n determines the disutility of producing output z. An agent's preference for future consumption depends on the discount factor δ.

We consider heterogeneity in both ability n and preference for future consumption δ. Although robust insights for optimal taxation have been derived in models with two types, considering heterogeneity in two parameters in a model with two types implies perfect correlation between the two parameters. The inference based on a simple two-types economy, although simple, may therefore be misleading. In order to allow for imperfect correlation, we consider a four-types model. We denote the four types by ll, lh, hl, hh with frequencies f_{ij} and welfare-weights η_{ij}. The first two types have low ability n_l, but differ in discount factors δ_l and δ_h. The second two types have high ability n_h and also differ in discount factors δ_l and δ_h. We have

	high discount factor δ_h	low discount factor δ_l
high ability n_h	**hh**	**hl**
low ability n_l	**lh**	**ll**

with $n_h > n_l$ and $\delta_h > \delta_l$.

There are only two jobs in the economy, h and l. The output from a job is independent of the worker's type, while the disutility of holding a job varies with ability. The low-ability types can only hold the low job. The high-ability types can hold either

4

job. We assume that redistribution to the low-skilled types is sufficiently important and the type mix sufficiently balanced that all high-skilled workers hold high-skilled jobs at the various optima analyzed. This requires a restriction on the weights in the social welfare functions and the population distribution, which we do not explore.

We begin with the first best, which differs from the usual treatment in that the output produced on a job must be the same for everyone holding the job. We assume a linear technology. The first best has the property that there is no marginal taxation of savings. Then we consider a *restricted first best* (the term 'first best' refers to a lack of incentive compatibility constraints, the term 'restricted' means limited tax tools, but not limited by IC constraints) with zero taxation of savings and the requirement that everyone holding a job receives the same pay (no taxes based on identity, only on potential earnings). We note whether social welfare can be improved by taxing or subsidizing savings in some form.

We then turn to the second best, with taxes based on earnings, not potential earnings, so that there is an incentive compatibility constraint. We assume a zero taxation of savings restriction, thus preserving the condition of equal pay for equal work. Again we ask about potential gains from taxing or subsidizing savings.

3   First Best

In the first best, the social welfare function is maximized with respect to the type-specific consumption levels in the first and second periods and the job-specific output level, subject to a resource constraint. Each worker is assigned to the matching job,

$$\text{Maximize}_{x,c,z} \quad \sum f_{ij}\eta_{ij}\left(u\left[x_{ij}\right] + \delta_j u\left[c_{ij}\right] - v\left[z_i/n_i\right]\right)$$

$$\text{subject to:} \qquad E + \sum f_{ij}\left(x_{ij} + R^{-1}c_{ij} - z_i\right) \leq 0$$

(1)

Forming a Lagrangian, we have

$$\mathcal{L} = \sum_{i,j} f_{ij}\eta_{ij}\left(u\left[x_{ij}\right] + \delta_j u\left[c_{ij}\right] - v\left[z_i/n_i\right]\right) - \lambda \sum_{i,j} f_{ij}\left(x_{ij} + R^{-1}c_{ij} - z_i\right)$$

With the welfare weight of type ij of η_{ij}, we define the net marginal social value of first period consumption for an individual of type ij as

$$g_{ij} \equiv \eta_{ij}u'\left[x_{ij}\right] - \lambda.$$

Along the relevant portion of the social welfare optima, we have the following properties:

$$g_{ij} = 0 \text{ and } u'\left[x_{ij}\right] = \delta_j R u'\left[c_{ij}\right] \text{ for all } i,j, \text{ and}$$

$$(f_{il}\eta_{il} + f_{ih}\eta_{ih}) \frac{v'\left[z_i/n_i\right]}{n_i} = (f_{il} + f_{ih})\lambda = f_{il}\eta_{il}u'\left[x_{il}\right] + f_{ih}\eta_{ih}u'\left[x_{ih}\right],$$

for both the high-skilled and the low-skilled jobs. The net marginal social value of first period consumption for each type equals 0 and the saving of each type is undistorted. Given that the required output for a given job is independent of an individual's type, the earnings are marginally distorted upward for one discount-factor type and downward for the other type, unless the welfare weights satisfy $\eta_{il} = \eta_{ih}$ since, otherwise, $u'\left[x_{il}\right] \neq u'\left[x_{ih}\right]$. The output is undistorted 'on average' though.

3.1 Restricted First Best with no Taxation of Savings

If the (after-tax) earnings on a job, y_i, is restricted to be type-independent and savings can not be taxed, there are further constraints which we approach using the indirect utility-of-consumption function. This function, $w_j\left[y, R\right]$, satisfies

$$w_j\left[y, R\right] \equiv \quad \max u\left[x\right] + \delta_j u\left[c\right]$$

$$\text{subject to} \quad x + R^{-1}c = y.$$

We continue to assume that the welfare weights and population fractions are such that all high skilled are on the more productive job at the optimum. The restricted first best solves the following problem,

$$\text{Maximize}_{y,z} \quad \sum f_{ij}\eta_{ij}\left(w_j\left[y_i, R\right] - v\left[z_i/n_i\right]\right)$$

(2)

$$\text{subject to:} \quad E + \sum f_{ij}\left(y_i - z_i\right) \leq 0$$

Forming a Lagrangian, we have

$$\mathcal{L} = \sum_{i,j} f_{ij}\eta_{ij}\left(w_j\left[y_i, R\right] - v\left[z_i/n_i\right]\right) - \lambda \sum_{i,j} f_{ij}\left(y_i - z_i\right).$$

The first order conditions are

$$\sum_{j} f_{ij}\eta_{ij}u'\left[x_{ij}\right] \quad = \quad \lambda \sum_{j} f_{ij} \text{ for } i = h, l,$$

$$\sum_{j} f_{ij}\eta_{ij}v'\left[z_i/n_i\right]/n_i \quad = \quad \lambda \sum_{j} f_{ij} \text{ for } i = h, l.$$

Recalling the definition of the net marginal social utility, $g_{ij} \equiv \eta_{ij}u'\left[x_{ij}\right] - \lambda$, the population-weighted values add to zero at each job

$$\sum_{j} f_{ij}g_{ij} = 0 \text{ for } i = h, l.$$

Thus, the welfare weights determine the direction of desired redistribution (given the equal pay condition) between workers on each job. Also, in the absence of savings taxation,

$$u' [x_{ij}] = \delta_j R u' [c_{ij}] \text{ for all } i, j.$$

The FOC for job outputs, z_i, are the same as given above.

3.2 Restricted First Best with Small Earnings-Dependent Savings Taxes

Given the observability of earnings, small linear taxes on savings (collected in the first period) could be set differently for high and low earners. This can be implemented by the rules on retirement savings accounts (IRA, 401(k)). The (local) desire to redistribute can be met by a small linear tax or subsidy on savings by types on a given job with the revenues returned equally to them by raising net-of-tax earnings on the job.[4]

Differentiating the Lagrangian with respect to a savings tax rate, τ_i, on those with earnings level y_i, evaluated at a zero tax level:

$$\frac{\partial \mathcal{L}}{\partial \tau_i} = \lambda \left(\sum_j f_{ij} (y_i - x_{ij}) \right) - \sum_j f_{ij} \eta_{ij} u' [x_{ij}] (y_i - x_{ij}).$$

The impact of a savings tax on the Lagrangian is made up of two pieces: the impact on the revenue constraint and the impact on utilities. Using the FOC, multiplied by y_i, the derivative can be written as:

$$\frac{\partial \mathcal{L}}{\partial \tau_i} = \sum_j f_{ij} \left(\eta_{ij} u' [x_{ij}] - \lambda \right) x_{ij} = \sum_j f_{ij} g_{ij} x_{ij}.$$

Recall that

$$\sum_j f_{ij} g_{ij} = 0 \text{ for } i = h, l \text{ and } x_{ih} < x_{il}, c_{ih} > c_{il}.$$

This implies that a tax on the savings by the two types on a given job increases welfare if the savings of the one type towards which redistribution is desirable saves sufficiently little compared to the other type.

The welfare weights imply the desired direction of redistribution within productivity types and so the signs of g_{ij}. With equal incomes and different discount factors, we have

$$x_{ih} < x_{il}$$
$$c_{ih} > c_{il}$$

---

[4] In section 7 we briefly consider how the sign of the welfare impact of a small savings tax is related to the sign of the optimal linear tax.

Thus, if first period utilities get the same weights for both types, $\eta_{il} = \eta_{ih}$, $g_{il} < 0 < g_{ih}$, implying a desire to redistribute to the high saver. In contrast, if second period utilities get the same weights for both types, $\eta_{il}\delta_l = \eta_{ih}\delta_h$, the signs are reversed, implying a desire to redistribute to the low saver. If there is no desire to redistribute for high (low) skill types we have $\eta_{hh}u'[x_{hh}] = \eta_{hl}u'[x_{hl}]$ $(\eta_{lh}u'[x_{lh}] = \eta_{ll}u'[x_{ll}])$. In general, with uniform weights for given discount factors, $\eta_{hi} = \eta_{li}$, we do not satisfy both conditions.

4 Second Best

We draw a distinction between restricted first-best analyses and second-best ones based on the absence or presence of IC constraints involving taking a job with lower productivity (the reverse having been ruled out by assumption). That is, the distinction depends on the observability of productivity. The prime issue in second-best analyses is determining which IC constraints are binding. We start with the further restriction, as above, that savings not be taxed. With no taxation of savings, second-best implies equal pay for equal work and the IC constraint of not imitating the other discount rate type who is holding the same job does not bind. Similarly, if a high productivity worker were to take the low productivity job, the person imitated would be the one with the same discount factor.

We add the critical assumption that earnings distribution issues are sufficiently important that at the second-best optimum (with IC constraints) the net marginal social value of first period consumption $g_{ij} \equiv \eta_{ij}u'[x_{ij}] - \lambda$ is negative at the earnings level of a high-skill job and positive at the earnings level of a low-skilled job. Without an IC constraint, this condition could not hold at the optimum as noted above.

Assumption 1 *The net marginal social values of first period consumption satisfy*

$$g_{hj} < 0, g_{lj} > 0, \text{ for } j = h, l.$$

4.1 Second best with No Taxation of Savings

We assume that the Pareto-weights and population fractions are such that all high-skilled workers work at the high-skilled job and the desired level of redistribution to lower earners is sufficient that at least one IC constraint is binding.

$$\text{Maximize}_{y,z} \quad \sum f_{ij}\eta_{ij}\left(w_j\left[y_i, R\right] - v\left[z_i/n_i\right]\right)$$

$$\text{subject to:} \quad E + \sum f_{ij}\left(y_i - z_i\right) \leq 0 \tag{3}$$
$$w_h\left[y_h, R\right] - v\left[z_h/n_h\right] \geq w_h\left[y_l, R\right] - v\left[z_l/n_h\right]$$
$$w_l\left[y_h, R\right] - v\left[z_h/n_h\right] \geq w_l\left[y_l, R\right] - v\left[z_l/n_h\right].$$

Forming a Lagrangian, and assuming that at the optimum each worker is assigned to the matching job, we have

$$\mathcal{L} = \sum_{i,j} f_{ij}\eta_{ij}\left(w_j\left[y_i, R\right] - v\left[z_i/n_i\right]\right) - \lambda \sum_{i,j} f_{ij}\left(y_i - z_i\right)$$
$$+ \sum_j \mu_j\left(w_j\left[y_h, R\right] - v\left[z_h/n_h\right] - w_j\left[y_l, R\right] + v\left[z_l/n_h\right]\right).$$

Since the first-period consumption of type hj if switching to the low job equals the first-period consumption of type lj, the FOC with respect to earnings are

$$\sum_j f_{hj}\eta_{hj}u'\left[x_{hj}\right] + \sum_j \mu_j u'\left[x_{hj}\right] - \lambda \sum_j f_{hj} = 0,$$
$$\sum_j f_{lj}\eta_{lj}u'\left[x_{lj}\right] - \sum_j \mu_j u'\left[x_{lj}\right] - \lambda \sum_j f_{lj} = 0.$$

Given the definition of the net social utility $g_{ij} \equiv \eta_{ij}u'\left[x_{ij}\right] - \lambda$, this implies

$$\sum_j f_{hj}g_{hj} = -\sum_j \mu_j u'\left[x_{hj}\right] < 0,$$
$$\sum_j f_{lj}g_{lj} = \sum_j \mu_j u'\left[x_{lj}\right] > 0.$$

The population-weighted values add to a positive expression

$$\sum_{i,j} f_{ij}g_{ij} = \sum_j \mu_j\left(u'\left[x_{lj}\right] - u'\left[x_{hj}\right]\right) > 0.$$

That is, transfers which would be worth doing without an IC constraint are restricted, raising the social marginal utilities of consumption, on average, above the value of resources in the hands of the government. Since the IC constraints are on the high skilled types, on average more redistribution from the high earners to the low earners is desirable.

IC constraints   Given the equal pay constraint, it follows that only one of the IC constraints is binding, and it is the one on the low discount factor type. To see this consider the difference in consumption utility from different incomes,

$$\Delta\left[y_h, y_l, \delta_j, R\right] \equiv w_j\left[y_h, R\right] - w_j\left[y_l, R\right].$$

This difference in consumption utility is increasing in the discount factor,

$$\frac{\partial\Delta\left[y_h, y_l, \delta_j, R\right]}{\partial\delta} = u\left[c_{hj}\right] - u\left[c_{lj}\right] > 0.$$

The difference in labor disutility does not depend on the discount factor. Thus if the IC constraint is binding on the low discount factor type, it is not binding on the high discount factor type. The low discount factor type values earnings in the first period less and is therefore more tempted to switch to the less productive job.

4.2 Second Best with Small Earnings-Dependent Taxes on Savings

As above, the sign of the welfare impact of introducing a small linear savings tax or subsidy depends on the welfare weights. Given observability of earnings, the small linear tax on savings could be different for high and low earners. The welfare impacts of introducing a tax on savings (collected in the first period) are obtained by differentiating the Lagrangian (with savings taxation included, and the tax rates, τ_i, set at zero):

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = \lambda \left(\sum_j f_{hj} (y_h - x_{hj}) \right) - \sum_j f_{hj}\eta_{hj}u'[x_{hj}](y_h - x_{hj}) - \mu_l u'[x_{hl}](y_h - x_{hl}),$$

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = \lambda \left(\sum_j f_{lj} (y_l - x_{lj}) \right) - \sum_j f_{lj}\eta_{lj}u'[x_{lj}](y_l - x_{lj}) + \mu_l u'[x_{ll}](y_l - x_{ll}).$$

That is, the impact on the Lagrangian is made up of three pieces: the impact on the revenue constraint, the impact on utilities, and the impact on the binding IC constraint.

The FOC for earnings are

$$\sum_j f_{hj}g_{hj} + \mu_l u'[x_{hl}] = 0,$$

$$\sum_j f_{lj}g_{lj} - \mu_l u'[x_{ll}] = 0.$$

Multiplying these by the earnings level at the job, y_i, and substituting, we have

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = \sum_j f_{hj}g_{hj}x_{hj} + \mu_l u'[x_{hl}]x_{hl},$$

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = \sum_j f_{lj}g_{lj}x_{lj} - \mu_l u'[x_{ll}]x_{ll}.$$

Substituting for $\mu_i u'[x_{il}]$ from the FOC for earnings, we find

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = f_{hh}g_{hh}(x_{hh} - x_{hl}) > 0$$

and

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = f_{lh}g_{lh}(x_{lh} - x_{ll}) < 0.$$

The signs follow from the assumption on the net social marginal utilities and the differences in savings behavior by types ih and il for $i = h, l$. The correlation between skill
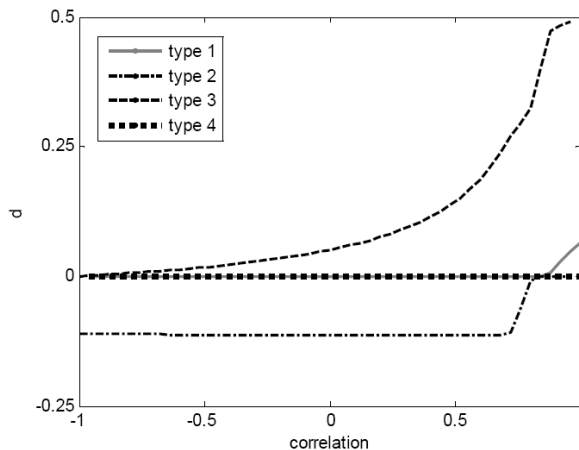
FIGURE 1: Marginal tax on savings in the welfarist case

and discount plays no role in signing these expressions. The Proposition immediately follows.

Proposition 1 *At the second best optimum, assuming that all high skill workers hold high skill jobs and*

$$g_{hj} < 0, g_{lj} > 0, \text{ for } j = h, l,$$

then introduction of a small linear tax on savings that falls on high earners is welfare improving; and introduction of a small linear subsidy on savings that falls on low earners is welfare improving.

One can increase the redistribution from high earners/high savers by taxing savings, but increasing net-of-tax earnings just enough that the high earners/low savers remain indifferent to job change and thus the binding IC constraint is unchanged. One can also increase the redistribution towards the low earners/high savers by subsidizing their savings, but decreasing net-of-tax earnings such that the low earners/low savers remain indifferent so that it doesn't become more attractive for the high earners/low savers to take the low job.

Comparison with Tenhunen and Tuomala (2008) As noted above, Tenhunen and Tuomala (2008) consider two-, three- and four-types models with hours chosen by workers. They derive the mechanism design optimal allocations assuming CES preferences with varying correlations between discount and skill, with implicit marginal taxes shown in their Figure 1. For all but the highest correlation, they find that savings are implicitly taxed for the high skill worker with low discount factor (type 3), savings are implicitly subsidized for the low skill worker with high discount factor (type 2), and there are no other marginal savings distortions. With very high correlation, the low skill worker with low discount factor (type 1) is implicitly taxed, which also happens in the

11

two type model, which has perfect correlation. The potential relevance of the pattern we find is enhanced by their findings. In contrast with our model, the mechanism design optimal allocation allows to distort the savings of each type separately. However, on average the savings tax is positive for the high skill and negative for the low skill types.

4.3   Robustness

We consider three extensions to highlight the extent of robustness of the main proposition. First, we allow different ability levels for the two high earner types. Second, we allow different discount factors for all four types. Third, we show how the analysis extends to three skill levels in workers and jobs, preserving the various assumptions.

Different Ability Levels of the High Earners   In the four-types model above, we assume that the two types with high skill have exactly the same skill. As long as the high skill type with high discount factor has higher skill than the high skill type with low discount factor, Proposition 1 continues to hold. However, if the type with low discount factor is sufficiently more skilled, the type with high discount factor may be more tempted to switch to the low earner job for which less output is required. For given skill of type hh, n_{hh}, this reversal of which IC constraint is binding holds when the ability level n_{hl} of type hl is higher than $\hat{n}_{hl}$, where the cut-off level $\hat{n}_{hl}$ is such that the IC constraint is just binding on both types.

$$\{w_l\,[y_h, R] - v\,(z_h/\hat{n}_{hl})\} - \{w_l\,[y_l, R] - v\,(z_l/\hat{n}_{hl})\} =$$
$$\{w_h\,[y_h, R] - v\,(z_h/n_{hh})\} - \{w_h\,[y_l, R] - v\,(z_l/n_{hh})\}$$

With $v\,[z/n]$ convex, the difference in labor disutilities between jobs, $\{v\,(z_l/n) - v\,(z_h/n)\}$, is decreasing in n. Hence, for values of n_{hl} higher than $\hat{n}_{hl}$ the IC constraint is more stringent for the high discount saver. In this case, a savings subsidy on the high earners and a savings tax on the low earners are welfare improving. This is the opposite of Proposition 1.

Different Discount Rates among the High and Low Savers   With job-specific earnings and no taxation of savings, a high skill worker considering switching to the low job chooses optimal savings without needing to match any particular worker holding the low job. Thus, with the same skill among high earners, the gain from switching to the low job is always higher for the high skill worker who has lower preference for savings, regardless of the discount rates among the low skill workers. We continue to have a welfare gain from introducing taxation of savings among high earners as in Proposition 1.

Subsidization of savings of low earners will continue to generate a welfare gain as long as the discount factor of the high-skill low-saver is small enough relative to the

distribution of discount factors among holders of the low skill job. Denoting by $\widetilde{x}_{hl}$ the first-period consumption of the high-skilled low saver if taking the low skill job, the FOC for earnings on that job is:

$$\sum_j f_{lj} g_{lj} - \mu_l u' [\widetilde{x}_{hl}] = 0.$$

The impact of a savings tax on low earners is

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = \sum_j f_{lj} g_{lj} x_{lj} - \mu_l u' [\widetilde{x}_{hl}] \widetilde{x}_{hl}.$$

Comparing the consumption in the IC constraint with a weighted average of consumptions among low earners, this derivative is negative (and the gain from the subsidization of the savings of low earners in Proposition 1 continues to hold) if and only if $\widetilde{x}_{hl} > \overline{x}_l$, where

$$\overline{x}_l = \frac{\sum_j f_{lj} g_{lj} x_{lj}}{\sum_j f_{lj} g_{lj}}.$$

With the net marginal social values assumption, $g_{lj} > 0$, for $j = h, l$, $\overline{x}_l$ is a proper weighted average of the x_{lj}.

Three Ability Levels, Three Jobs   We introduce an intermediate skill level in the model. We extend the assumption that welfare weights and population fractions are such that at the optimum all the high skilled are on the most productive job and all the intermediate skilled are on the intermediate job. We again consider the case in which agents may be tempted to switch to jobs designed for less skilled people. Only two downward constraints are relevant though.

First, as above, for two agents with the same skill, but different discount factors, the IC constraint is slack for the type with the higher discount factor if it is binding for the type with the lower discount factor. The reason is again that both

$$\frac{\partial \Delta [y_m, y_l, \delta_j, R]}{\partial \delta} > 0 \text{ and } \frac{\partial \Delta [y_h, y_i, \delta_j, R]}{\partial \delta} > 0 \text{ for } i = m, l,$$

with

$$\Delta [y_1, y_2, \delta_j, R] \equiv w_j [y_1, R] - w_j [y_2, R].$$

Second, with $v [z/n]$ convex, we have a similar condition for the difference in the disutility of labor between jobs. That is, with

$$\Delta' [z_h, z_l, n] \equiv v [z_h/n] - v [z_l/n],$$

$$\frac{\partial \Delta' [z_h, z_l, n]}{\partial n} = \left(-v' [z_h/n] z_h + v' [z_l/n] z_l \right) /n^2 < 0.$$

Thus, for two agents with the same discount factor, the IC constraint of switching to

the low-skilled job is slack for the type with the highest ability if it is satisfied for the type with the intermediate ability switching to the low-skilled job and for the type with highest ability switching to the intermediate job. That is, the local IC constraints imply the global IC constraint.

In a similar way as for the four-types model, we can set up the Lagrangian for the constrained maximization problem. The two relevant IC constraints are

$$
\begin{aligned}
w_l\left[y_h, R\right] - v\left[z_h/n_h\right] &\geq w_l\left[y_m, R\right] - v\left[z_m/n_h\right] \\
w_l\left[y_m, R\right] - v\left[z_m/n_m\right] &\geq w_l\left[y_l, R\right] - v\left[z_l/n_m\right].
\end{aligned}
$$

The impact of the introduction of earnings-dependent savings taxes on the Lagrangian equals respectively

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \tau_h} &= f_{hh}g_{hh}\left(x_{hh} - x_{hl}\right) > 0, \\
\frac{\partial \mathcal{L}}{\partial \tau_m} &= f_{mh}g_{mh}\left(x_{mh} - x_{ml}\right) \gtrless 0, \\
\frac{\partial \mathcal{L}}{\partial \tau_l} &= f_{lh}g_{lh}\left(x_{lh} - x_{ll}\right) < 0.
\end{aligned}
$$

This implies that Proposition 1 continues to hold for the high earners and the low earners. The following Proposition applies for the intermediate earners.

Proposition 2 *In a model with three ability levels and three jobs, the introduction of a small linear tax (subsidy) on savings that falls on the intermediate earners is welfare improving if redistribution from (to) the intermediate earners to (from) general revenues is welfare improving.*

5 Second Best with Small Uniform Taxes on Savings

Proposition 1 leaves the natural question of what to do with a Nordic dual income tax where the tax rate on savings is required to be the same for both earnings levels. Adding the responses to the two separate tax changes, we have

$$
\frac{\partial \mathcal{L}}{\partial \tau} = \frac{\partial \mathcal{L}}{\partial \tau_h} + \frac{\partial \mathcal{L}}{\partial \tau_l} = f_{hh}g_{hh}\left(x_{hh} - x_{hl}\right) + f_{lh}g_{lh}\left(x_{lh} - x_{ll}\right).
$$

In contrast with the earnings-varying tax on savings, the correlation between skill and discount factor plays a role here.

If there is no desire to redistribute within a job, $g_{ih} = g_{il}$, for $i = h, l$, then

$$f_{hh}g_{hh} = \frac{f_{hh}}{\sum_j f_{hj}} \sum_j f_{hj}g_{hj} = -\frac{f_{hh}}{\sum_j f_{hj}}\mu_l u'[x_{hl}],$$

$$f_{lh}g_{lh} = \frac{f_{lh}}{\sum_j f_{lj}} \sum_j f_{lj}g_{lj} = \frac{f_{lh}}{\sum_j f_{lj}}\mu_l u'[x_{ll}].$$

Thus

$$\frac{\partial \mathcal{L}}{\partial \tau} = \mu_l \left(\frac{f_{hh}}{\sum_j f_{hj}} u'[x_{hl}](x_{hl} - x_{hh}) - \frac{f_{lh}}{\sum_j f_{lj}} u'[x_{ll}](x_{ll} - x_{lh}) \right).$$

It is convenient to write this as

$$\frac{\partial \mathcal{L}}{\partial \tau} = \mu_l \frac{f_{lh}}{\sum_j f_{lj}} u'[x_{ll}](x_{ll} - x_{lh}) \left(\frac{f_{hh}/\sum_j f_{hj}}{f_{lh}/\sum_j f_{lj}} \Omega - 1 \right),$$

with

$$\Omega \equiv \frac{u'[x_{hl}](x_{hl} - x_{hh})}{u'[x_{ll}](x_{ll} - x_{lh})} > 0.$$

Since $x_{ll} > x_{lh}$, $sign\left(\frac{\partial \mathcal{L}}{\partial \tau}\right) = sign\left(\frac{f_{hh}/\sum_j f_{hj}}{f_{lh}/\sum_j f_{lj}} \Omega - 1\right)$. The sign of this expression depends on the distribution of types and the ratio of the weights, Ω. That is, a sufficient condition for a positive correlation between aspects, $\frac{f_{hh}}{\sum_j f_{hj}} > \frac{f_{lh}}{\sum_j f_{lj}}$, to imply a positive impact on social welfare from introducing a savings tax is that $\Omega \geq 1$.

Assuming homothetic preferences, so that $\frac{x_{hl}}{x_{ll}} = \frac{x_{hh}}{x_{lh}}$, the expression for Ω becomes $\frac{u'[x_{hl}]x_{hl}}{u'[x_{ll}]x_{ll}}$. This expression is equal to one for the log utility function. For CRRA preferences, $u[x] = \frac{x^{1-\gamma}}{1-\gamma}$, we find

$$\Omega = \left(\frac{x_{hl}}{x_{ll}} \right)^{1-\gamma}.$$

Thus, if the relative risk aversion γ is smaller than or equal to 1, when the correlation between ability and discount factor is positive (i.e. $\frac{f_{hh}}{\sum_j f_{hj}} > \frac{f_{lh}}{\sum_j f_{lj}}$), $\Omega \geq 1$, and thus $\frac{\partial \mathcal{L}}{\partial \tau}$ is positive. If γ is larger than 1, the sign of $\frac{\partial \mathcal{L}}{\partial \tau}$ depends on the size of the correlation and the magnitude of the earnings difference between jobs. Conversely, when the correlation is negative, $\frac{\partial \mathcal{L}}{\partial \tau}$ is negative if γ is greater than or equal to 1.[5] This implies the following proposition.

Proposition 3 *With CRRA preferences, a uniform small tax on savings increases welfare if the relative risk aversion is smaller than one and the correlation between ability and discount factor is positive. A uniform small subsidy on savings increases welfare if the relative risk aversion is greater than one and the correlation between*

---

[5] For CARA preferences, $\frac{\partial \mathcal{L}}{\partial \tau}$ is negative when the correlation between ability and discount factor is negative. When the correlation is positive, $\frac{\partial \mathcal{L}}{\partial \tau}$ is positive if the absolute risk aversion is sufficiently small.

ability and discount factor is negative.

6 Continuous Ability or Discount

We explore the role of the nonbinding IC constraint by considering separately a continuum of skill levels, and a continuum of discount factors.

6.1 A Continuum of Ability Levels

As above, the discount factor can take two values, either δ_h or δ_l with $\delta_h > \delta_l$. However, ability n is continuously distributed in the population between $\underline{n}$ and $\overline{n}$, with densities $f_j(n)$ for $j = l, h$ and total numbers N_j. The agent with discount factor δ_j chooses job h if her ability is higher than or equal to the cutoff level $\hat{n}_j$ and job l otherwise. The cutoff levels satisfy the conditions for indifference between the jobs:

$$
\begin{aligned}
w_h[y_h, R] - v[z_h/\hat{n}_h] &= w_h[y_l, R] - v[z_l/\hat{n}_h] \\
w_l[y_h, R] - v[z_h/\hat{n}_l] &= w_l[y_l, R] - v[z_l/\hat{n}_l].
\end{aligned}
$$

Notice that $\hat{n}_h < \hat{n}_l$, since the types with low discount factor have a lower marginal valuation of earnings.

The index ij refers to the individual who chooses job $i \in \{h, l\}$, and has discount factor $\delta_j \in \{\delta_h, \delta_l\}$. We assume that the Pareto weights depend only on the discount factor.[6] Setting up the problem as before, but with shifting fractions of workers on the two jobs,

$$
\text{Maximize}_{z,y,n} \sum_j N_j \int_{\hat{n}_j}^{\overline{n}} \eta_j \{w_j[y_h, R] - v[z_h/n]\} \, dF_j(n)
$$

$$
+ \sum_j N_j \int_{\underline{n}}^{\hat{n}_j} \eta_j \{w_j[y_l, R] - v[z_l/n]\} \, dF_j(n)
$$

subject to

$$
E + \sum_j N_j \int_{\hat{n}_j}^{\overline{n}} \{y_h - z_h\} \, dF_j(n) + \sum_j N_j \int_{\underline{n}}^{\hat{n}_j} \{y_l - z_l\} \, dF_j(n) \leq 0
$$

$$
w_h[y_h, R] - v[z_h/\hat{n}_h] = w_h[y_l, R] - v[z_l/\hat{n}_h]
$$

$$
w_l[y_h, R] - v[z_h/\hat{n}_l] = w_l[y_l, R] - v[z_l/\hat{n}_l]
$$

---

[6] Thus a change in jobs that does not affect lifetime utility does not directly affect the social welfare objective function.

We find the FOC with respect to earnings,

$$\sum_j N_j \left(1 - F_j \left(\hat{n}_j\right)\right) g_{hj} + \mu_l u' \left[x_{hl}\right] + \mu_h u' \left[x_{hh}\right] = 0$$

$$\sum_j N_j F_j \left(\hat{n}_j\right) g_{lj} - \mu_l u' \left[x_{ll}\right] - \mu_h u' \left[x_{lh}\right] = 0.$$

The welfare impacts of introducing earnings-dependent taxes on savings are again obtained by differentiating the Lagrangian

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = \sum_j N_j \left(1 - F_j \left(\hat{n}_j\right)\right) g_{hj} x_{hj} + \mu_l u' \left[x_{hl}\right] x_{hl} + \mu_h u' \left[x_{hh}\right] x_{hh}$$

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = \sum_j N_j F_j \left(\hat{n}_j\right) g_{lj} x_{lj} - \mu_l u' \left[x_{ll}\right] x_{ll} - \mu_h u' \left[x_{lh}\right] x_{lh}.$$

Using the FOC for earnings as above, with $g_{hj} < 0, g_{lj} > 0, \mu_h > 0$, we find

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = \underbrace{N_h \left(1 - F_h \left(\hat{n}_h\right)\right) g_{hh} \left(x_{hh} - x_{hl}\right)}_{>0} + \underbrace{\mu_h u' \left[x_{hh}\right] \left(x_{hh} - x_{hl}\right)}_{<0} \tag{4}$$

$$= \left\{N_h \left(1 - F_h \left(\hat{n}_h\right)\right) g_{hh} + \mu_h u' \left[x_{hh}\right]\right\} \left(x_{hh} - x_{hl}\right) \tag{5}$$

and

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = \underbrace{N_h F_h \left(\hat{n}_h\right) g_{lh} \left(x_{lh} - x_{ll}\right)}_{<0} + \underbrace{\left(-\mu_h\right) u' \left[x_{lh}\right] \left(x_{lh} - x_{ll}\right)}_{>0} \tag{6}$$

$$= \left\{N_h F_h \left(\hat{n}_h\right) g_{lh} - \mu_h u' \left[x_{lh}\right]\right\} \left(x_{lh} - x_{ll}\right). \tag{7}$$

This can be interpreted as the effect of a savings tax and a change in the earnings that keep the types with low discount factor and therefore $\hat{n}_l$ unaffected. First, a small tax redistributes away from the types with high discount factor. This is the first term in (4) and (6). Second, a tax changes the cut-off ability $\hat{n}_h$. The tax change for the high earners involving a savings tax discourages some types with high discount factor from staying on the high job, given that it leaves the types with low discount factor indifferent. Similarly, the tax change for the low earners involving a savings tax encourages some types with high discount factor to switch to the high job. These are the second terms in respectively (4) and (6). Hence, with $g_{hh} < 0$ and $g_{lh} > 0$, the two terms have opposite signs and the signs of both $\frac{\partial \mathcal{L}}{\partial \tau_h}$ and $\frac{\partial \mathcal{L}}{\partial \tau_l}$ are ambiguous in general. We examine the signs under further restrictions.

With no concern for redistribution within earnings groups, $g_{hh} = g_{hl} < 0, g_{lh} =$

$g_{ll} > 0$, (4) and (6) become

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = (x_{hl} - x_{hh}) u'[x_{hh}] \mu_l \frac{N_l(1-F_l(\hat{n}_l))}{\sum_j N_j(1-F_j(\hat{n}_j))} \left\{ \frac{u'[x_{hl}]}{u'[x_{hh}]} \frac{N_h(1-F_h(\hat{n}_h))}{N_l(1-F_l(\hat{n}_l))} - \frac{\mu_h}{\mu_l} \right\}$$

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = -(x_{ll} - x_{lh}) u'[x_{lh}] \mu_l \frac{N_l F_l(\hat{n}_l)}{\sum_j N_j F_j(\hat{n}_j)} \left\{ \frac{u'[x_{ll}]}{u'[x_{lh}]} \frac{N_h F_h(\hat{n}_h)}{N_l F_l(\hat{n}_l)} - \frac{\mu_h}{\mu_l} \right\}.$$

Earnings-dependent taxes on savings affect the job choices of both the high savers and the low savers. The relative importance of these impacts, μ_h/μ_l plays a key role in determining the signs of the tax changes that increase welfare. We distinguish three effects in the determination of the signs of the bracketed expressions.

First, the expressions in brackets are increasing in $\frac{N_h(1-F_h(\hat{n}_h))}{N_l(1-F_l(\hat{n}_l))}$ and $\frac{N_h F_h(\hat{n}_h)}{N_l F_l(\hat{n}_l)}$ respectively. These terms reflect the relative distributions of high savers and low savers on the respective jobs. Assuming that high savers are more likely to hold the high job, then $\frac{1-F_h(\hat{n}_h)}{1-F_l(\hat{n}_l)} > \frac{F_h(\hat{n}_h)}{F_l(\hat{n}_l)}$. Since $\hat{n}_h < \hat{n}_l$, this holds if F_l and F_h follow the same distribution. Positive correlation between ability and discount factor would only strengthen the inequality.

Second, the expressions in brackets are increasing in $\frac{u'[x_{hl}]}{u'[x_{hh}]}$ and $\frac{u'[x_{ll}]}{u'[x_{lh}]}$ respectively. These terms depend on the differences in savings behavior among workers on each job. With $u = \log$,

$$x_{ij} = \frac{y_i}{1 + \delta_j}.$$

Thus

$$\frac{u'[x_{hl}]}{u'[x_{hh}]} = \frac{x_{hh}}{x_{hl}} = \frac{1 + \delta_l}{1 + \delta_h} = \frac{x_{lh}}{x_{ll}} = \frac{u'[x_{ll}]}{u'[x_{lh}]},$$

implying

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = \frac{x_{hl} - x_{hh}}{x_{hh}} \mu_l \frac{N_l(1-F_l(\hat{n}_l))}{\sum_j N_j(1-F_j(\hat{n}_j))} \left\{ \frac{x_{hh}}{x_{hl}} \frac{N_h}{N_l} \frac{(1-F_h(\hat{n}_h))}{(1-F_l(\hat{n}_l))} - \frac{\mu_h}{\mu_l} \right\}$$

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = -\frac{x_{hl} - x_{hh}}{x_{hh}} \mu_l \frac{N_l F_l(\hat{n}_l)}{\sum_j N_j F_j(\hat{n}_j)} \left\{ \frac{x_{hh}}{x_{hl}} \frac{N_h}{N_l} \frac{F_h(\hat{n}_h)}{F_l(\hat{n}_l)} - \frac{\mu_h}{\mu_l} \right\}.$$

Thus, with $\frac{1-F_h(\hat{n}_h)}{1-F_l(\hat{n}_l)} > \frac{F_h(\hat{n}_h)}{F_l(\hat{n}_l)}$, we can rule out one of the four possible sign patterns, that subsidization of savings is welfare improving at both earnings levels:

$$\left\{ \frac{\partial \mathcal{L}}{\partial \tau_l} < 0 \right\} \Rightarrow \left\{ \frac{\partial \mathcal{L}}{\partial \tau_h} > 0 \right\}$$

$$\left\{ \frac{\partial \mathcal{L}}{\partial \tau_h} < 0 \right\} \Rightarrow \left\{ \frac{\partial \mathcal{L}}{\partial \tau_l} > 0 \right\}.$$

Third, the expressions in brackets are decreasing in $\frac{\mu_h}{\mu_l}$. This gives the relative welfare losses from encouraging workers to switch to the low earnings job. This ratio can be calculated from the FOC for the cutoff levels of skill, $\hat{n}_h, \hat{n}_l$. With the additional

assumption of a power function disutility of labor: $v\left[z/n\right] = k\left(z/n\right)^{\beta+1}/\left(\beta+1\right),\ \beta > 0$,

$$\frac{\mu_h}{\mu_l} = \frac{N_h f_h\left(\hat{n}_h\right)}{N_l f_l\left(\hat{n}_l\right)}\left(\frac{\hat{n}_h}{\hat{n}_l}\right)^{\beta+2}.$$

From the equations determining the cutoff levels of skill, we have

$$\left(\frac{\hat{n}_h}{\hat{n}_l}\right)^{\beta+1} = \frac{1+\delta_l}{1+\delta_h}.$$

Thus

$$\frac{\mu_h}{\mu_l} = \frac{N_h f_h\left(\hat{n}_h\right)}{N_l f_l\left(\hat{n}_l\right)}\left(\frac{1+\delta_l}{1+\delta_h}\right)^{\frac{\beta+2}{\beta+1}}.$$

Under these assumptions on preferences, we can express the impact of savings taxes as

$$\frac{\partial\mathcal{L}}{\partial\tau_h} = \frac{x_{hl}-x_{hh}}{x_{hh}}\mu_l\frac{N_l(1\text{-}F_l(\hat{n}_l))}{\sum_j N_j(1-F_j(\hat{n}_j))}\left\{\frac{\left(1-F_h\left(\hat{n}_h\right)\right)}{\left(1-F_l\left(\hat{n}_l\right)\right)}-\frac{f_h\left(\hat{n}_h\right)}{f_l\left(\hat{n}_l\right)}\left(\frac{1+\delta_l}{1+\delta_h}\right)^{\frac{1}{\beta+1}}\right\}\left(\frac{1+\delta_l}{1+\delta_h}\right)\left(\frac{N_h}{N_l}\right)$$

$$\frac{\partial\mathcal{L}}{\partial\tau_l} = -\frac{x_{hl}-x_{hh}}{x_{hh}}\mu_l\frac{N_l F_l(\hat{n}_l)}{\sum_j N_j F_j(\hat{n}_j)}\left\{\frac{F_h\left(\hat{n}_h\right)}{F_l\left(\hat{n}_l\right)}-\frac{f_h\left(\hat{n}_h\right)}{f_l\left(\hat{n}_l\right)}\left(\frac{1+\delta_l}{1+\delta_h}\right)^{\frac{1}{\beta+1}}\right\}\left(\frac{1+\delta_l}{1+\delta_h}\right)\left(\frac{N_h}{N_l}\right).$$

With $\frac{1\text{-}F_h(\hat{n}_h)}{1\text{-}F_l(\hat{n}_l)} > \frac{F_h(\hat{n}_h)}{F_l(\hat{n}_l)}$ and relatively equal densities of skills in the range between $\hat{n}_h$ and $\hat{n}_l$, we have applicability of the following proposition.

Proposition 4 *With a continuum of skills and two discount factors, a small tax on the savings of the high earners increases welfare if* $\frac{f_h(\hat{n}_h)}{1-F_h(\hat{n}_h)} < \frac{f_l(\hat{n}_l)}{1-F_l(\hat{n}_l)}$, $v\left[z/n\right] = k\left(z/n\right)^{\beta+1}/\left(\beta+1\right),\ \beta > 0,$ *and* $u = \log$.

The similar condition on low earners, $\frac{f_h(\hat{n}_h)}{F_h(\hat{n}_h)} < \frac{f_l(\hat{n}_l)}{F_l(\hat{n}_l)}$, which is sufficient to imply that a savings subsidy for low earners increases welfare, is not applicable if the densities are about the same.

Small Uniform Taxes on Savings   With no concern for redistribution within earnings groups, $g_{hh} = g_{hl} < 0$, $g_{lh} = g_{ll} > 0$, and $u = \log$, the condition for uniform savings taxation is

$$\frac{\partial\mathcal{L}}{\partial\tau} = \frac{x_{hl}-x_{hh}}{x_{hh}}\mu_l\left\{\left(\frac{x_{hh}}{x_{hl}}+\frac{\mu_h}{\mu_l}\right)\left(\frac{N_h\left(1\text{-}F_h\left(\hat{n}_h\right)\right)}{\sum_j N_j\left(1-F_j\left(\hat{n}_j\right)\right)}-\frac{N_h F_h\left(\hat{n}_h\right)}{\sum_j N_j F_j\left(\hat{n}_j\right)}\right)\right\} > 0,$$

with the sign following from the assumption that $\frac{1\text{-}F_h(\hat{n}_h)}{1\text{-}F_l(\hat{n}_l)} > \frac{F_h(\hat{n}_h)}{F_l(\hat{n}_l)}$.

Proposition 5 *With a continuum of skills and two discount factors, a uniform small tax on savings increases welfare if* $g_{hh} = g_{hl} < 0$, $g_{lh} = g_{ll} > 0$, $\frac{1\text{-}F_h(\hat{n}_h)}{1\text{-}F_l(\hat{n}_l)} > \frac{F_h(\hat{n}_h)}{F_l(\hat{n}_l)}$, *and* $u = \log$.

6.2 A Continuum of Discount Factors

We assume again that skill can only take two values, n_h and n_l with $n_h > n_l$. However, the discount factor δ is assumed to be continuously distributed in the population between $\underline{\delta}$ and $\overline{\delta}$, with densities $f_i(\delta)$ for $i = l, h$ and total numbers N_i. We still assume two jobs. We continue to assume that the high job can only be held by a high skilled worker. The high skilled agent chooses job h if her discount factor is higher than or equal to the cutoff level $\hat{\delta}$ and job l otherwise. The cutoff level satisfies the condition for indifference between the jobs:

$$w_{\hat{\delta}}[y_h, R] - v[z_h/n_h] = w_{\hat{\delta}}[y_l, R] - v[z_l/n_h].$$

The index $i\delta$ refers to the individual who chooses job $i \in \{h, l\}$, and has discount factor δ.

The FOC with respect to earnings are

$$N_h \int_{\hat{\delta}}^{\overline{\delta}} g_{h\delta} dF_h(\delta) + \mu_h u'[x_{h\hat{\delta}}] = 0$$

$$N_h \int_{\underline{\delta}}^{\hat{\delta}} g_{l\delta} dF_h(\delta) + N_l \int_{\underline{\delta}}^{\overline{\delta}} g_{l\delta} dF_l(\delta) - \mu_h u'[x_{l\hat{\delta}}] = 0.$$

The welfare impacts of introducing a tax on savings are

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = N_h \int_{\hat{\delta}}^{\overline{\delta}} g_{h\delta} x_{h\delta} dF_h(\delta) + \mu_h u'[x_{h\hat{\delta}}] x_{h\hat{\delta}}$$

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = N_h \int_{\underline{\delta}}^{\hat{\delta}} g_{l\delta} x_{l\delta} dF_h(\delta) + N_l \int_{\underline{\delta}}^{\overline{\delta}} g_{l\delta} x_{l\delta} dF_l(\delta) - \mu_h u'[x_{l\hat{\delta}}] x_{l\hat{\delta}}.$$

Using the FOC with respect to earnings, with $g_{h\delta} < 0$, $g_{l\delta} > 0$, and $\mu_h > 0$, we find

$$\frac{\partial \mathcal{L}}{\partial \tau_h} = N_h \int_{\hat{\delta}}^{\overline{\delta}} g_{h\delta} \left(x_{h\delta} - x_{h\hat{\delta}} \right) dF_h(\delta) > 0 \tag{8}$$

$$\frac{\partial \mathcal{L}}{\partial \tau_l} = N_h \int_{\underline{\delta}}^{\hat{\delta}} g_{l\delta} \left(x_{l\delta} - x_{l\hat{\delta}} \right) dF_h(\delta) + N_l \int_{\underline{\delta}}^{\overline{\delta}} g_{l\delta} \left(x_{l\delta} - x_{l\hat{\delta}} \right) dF_l(\delta). \tag{9}$$

The Proposition immediately follows.

Proposition 6 *With a continuum of discount factors and two skill types, a small savings tax on the high earners increases welfare.*

A savings tax on the high earners unambiguously increases welfare as the marginal saver has lower savings than the inframarginal savers holding the high job. The marginal saver does have higher savings than the inframarginal high skilled workers holding the low job. This is described by the first term in (9) and decreases the welfare gain

of a savings subsidy on the low earners. Whether such savings subsidy is still welfare improving as in Proposition 1, depends on whether the savings of the marginal high ability worker choosing the low job are sufficiently low compared to the savings of the low ability workers, who also hold the low job. This is described by the second term in **(9).** If agents with both skills were capable of holding both jobs, the analysis would be more complex.

7   Optimal Linear Earnings-Varying Savings Tax

We have considered the introduction of small savings taxes on high and low earners. Part of the interest in this analysis comes from the possible link to the signs of the optimal taxes. Derivation of the FOC for the optimal linear savings taxes is straightforward; we show that it matches the signs of the small improvements for a special case.[7]

The only difference in analysis is that changes in both the earnings and savings taxes have a first order effect on tax revenues through the behavioral change in savings. Denote by $R_h \equiv R(1 - \tau_h)$ and $R_l \equiv R(1 - \tau_l)$ the after-tax returns to savings for respectively the high type and the low type. Denote by $s_{ij}[y, R]$ the savings function of type ij as a function of after-tax earnings and the after-tax interest rate. (Given preference separability, there is no dependence on gross earnings.) Following the same approach as before, the derivative of the Lagrangian is the expression above plus the impact on revenue from the taxation of savings. Setting the derivative of the Lagrangian to zero, the optimal linear savings tax is such that

$$f_{hh}g_{hh}\left(x_{hh} - x_{hl}\right) = -\frac{R_h - R}{R}\lambda \sum_j f_{hj}\left\{\frac{\partial s_{hj}}{\partial y_h}x_{hl} + s_{hj} + \frac{\partial s_{hj}}{\partial R_h}R_h - \frac{\partial s_{hj}}{\partial y_h}y_h\right\}$$

and

$$f_{lh}g_{lh}\left(x_{lh} - x_{ll}\right) = -\frac{R_l - R}{R}\lambda \sum_j f_{lj}\left\{\frac{\partial s_{lj}}{\partial y_l}x_{ll} + s_{lj} + \frac{\partial s_{lj}}{\partial R_l}R_l - \frac{\partial s_{lj}}{\partial y_l}y_l\right\}.$$

Thus, if the sum of the terms in brackets are positive, we have the same signs for the optimal taxes as for the welfare-improving introductions of taxes, which are the same as that of the left-hand side.

In the case of logarithmic preferences $u[x] = \log[x]$, savings are linearly increasing

---

[7]Consideration of earnings-dependent nonlinear savings taxation would raise the issue of the degree of complexity that is interesting for policy purposes.

in income and independent of the interest rate, thus

$$s_{ij} = \frac{\partial s_{ij}}{\partial y_i} y_i \text{ and}$$

$$\frac{\partial s_{ij}}{\partial R_j} R_j = 0 \text{ and}$$

$$\frac{\partial s_{ij}}{\partial y_i} x_{ij} > 0 \text{ for all } i,j.$$

This implies that under the assumption that $g_{hj} < 0$ and $g_{lj} > 0$ for $j = h, l$, the optimal linear savings tax is positive for the high earners and negative for the low earners. This result holds more generally for CRRA preferences $u[x] = \frac{x^{1-\gamma}}{1-\gamma}$ with $\gamma < 1$ so that $\frac{\partial s_{ij}}{\partial R_j} R_j > 0$. When $\gamma > 1$, $\frac{\partial s_{ij}}{\partial R_j} R_j < 0$ and the relative population weights at each skill level matter.

8 Two Types

Using a model with two types of workers, a high-skilled worker with high discount factor and a low-skilled worker with low discount factor, the empirical finding of a positive correlation between skill and savings rates is treated as a perfect correlation. In this model, if there is positive (negative) marginal earnings taxation then there is a gain from introducing positive (negative) marginal savings taxation. The corollary is that introducing savings taxation is a gain if redistribution goes from the high earner to the low earner. The full mechanism design optimum has the same property. The source of this inference does not seem robust to realistic diversity in the economy. With two-dimensional heterogeneity, there are low earners with both high and low savings rates. If a high earner can imitate the savings of a low earner with the same savings propensities, a savings tax on the low earner does not work to discourage the high earner from imitating. Thus to model less-than-perfect correlation, we use the four-types model with high and low earners with both high and low concern for the future. Nevertheless, we report the results here, to mark the contrast with the four types model. The proof parallels that of the same result for the mechanism design optimum, which is in Diamond (2003). We consider the second-best Pareto frontier with the types referred to as 1 and 2.

Proposition 7 *In a two-types model without taxes on savings and with* $sign\,(\delta_1 - \delta_2) = sign\,(n_1 - n_2)$, *the introduction of a small linear tax (subsidy) on savings at a given earnings level is welfare improving if and only if earnings at that level are marginally taxed (subsidized).*

Corollary 1 *In a two-types model without taxes on savings and with* $sign\,(\delta_1 - \delta_2) = sign\,(n_1 - n_2)$, *the introduction of a savings tax on the lower earner is welfare improving if redistribution goes from the higher earner to the lower earner.*

The Proposition combines the properties of the mechanism design optima in the two separate two-types models with heterogeneity in one dimension. When both types have the same discount factor, but different abilities, the earnings of the potentially imitated type are marginally taxed or subsidized if the ability of that type is lower or higher respectively (Mirrlees, 1971). Similarly, when both types have the same ability, but different discount factors, the savings of the potentially imitated type are marginally taxed or subsidized if the discount factor for that type is lower or higher respectively. If both types have the same discount factor, distorting savings does not help separate the two types (Atkinson and Stiglitz, 1976). Similarly, if both types have the same ability, earnings are not subject to marginal taxation.[8] However, when the two types differ in both ability and discount factor, both the marginal taxation (or subsidization) of earnings and the marginal taxation (or subsidization) of savings is used to separate types.

9   Preferences and IC Constraints

Above we used the utility functions $u[x] + \delta_j u[c] - v[z/n_j]$. This family of utility functions has the property that those with higher savings rates (larger values of δ_j) are more willing to increase work for a given amount of additional pay. But that is not the only way in which the savings and labor supply decisions can be connected. For example, with the utility functions $(u[x] + \delta_j u[c])/\delta_j - v[z/n_j] = u[x]/\delta_j + u[c] - v[z/n_j]$, the relationship is reversed - those with higher savings rates are less willing to increase work for additional pay. If we had assumed this class of functions, then we would have reversed the pattern of desirable savings taxes in Proposition 1 - having the IC constraint bind for the high saver would imply that it is not binding for the low saver, implying, in turn, that there should be a subsidy of savings for high earners and a tax on savings for low earners. More generally, a one-dimensional family of separable utility functions, $U[\phi[x,c,j],z,j]$, can have any pattern between the variation in the subutility function of consumption and the variation in the interaction between consumption and labor. This raises the question of identifying an empirical basis for distinguishing which case is more relevant. That it is standard practice to write utility in the form employed does not, by itself, shed light on its empirical reality.

While the formal model has consumption and work simultaneous in the first period, experience in real time is different. Generally, work precedes pay, which precedes spending it (but not borrowing against it). So modeling in continuous time would naturally have a similar role for discounting on both aspects - saving and willingness

---

[8]This can be considered an implication of the Atkinson-Stiglitz theorem, since there is separability and everyone has the same subutility function over first period earnings and consumption. Notice that if a savings tax is not allowed, the two types can be usefully separated by an earnings tax. The marginal tax on the earnings of the potentially imitated type is positive if and only if that type saves less for the same earnings.

to work. But that does not rule out the possibility that the preferences of high and low savers differ in other ways than just a discount rate on otherwise identical utility and disutility functions (assuming additivity). There are reasons based on casual empiricism for supporting the appropriateness of using the formulation employed above. Modeling savings with rationality and discounting combines underlying preferences and issues of self-control. As discussed in Banks and Diamond (2008), psychological analyses suggest these are mixed together. We see no reason to think that this does not apply to working as well as to consuming - whether that is working for later consumption or working to influence future work opportunities. That is, working (at a job with disutility) involves self-control for a future payoff. And saving involves self-control. So those with less difficulty in self-control may show greater willingness to both work and save, which would be captured in the standard utility function expression. In a richer model, human capital investment involves discounting in a similar way to savings decisions and so may generate the pattern in the standard model structure, although formal modeling would distinguish between human and financial capital accumulations.

It is not easy to find data applying directly to this issue. The question we want to answer is whether, for a given level of skill, those with higher savings rates tend to have greater labor supply functions. A complication in looking at data comes from the differences in circumstances with age, which we address by considering separate age cells. We report a few correlations supportive of a positive correlation among savings propensities, discounting and earnings abilities using the Survey of Consumer Finances, which includes some questions on time horizon and savings practice.[9] We also report some correlations with work effort. Before turning to the data, we briefly consider a three-period version of the two-period model we have been considering. This will bring out some of the complications in interpreting the data at different ages. There is also a complication in interpreting the data across education levels. Education choices reflect both ex ante "skill" and discount rate and then affect wage rates, which matter for later taxation. Presumably, the level of completed education is increasing in both ex ante skill and discount factor, on average. In addition to ex post skill, education may affect one's discount rate thereafter. Thus education is a proxy for both skill and discount rate and can not be used in a simple way to distinguish between them. A further difficulty in interpreting the correlations is that education is a discrete variable while skill is continuous and varying within education classes.

9.1 Three-period Model

We set up a three-period model with the same preference structure as the two-period model analyzed, assuming the same discounting for the utility of consumption and the

---

[9]For discussion of correlations with experimentally measured discount rates, see Chabris et al, 2008.

disutility of work and allowing for different skills in the two working periods.

$$u[x_1] + \delta u[x_2] + \delta^2 u[c] - v[z_1/n_1] - \delta v[z_2/n_2]$$

Considering the special case analyzed above, with $u[x] = \log[x]$, and $v[z/n] = k(z/n)^{\beta+1}/(\beta+1)$, we have the time series of consumption and earnings behavior, assuming that hours are a control variable and the marginal tax rate is constant over time (with derivation in the Appendix):

$$\frac{x_2}{x_1} = \frac{c}{x_2} = \delta R$$
$$\frac{z_2}{z_1} = (\delta R)^{-1/\beta}\left(\frac{n_2}{n_1}\right)^{1+1/\beta}$$
$$\frac{z_2/n_2}{z_1/n_1} = (\delta R)^{-1/\beta}\left(\frac{n_2}{n_1}\right)^{1/\beta}.$$

Thus, those with higher discount factors have more rapidly growing consumption but less rapidly growing earnings (for given skills). This suggests that the cross section pattern of earnings and work effort may be different at different ages.

The cross-section pattern of time series behavior may be more illuminating than that of single-period behavior since the single-period cross section patterns are dependent on the full pattern over time in skills. If we added uncertain rates of return to the model, we would also be concerned about income effects in both consumption and earnings choices. Consideration of wealth or wealth/earnings ratios are also affected by the time series pattern of skills. But we do not explore these issues, just reporting simple correlations

9.2 Data Analysis

We first consider the relations among discounting, saving, education and age. We use the SCF panels in 1998, 2001 and 2004, containing information on 13,266 households in total. For savings rates we consider two proxies. The first proxy is the logarithm of the ratio of net worth to earnings for households. The second proxy is whether people report that they save regularly or not.[10] The sample is divided into age-education cells (5-year age groupings from 30 to 59, 5 education groups). Tables 1 and 2 show the average savings preferences measured by the two proxies respectively.[11]

---

[10] Subject can choose among different statements. We use whether subjects confirm the statement: "Save regularly by putting money aside each month." The results are similar (with sign reversal) with the statement "Don't save - usually spend about as much as income."

[11] Population weights are used to convert the SCF sample to a representative national sample.

TABLE 1: Average of log(net worth/earnings)

	30-34	35-39	40-44	45-49	50-54	55-59
< High School	−.55	−.85	−.13	−.10	.38	.84
High School	−.70	−.11	.25	.47	.78	1.00
Some College	−.26	.15	.57	.74	1.09	1.47
College Degree	.24	.80	1.10	1.49	1.78	2.05
Graduate Degree	.44	.82	1.41	1.59	1.79	2.10

TABLE 2: Proportion of regular savers

	30-34	35-39	40-44	45-49	50-54	55-59
< High School	.21	.16	.26	.25	.28	.27
High School	.34	.40	.37	.39	.41	.48
Some College	.37	.45	.50	.42	.45	.51
College Degree	.58	.54	.51	.57	.57	.50
Graduate Degree	.64	.56	.62	.59	.54	.52

Both tables show mostly rising savings with education, and so wage rate, given age, although we have not tested for statistical significance.

Past accumulations, including high realized rates of return and inheritances, may play a confounding role in savings behavior. We therefore also consider the time horizon people have in mind when making saving decisions, as reported in the SCF.[12] This question allows analysis of the positive correlation between education and discount factor more directly. Table 3 shows the average time horizon per cell.

TABLE 3: Average time horizon (converted into years)

	30-34	35-39	40-44	45-49	50-54	55-59
< High School	2.5	2.4	3.3	3.3	3.3	3.9
High School	3.2	3.6	4.0	4.1	3.8	4.0
Some College	3.5	4.2	4.5	4.1	4.3	4.5
College Degree	4.7	5.0	5.1	5.0	5.5	5.2
Graduate Degree	5.2	6.0	5.9	5.8	5.8	5.7

Time horizon increases with education and so wage rate, which is supportive of the

---

[12] The question asks: "In planning (your/your family's) saving and spending, which of the following is most important to [you/you and your (husband/wife/partner)]: the next few months, the next year, the next few years, the next 5 to 10 years, or longer than 10 years?" To interpret the averages, we quantify this variable assigning $0.5, 1, 3, 5$ and 10 to the respective answers.

assumed positive correlation between discount factor and skill used in analysis of the taxation of savings that is not earnings-varying.

It is harder to measure whether for a given level of skill, those with higher savings rates tend to have greater labor supply functions. Moreover, although this is a robust prediction of the two-period model, the three-period example above shows that types with higher discount factors will work relatively more while they are young but may not when older. For greater labor supply curves, we look at both earnings and hours worked per week.[13] We consider again whether one reports to save regularly and the reported time horizon.[14] For each education-age cell, we calculate the correlation between the proxy for labor supply and the proxy for the preference for savings.

Hours worked   The results for the correlation between hours worked and the two proxies for the savings preferences are in Table 4 and Table 5 respectively.

TABLE 4: Correlation hours worked and saving regularly

	30-34	35-39	40-44	45-49	50-54	55-59
< High School	.05	$-$.00	.17*	.02	.24*	.04
High School	.07*	.15*	.05	$-$.01	.10*	.02
Some College	$-$.03	.04	.18*	.13*	.06*	$-$.18*
College Degree	.11*	.10*	.05*	$-$.03	.10*	.14*
Graduate Degree	.08	$-$.04	.07*	.04*	.01	.10*

* denotes statistical significance at the 5 percent level

---

[13] The question asks: How many hours (do you/does [he/she]) work on (your/her/his) main job in a normal week? (if not self-employed) How many hours (do you/does [he/she]) work in this business in a normal week? (self-employed). The result for hours a year are very similar. This leaves out the role of second jobs.

[14] Using the logarithm of the rate of net worth to earnings as a savings proxy, the correlation between hours worked and log(net worth/earnings) shows a lot of variation among the different age-education cells. For the older sample in particular, the correlation is often negative. This is consistent with the predictions of the three-period example. Also, working more hours mechanically increases earnings and therefore biases the estimated correlations between hours worked and savings preferences downward. The number of observations per cell is small, varying between 25 and 135. The correlation for the full sample between the hours worked and the rate of net worth to earnings, conditional on cell dummies, equals .03. The educations cells do not completely capture the variation in ability. To the extent that those who are more able tend to work more and save more, this estimate is biased upward. The same is true for the age cells if, within a cell, those who are older tend to work less and save less.

TABLE 5: Correlation hours worked and time horizon

	30-34	35-39	40-44	45-49	50-54	55-59
< High School	.22*	.26*	−.06	−.15*	−.01	.05
High School	−.03	.04	.07*	.02	.18*	.11*
Some College	−.01	.08*	.13*	.15*	−.00	.15*
College Degree	−.01	−.02	.03	.12*	.06*	.04
Graduate Degree	.15*	.14*	.12*	−.01	.05*	−.12*

* denotes statistical significance at the 5 percent level

Most of the correlations in Table 4 and Table 5 are positive. Many of them are significant as well. The correlations for the full-sample using the two different proxies for saving preference, conditional on cell dummies, are both .06. Chabris et al. (2008) find that experimentally measured discount rates have strong predictive power for field behavior relative to other variables in their sample (e.g., sex, age, education). However, they find that the correlation between the discount rate and each field behavior is small; none exceeds 0.28 and many are near 0.

Earnings   The correlations between earnings and the two proxies of preference for savings are in Table 6 and Table 7.[15]

TABLE 6: Correlation log(earnings) and saving regularly

	30-34	35-39	40-44	45-49	50-54	55-59
< High School	.30*	.17*	.22*	.20*	.31*	.40*
High School	.24*	.17*	.29*	.22*	.32*	.34*
Some College	.18*	.21*	.20*	.23*	.20*	.17*
College Degree	.22*	.34*	.20*	.28*	.19*	.10*
Graduate Degree	.30*	−.01	.17*	.15*	.15*	.13*

* denotes statistical significance at the 5 percent level

TABLE 7: Correlation log(earnings) and time horizon

	30-34	35-39	40-44	45-49	50-54	55-59
< High School	−.01	.12*	.12*	.19*	.22*	.28*
High School	.14*	.14*	.20*	.09*	.14*	.25*
Some College	.13*	.11*	.15*	.23*	.16*	.12*
College Degree	.17*	.36*	.19*	.29*	.28*	.22*
Graduate Degree	.33*	.18*	.23*	.28*	.24*	.23*

* denotes statistical significance at the 5 percent level.

---

[15] The correlations between earnings and log(net worth/earnings) per cell are mostly positive. The conditional correlation for the full-sample is 0.06.

Almost all of the correlations per cell are positive and significant. The correlation for the full sample between hours worked and saving regularly, conditional on cell dummies, equals .22. The conditional correlation between hours worked and the time horizon equals .19. Thus, those with longer time horizons and greater likelihood to save regularly work longer hours and earn more. Presumably this reflects a greater willingness to work and so is supportive of the basic preference structure we analyze. There are many factors that affect savings and work that are not in the basic model used for analysis. Thus, this evidence is merely suggestive.

10   Conclusion

Design of the taxation of capital income needs to reflect many factors. The paper focuses on heterogeneity in savings rates, an important dimension of heterogeneity for tax setting. The paper uses a model with jobs, rather than one with individual worker choices of hours. Neither labor market model describes the nature of opportunities for all workers, making room for learning from both types of models. In an hours model, workers make changes in response to small changes in marginal taxes. In a jobs model, there are many workers who are not at the margin of switching to a different job. Their lack of labor supply response to small changes in taxes is important for tax policy and seems plausible for many workers.

In keeping with the optimal tax literature the objective function in this paper has been defined in terms of individual lifetime utilities. Rather than considering how to weight the utilities of those with different discount factors (or preference differences more generally), we have reported some results in terms of different weights.[16] And we have used a model where the social objective function respects all preferences, not allowing for concerns that some people save too little for their own good.[17] Moving from this analysis toward concrete policy recommendations calls for addressing these issues as well as the more complex need to move from analyses based on lifetimes to taxes that are set annually, primarily on annual tax bases.

11   Appendix: Three-period model

Preferences are given by

$$u\left[x_1\right] + \delta u\left[x_2\right] + \delta^2 u\left[c\right] - v\left[z_1/n_1\right] - \delta v\left[z_2/n_2\right],$$

----

[16] A number of other papers have considered optimal taxes with heterogeneous preferences. See, for example, Blomquist and Christiansen, 2004, Boadway et al, 2002, Cuff, 2000, Kaplow, 2008, Sandmo, 1993, Tarkiainen and Tuomala, 2007.

[17] Addressing this concern would include consideration of mandatory retirement income programs and the effect of the design of savings incentives on consumer behavior, beyond the standard model of lifetime utility maximization.

with $u[x] = \log[x]$ and $v[z/n] = k(z/n)^{\beta+1}/(\beta+1)$, $\beta > 0$. With net-of-tax earnings written as a function of gross earnings, the budget constraint is

$$x_1 + R^{-1}x_2 + R^{-2}c - y_1(z_1) - R^{-1}y_2(z_2) = 0.$$

From the FOC, we find

$$
\begin{aligned}
x_1 &= \lambda^{-1} \\
x_2 &= \delta R \lambda^{-1} \\
c &= \delta^2 R^2 \lambda^{-1} \\
k(z_1/n_1)^{\beta} &= \lambda n_1 y_1' \\
\delta k(z_2/n_2)^{\beta} &= R^{-1}\lambda n_2 y_2'.
\end{aligned}
$$

Assume that the marginal tax rates on earnings are the same in both periods, so that an increase in output increases the earnings at the same rate: $y_1' = y_2'$. Then the time series of behavior satisfies

$$
\begin{aligned}
\frac{x_2}{x_1} &= \frac{c}{x_2} = \delta R \\
\frac{z_2/n_2}{z_1/n_1} &= (\delta R)^{-1/\beta}\left(\frac{n_2}{n_1}\right)^{1/\beta}.
\end{aligned}
$$

References

[1] 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.

[2] Banks, James and Peter Diamond. 2008. The Base for Direct Taxation, written for Reforming the Tax System for the 21st Century: The Mirrlees Review, Institute for Fiscal Studies. http://www.ifs.org.uk/mirrleesreview/reports/base.pdf.

[3] Blomquist, Sören and Vidar Christiansen. 2008. "Taxation and Heterogeneous Preferences." FinanzArchiv, 64: 218-244.

[4] Boadway, Robin, Maurice Marchand, Pierre Pestieau, and María del Mar Racionero. 2002. "Optimal Redistribution with Heterogeneous Preferences for Leisure." Journal of Public Economic Theory, 4(4): 475-498.

[5] Chabris, Christopher F., David Laibson, Carrie L. Morris, Jonathan P. Schuldt, and Dmitry Taubinsky. 2008. "Individual Laboratory-Measured Discount Rates Predict Field Behavior." NBER Working Paper 14270.

[6] Cuff, Katherine. 2000. "Optimality of workfare with heterogeneous preferences." Canadian Journal of Economics, 33(1): 149-74.

[7] Diamond, Peter. 2003. Taxation, Incomplete Markets, and Social Security. Cambridge: MIT Press.

[8] 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

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

[10] Gordon, Roger H. 2004 "Taxation of Interest Income." International Tax and Public Finance, 11(1): 5-15.

[11] Gordon, Roger H. and Wojciech Kopczuk. 2008. "The Choice of the Personal Income Tax Base." working paper

[12] Kaplow, Louis. 2008. "Optimal Policy with Heterogeneous Preferences." B.E. Journal of Economic Analysis and Policy, Vol. 8, Iss. 1 (Advances), Article 40: 1-28.

[13] Kocherlakota, N. 2005. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation." Econometrica, 73(5): 1587-1621.

[14] Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." Review of Economic Studies, 38: 175-208.

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

[16] Sandmo, Agnar. 1993. "Optimal redistribution when tastes differ." Finanz Archiv, 50(2): 149-163.

[17] Tarkiainen, Ritva and Matti Tuomala. 2007. "On optimal income taxation with heterogeneous work preferences." International Journal of Economic Theory, 3(1): 35 - 46.

[18] Tenhunen, Sanna and Matti Tuomala. 2008. "On Optimal Lifetime Redistribution Policy." forthcoming Journal of Public Economic Theory