On the desirability of taxing capital income in optimal social insurance☆

Bas Jacobs [a,c,*], Dirk Schindler [b,c]

[a] *Erasmus University Rotterdam, Tinbergen Institute, The Netherlands*
[b] *Universität Konstanz, Germany*
[c] *CESifo, Germany*

A R T I C L E   I N F O

A B S T R A C T

This paper analyzes optimal linear taxes on labor income and savings in a two-period life-cycle model with ex ante identical households, endogenous leisure demands in both periods, and general processes of skill shocks over the life cycle. We demonstrate that the Atkinson–Stiglitz theorem breaks down under risk. Capital taxes are employed besides labor income taxes for two distinct reasons: i) capital taxes reduce labor supply distortions on second-period labor supply, since second-period labor supply and saving are substitutes, ii) capital taxes insure first-period income risk, although this benefit is partially off-set because first-period labor supply and saving are complements. Our results imply that (retirement) saving should not be actuarially fair.

© 2012 Elsevier B.V. All rights reserved.

"[T]he conventional argument ignores the possibility that a tax on interest income might be desirable in order to offset the distortions introduced by a tax on labour earnings" (Atkinson and Sandmo, 1980, p. 529).

1. Introduction

Should capital income be taxed or not? This is one of the oldest and most important questions in public finance. Ever since the seminal work of Pigou (1928), the desirability of taxing capital income has been a controversial issue. And, as of today, controversies still abound

in the economics literature. Mirrlees et al. (2011) recommend in the final chapter of the *Mirrlees Review* not to tax the (normal) returns on saving, while Banks and Diamond (2010), writing a chapter in that very same *Mirrlees Review*, argue in favor of taxing the (normal) return to savings. Similarly, Mankiw et al. (2009) in the *Journal of Economic Perspectives* argue that capital income should remain untaxed, whereas Diamond and Saez (2011), in that very same journal, strongly recommend taxing capital incomes.

Our paper contributes to this long-standing debate in public finance by highlighting the role of non-insurable labor income risks. We demonstrate that under risk the optimal capital tax is always non-zero, and should be positive under empirically grounded conditions. We believe that our findings have potentially important policy implications for the debate on whether capital income should be taxed, how retirement savings should be taxed, and how pension reforms should be designed.

The argument against taxing capital incomes relies on two strands in the literature. The first strand originates from Chamley (1986) and Judd (1985) who analyze deterministic infinite-horizon models. These authors show that taxes on capital income can be seen as an ever-increasing tax on consumption in the more distant future. Ramsey principles therefore insist that in the long-run capital income should not be taxed. Since taxes on capital incomes are differentiated consumption taxes, these results are intimately linked to the second strand in the literature, which analyzes the desirability of differentiated commodity taxes. In particular, Sandmo (1974, 1976), Atkinson and Stiglitz (1976), and Deaton (1979) have demonstrated that commodity taxes should not be differentiated, even in finite-horizon models, as

long as preferences over consumption goods are weakly separable from leisure under non-linear income taxation. The requirements are stronger when only linear instruments are available. In that case, the subutility function over consumption goods needs to be homothetic as well. This result is generally referred to in the literature as the Atkinson–Stiglitz theorem.[1]

Our paper investigates the desirability of capital-income taxes when insurance markets are missing and individuals are subject to earnings risk. To that end, we develop a two-period life-cycle model where individuals optimally decide on consumption and leisure choices in both periods. Individuals could be hit by a non-insurable skill shock in each period of their life cycle. Ex ante all individuals are identical. Ex post they differ due to the realizations of these skill shocks. We allow for completely general skill-processes that could feature persistence or mean reversion over the life cycle. Capital markets are assumed to be perfect. A government with full commitment designs an optimal social insurance package. Since the government is unable to verify the skill shocks, it cannot employ state-dependent instruments. Consequently, individualized lump-sum taxes are ruled out, and the government has to resort to distortionary tax instruments. In addition, we assume that taxes on labor and capital incomes are linear and age-independent. Designing an optimal social insurance policy is thus a second-best problem and the government needs to trade-off incentives to work or to save against the benefits of social insurance.

We demonstrate that capital-income taxes are optimally non-zero in an optimal social insurance package in a wide class of standard two-period life-cycle models with risk. Therefore, capital-income taxes should always be employed even when adopting preferences that render capital-income taxes zero in the absence of risk. This result demonstrates that the Atkinson and Stiglitz (1976) theorem breaks down under risk. In addition, capital-income taxes should be positive under weak conditions that are likely to be fulfilled in practice.

We identify three mechanisms that determine whether capital-income taxes should optimally be employed in an optimal social insurance program: i) reducing labor-tax distortions on second-period labor supply, ii) (co-)insuring first-period labor-income risk, and iii) containing intertemporal labor supply distortions. To understand these three mechanisms, we will analyze two sub-models of our more general model.

First, we assume that labor supply in the first period is exogenous. In this setting, the second-period labor supply can be interpreted as the retirement decision. We will refer to this case as the 'working-for-retirement' model. It corresponds to the setting analyzed in Cremer and Gahvari (1995a) and also resembles the models of Diamond and Mirrlees (1978, 1986) in which individuals can retire early. The first role of taxing capital income is to off-set the tax distortions on labor supply (retirement), since a lower level of saving stimulates labor supply (later retirement) due to intertemporal wealth effects. Consequently, capital-income taxes directly alleviate the distortions of labor-income taxation on labor supply. The government trades off lower distortions in labor supply against larger distortions in saving.

Second, we analyze the case where second-period labor supply is assumed to be exogenous, and individuals only choose leisure in the first period. This is what we call the 'saving-for-retirement' model. This case extends the models analyzed by Ordover and Phelps (1979) and Atkinson and Sandmo (1980) to a setting with income risk. In the saving-for-retirement model, subsidies on saving would reduce distortions in labor supply of the young workers. Based on the same reasoning from the working-for-retirement model, one could therefore be tempted to conclude that subsidies on saving would be optimal, as this would alleviate tax distortions on labor supply. However, this turns out not to be the case. The reason is that capital-income taxes feature social insurance gains, since savings are stochastic, and reflect the severity of the skill shock individuals experience during the first period of their life cycle. Hence, the second role of capital-income taxes is to complement the labor-income tax in insuring income risk. The optimal capital-income tax trades off the distortions in both saving *and* labor supply, on the one hand, against the social gains of income insurance, on the other hand. Taxes on saving are thus optimal so as to smooth the dead-weight costs of social insurance over both the labor and capital tax bases.

In our full model, leisure demands are endogenous in both periods of the life cycle. The optimal capital tax tends to be positive for both reasons discussed in the two special cases. However, since both leisure demands are now endogenous, the capital tax entails an additional distortion in the intertemporal leisure decision, besides the saving decision. This third impact of capital-income taxes tends to reduce the optimal capital tax. The optimal capital tax always remains unambiguously positive, however, as long as life-time labor supply increases with the capital tax. This is the case as long as second-period labor supply increases more than first-period labor supply decreases. Empirical evidence suggests that this condition indeed holds.

Our paper contributes to the existing literature in a number of ways. First, by restricting the analysis to linear instruments without record keeping, and allowing for general skill-processes, we add relevant real-world features to the analysis of optimal capital-income taxation under risk. Non-linear policies have been extensively analyzed in, for example, Diamond and Mirrlees (1978, 1986) and papers in the 'new dynamic public finance' literature; see, e.g., da Costa and Werning (2002), Golosov et al. (2003, 2006), Kocherlakota (2005), Golosov and Tsyvinski (2006), Albanesi and Sleet (2006), and Diamond (2006). These papers show that labor supply optimally carries a wedge (i.e., is distorted) for insurance purposes. Moreover, there is an intertemporal wedge in consumption choices, indicating a potential role for capital-income taxation. However, non-linear instruments are very demanding in terms of information, as they require verifiability of labor incomes and savings at the individual level. Furthermore, in dynamic optimal-tax models with risk, optimal second-best allocations cannot be implemented with standard non-linear tax schedules unless specific assumptions are made on the dynamics of the skill process or on the set of available government instruments, such as record keeping. See, for example, Golosov et al. (2003), Kocherlakota (2005), Albanesi and Sleet (2006), Golosov and Tsyvinski (2006), and Blomquist and Micheletto (2008). In the real world, capital-income taxes are generally not based on record keeping and are often linear. Governments generally do not keep tax records, even in most advanced countries. Moreover, Blomquist and Micheletto (2008) and Bastani et al. (2010) argue that non-linear taxes on savings introduce arbitrage possibilities, which cannot be eliminated if the government cannot verify capital incomes at the individual level.[2] By analyzing linear instruments we avoid these implementability issues. Moreover, as we do not need to worry about implementation issues, we can allow for completely general

---

[1] Numerous other papers have elucidated the conditions under which capital-income taxes are not optimally zero in deterministic models. If horizons are not infinite and preferences do not meet the required separability conditions, capital income might be taxed or subsidized on a net basis. In particular, when marginal rate of substitution between future and current consumption decreases with labor effort, capital incomes should optimally be taxed so as to (partially) off-set the tax distortions of the income tax on labor supply. See for example Ordover and Phelps (1979), Atkinson and Sandmo (1980), Erosa and Gervais (2002), and Diamond (2006). Aiyagari (1995) allows for incomplete financial markets such that individuals can be borrowing constrained. Capital-income taxes redistribute resources from unconstrained toward constrained phases of the life-cycle, and thereby help to complete missing borrowing markets. Saez (2002), Boadway and Pestieau (2003), Diamond (2006), and Blomquist and Christiansen (2008) allow for heterogeneous preferences. They show that when discount rates decrease with ability, it is optimal to tax capital income in a redistributive program even under separable preferences. In case governments cannot commit to future tax plans, optimal time-consistent capital taxes might also be (very) high, see, for example, Kydland and Prescott (1977) and Fischer (1980).

[2] These authors analyze (age-dependent) non-linear labor-income taxes and also demonstrate that there is generally a role for (linear) capital taxation.

skill processes. We are able to demonstrate that the optimality of a cap-
ital wedge is robust to (very) large deviations from the informational
requirements adopted in non-linear tax frameworks.

Second, in a series of seminal papers, Cremer and Gahvari (1995a,b,
1999a,b) have investigated the desirability of commodity tax differ-
entiation in risky environments. Using linear policies, Cremer and
Gahvari (1995a) have shown that the Atkinson–Stiglitz theorem fails
in a special case of our more general model. In particular, Cremer and
Gahvari (1995a) argue that commodity tax differentiation helps to off-
set over- or underconsumption – relative to the first-best rules – of pre-
committed and post-committed goods, i.e., goods that are consumed
before or after the skill shock materializes. Translated to our setting,
this would imply that the government would like to tax precautionary
saving. However, in our view, their explanation for this result needs to
be revised. We demonstrate that in their setting, the capital tax does
not reduce the exposure to labor market risk. Hence, the capital tax
has no insurance gains, while upsetting the optimal private response
to earnings risk by taxing savings in a distortionary way. Instead, we
show that the capital tax boosts labor supply, and thereby reduces
labor supply distortions in social insurance. Consequently, positive cap-
ital taxes are optimal to reduce labor market distortions, and are not
employed to reduce precautionary saving.[3,4]

Third, by analyzing the optimal tax treatment of (pension) saving,
we also contribute to the analysis of Cremer et al. (2004, 2008) who
studied the optimal taxation of retirement, but did not analyze the
optimal tax treatment of saving. Our results strengthen their findings
by demonstrating that not only retirement choices should be distorted,
but also that (private) retirement saving should optimally be actuarially
unfair. A tax on saving alleviates the distortions in early retirement
choices caused by social insurance. This finding has substantial policy
relevance. In the upcoming decades, many countries are confronted
with the aging of work forces, resulting in financing problems for
PAYG-pensions and health care. Our results indicate that if govern-
ments aim to promote later retirement, they should not strengthen
incentives to save for retirement at the same time. Stronger incentives
for retirement saving would promote earlier retirement, not later re-
tirement. Similarly, if governments would like to promote labor supply
of working-age individuals, they should not stimulate (pension) savings
either. For a given level of social insurance, the rise in the tax burden
needed to compensate the saving subsidies would reduce labor supply
of working-age individuals more than the saving subsidies could offset.
Thus, the trade-off between incentives and insurance worsens.

The remainder of this paper is structured as follows. Section 2 pre-
sents the baseline model. Section 3 derives the optimal tax rules for
optimal labor and capital taxes. Section 4 derives the optimal tax struc-
ture in the 'saving-for-retirement' model. Section 5 derives the optimal
tax structure in the 'working-for-retirement' model. Section 6 gives the
solution to the complete model. Section 7 concludes. Three appendices
contain technical details of the derivations.

2. Model

2.1. Households

There is a continuum of infinitely small households who live for two
periods. In each period households decide upon their consumption and

labor supply. Perfect capital markets allow individuals to borrow and
lend at constant real interest rate r. In addition, labor markets are fric-
tionless and the wage per efficiency unit of labor equals one.[5] Insurance
markets to insure idiosyncratic labor income risks are missing, which
can be due to moral hazard, adverse selection, and contract incomplete-
ness (see, e.g., Sinn, 1996). By the law of large numbers idiosyncratic in-
dividual risk washes out in aggregate and there is no aggregate
(systematic) risk.

Households are identical ex ante, but not ex post. In each period
$i = 1,2$, their productivity per hour worked or 'skill' θ_i is stochastic.[6]
The joint set of possible realizations is denoted by $\Theta \equiv [\underline{\theta}_1, \bar{\theta}_1] \times$
$[\underline{\theta}_2, \bar{\theta}_2]$, where $\underline{\theta}_1 > 0$ and $\underline{\theta}_2 > 0$. It does not matter whether $\bar{\theta}_i$ is
bounded or infinite. $\theta \equiv \{\theta_1, \theta_2\} \in \Theta$ denotes a skill history of θ_1 and θ_2.
We will denote by $\Theta_i \equiv [\underline{\theta}_i, \bar{\theta}_i]$ the set of realizations of θ_i for $i = 1,2$.
$p(\theta)$ is the probability distribution function, which attaches a strictly
positive probability $p(\theta)$ to skill history θ. The conditional probability
that θ_2 is realized given θ_1 is denoted by $p(\theta_2 \mid \theta_1)$. The (life-time) ex-
pectation $E[.]$ over variable $x(\theta)$ as of period one is defined as
$E[x(\theta)] \equiv \sum_\Theta x(\theta)p(\theta)$, whereas the conditional expectation of a variable
as of period two, given a particular realization of the skill shock θ_1 in
period one, is denoted by $E[x(\theta_2)|\theta_1] \equiv \sum_{\theta_2} x(\theta_2)p(\theta_2|\theta_1)$. We allow for
fully general stochastic processes for the evolution of skills; hence,
there could be persistence or mean reversion in skill shocks over time.
There is strong empirical evidence for persistence in incomes of individ-
uals, see, e.g., Meghir and Pistaferri (2011). Persistence in income
shocks has significant effects on welfare as these shocks cannot be
perfectly smoothed out over the life cycle, in contrast to temporary,
idiosyncratic shocks. Furthermore, persistence in skills has important
consequences for the implementability of non-linear tax schedules,
see the introduction. While we allow for completely general skill
processes, we are able to implement allocations using simple, linear
tax instruments. For notational simplicity we harmlessly normalize
the expectation of the first skill shock to one: $E[\theta_1] \equiv 1$.

c_i denotes consumption in period $i = 1,2$. Similarly, l_i is labor supply
in period i. In period one, households choose labor supply and con-
sumption before the shock realizes, hence c_1 and l_1 are 'committed'
goods (Cremer and Gahvari, 1995a,b). When entering the second peri-
od, households carry forward a stochastic level of assets $a(\theta_1)$ and first
determine how much labor $l_2(\theta_1)$ to supply. Hence, second-period
labor supply only depends on shock θ_1 and not on θ_2. Second-period
consumption $c_2(\theta_1, \theta_2)$ is determined residually.

This particular sequencing of decisions and skill shocks has been
chosen based on the following considerations. First, the model would
collapse to a standard heterogeneous-agent model (without uncertain-
ty), if neither first-period consumption nor first-period leisure would be
committed before the first skill shock realizes (cf. Cremer and Gahvari,
1999a). We follow Cremer and Gahvari (1995a,b, 1999a,b) by assuming
that first-period consumption is the committed good.[7] This implies that
saving becomes stochastic, and absorbs part of the first-period skill
shock. Second, second-period consumption should be determined
residually, i.e., after the second skill shock realizes, so as to close the
model. This leaves us with a choice as to when labor supply is chosen
in both periods: either before or after the skill shocks realize. It turned
out to be technically slightly simpler to assume that labor supply is

chosen before the skill-shock realizes in each period, while doing so is without loss of generality.[8]

We assume that expected utility U is an additively separable function over consumption and labor supply in both periods (see also, e.g., Cremer and Gahvari, 1995a,b; Golosov et al., 2003, 2006; Diamond, 2006):

$$U \equiv u_1(c_1) - v_1(l_1) + \beta E[u_2(c_2(\theta_1, \theta_2)) - v_2(l_2(\theta_1))],$$
$$u_i', v_i' > 0, \quad u_i'', -v_i'' < 0, \quad u_i''' > 0, \quad 0 < \beta < 1, \quad i = 1, 2, \tag{1}$$

where sub-utilities u_i and v_i satisfy the Inada conditions. β is the discount factor, which captures the time-preference of the household. We assume decreasing absolute risk aversion in consumption, which necessarily implies $u_i''' > 0$. This utility function satisfies the conditions for the Atkinson–Stiglitz theorem in the absence of risk if sub-utility over consumption is homothetic. Hence, if we find a role for capital-income taxation assuming homothetic consumption preferences we directly establish that the Atkinson–Stiglitz theorem breaks down.

The government cannot verify individual skill shocks. Therefore, it is unable to employ individualized lump-sum taxes. Instead, it has to resort to taxing verifiable labor and capital incomes. Additionally, the government may employ a non-individualized lump-sum tax (transfer). Thus, the government cannot insure skill shocks without incurring efficiency losses in labor supply and saving. Hence, the optimal insurance problem features the well-known trade-off between insurance and incentives. We restrict the analysis to age-independent linear instruments.[9] The informational requirements for linear instruments are that the government only observes aggregate tax bases. In particular, the government levies a linear tax on labor earnings in both periods at rate t. In addition, the household receives a transfer T in the first period. We do not explicitly allow for a second-period income transfer. This instrument is redundant, since individuals can freely allocate the first-period transfer over the life cycle by having perfect access to capital markets. Finally, a linear tax at rate τ is levied on interest income from savings.[10]

In the first period, the household works and earns $\theta_1 l_1$ in gross labor income. The first-period budget constraint states that total consumption equals net labor income minus saving $a(\theta_1)$:

$$c_1 = (1 - t)\theta_1 l_1 + T - a(\theta_1), \quad \forall \theta_1 \in \Theta_1. \tag{2}$$

In the second-period, the household earns net labor income $(1 - t)\theta_2 l_2(\theta_1)$ and interest income $ra(\theta_1)$ on assets carried forward from period one. Interest income is taxed at flat rate τ. Hence, the second-period budget reads as

$$c_2(\theta_2, \theta_1) = (1 - t)\theta_2 l_2(\theta_1) + (1 + (1 - \tau)r)a(\theta_1), \quad \forall \theta \in \Theta. \tag{3}$$

In the remainder, we will employ $R \equiv 1 + (1 - \tau)r$ to denote the net interest factor.

The household maximizes life-time utility by choosing the optimal levels of consumption c_i and labor supply l_i. We solve this problem backwards. Individuals enter the second period with a stochastic level of assets $a(\theta_1)$. Given this level of assets, and before the second shock θ_2 materializes, the individual solves the subprogram:

$$\max_{\{l_2(\theta_1)\}} E[u_2((1 - t)\theta_2 l_2(\theta_1) + Ra(\theta_1)) - v_2(l_2)|\theta_1], \quad \forall \theta_1 \in \Theta_1, \tag{4}$$

which yields the following first-order condition for the second-period labor supply:

$$(1 - t)E[u_2'(\theta_2)\theta_2|\theta_1] = v_2'(l_2(\theta_1)), \quad \forall \theta_1 \in \Theta_1. \tag{5}$$

Consequently, we can write for the conditional expectation of second-period indirect utility:

$$E[W(\theta_2, a(\theta_1))|\theta_1] \equiv E\left[u_2(\hat{c}_2) - v_2(\hat{l}_2)\Big|\theta_1\right], \quad \forall \theta_1 \in \Theta_1, \tag{6}$$

where hats are used to denote the optimal values of c_2 and l_2. Taking expectations as of period one on both sides yields expected indirect utility in period two as a function of saving and the skill shocks:

$$E[Wa((\theta_1), \theta_1, \theta_2)] \equiv E\left[u_2(\hat{c}_2(\theta_1, \theta_2)) - v_2\left(\hat{l}_2(\theta_1)\right)\right]$$
$$= E\left[u_2\left((1 - t)\theta_2\hat{l}_2(\theta_1) + Ra(\theta_1)\right) - v_2\left(\hat{l}_2(\theta_1)\right)\right]. \tag{7}$$

Straightforward differentiation yields $\dfrac{\partial E[W(a(\theta_1), \theta_1, \theta_2)]}{\partial a(\theta_1)} = RE[u_2'(c_2(\theta_1, \theta_2))]$.

In the first stage, individuals choose c_1 and l_1 before the shock θ_1 realizes, conditional upon optimal choices in the second period. Hence, the individual solves the following subprogram:

$$\max_{\{c_1, l_1\}} U = u_1(c_1) - v_1(l_1) + \beta E[W(a(\theta_1), (\theta_1, \theta_2)]$$
$$= u_1(c_1) - v_1(l_1) + \beta E[W((1 - t)\theta_1 l_1 + T - c_1, \theta_1, \theta_2)], \tag{8}$$

where we substituted saving from the individual budget constraint in Eq. (2) in the second line. The first-period labor supply decision is governed by

$$v_1'(l_1) = (1 - t)\beta R E[u_2'(c_2(\theta_1, \theta_2))\theta_1]. \tag{9}$$

The first-order conditions also imply the standard stochastic Euler equation for consumption:

$$u_1'(c_1) = \beta R E[u_2'(c_2(\theta_1, \theta_2))]. \tag{10}$$

A higher real return on saving R, or a higher discount factor β, strengthens the incentive to save by substituting current for future consumption.[11]

We introduce the risk premia of first- and second-period labor supply as the normalized covariance between the marginal utility of second-period consumption and the skill shocks θ_1 and θ_2:

$$\pi_1 \equiv -\frac{cov[u_2'(c_2(\theta_1, \theta_2)), \theta_1]}{E[u_2'(c_2(\theta_1, \theta_2))]E[\theta_1]} \geq 0, \tag{11}$$

$$\pi_2 \equiv -\frac{cov[u_2'(c_2(\theta_1, \theta_2)), \theta_2]}{E[u_2'(c_2(\theta_1, \theta_2))]E[\theta_2]} \geq 0. \tag{12}$$

π_i denotes the marginal welfare loss due to skill risk in period i as a fraction of $E[\theta_i]$. Indeed, $(1 - \pi)E[\theta_i]$ denotes the certainty equivalent

---

[8] We have also derived the model when labor supply in each period is chosen after the shock has realized, and the optimal tax expressions basically remain the same. In that case, they contain the expected elasticity of first-period labor income rather than the deterministic elasticity, and the expected elasticity of second-period labor also depends on the second skill shock.

[9] An unpublished appendix – available upon request from the authors – demonstrates that the main findings are stronger if the government would be allowed to set age-dependent tax rates on labor income. Intuitively, allowing for age-dependent labor-income taxes allows the government to off-set the distortions of capital-income taxes on intertemporal labor supply choices by adjusting first-period labor taxation, thereby alleviating the distortion of the capital-income tax. Intertemporal distortions in labor supply will therefore no longer reduce the attractiveness of employing capital-income taxes. For given first-period insurance benefits and reductions in second-period labor-market distortions, optimal capital-income taxes will be higher. Indeed, capital taxes can be shown to be unambiguously positive under age-dependent labor-income taxation. See also the discussion at the end of Section 6.

[10] Cremer and Gahvari (1995a) study a similar setting using differentiated commodity taxes. In the absence of non-labor income, such as bequests, uniform commodity taxes are equivalent to a proportional tax on labor income, without taxes on capital income. Non-uniform commodity taxes are equivalent to a labor income tax supplemented with taxes or subsidies on capital income.

[11] Second-order conditions are always fulfilled due to the assumptions on preferences.

of $E[\theta_i]$. Because marginal utility of income is declining with income, the risk premia are non-negative in both periods. Given that risk affects labor earnings in a multiplicative way, larger labor supply raises the risk-exposure of households to labor market shocks.

Using these definitions, and recalling that $E[\theta_1] = 1$, we can derive that the labor supply equations in both periods can be written as:

$$\frac{v_1'(l_1)}{u_1'(c_1)} = (1-\pi_1)(1-t), \tag{13}$$

$$\frac{E[v_2'(l_2(\theta_1))]}{E[u_2'(c_2(\theta_1,\theta_2))]} = (1-\pi_2)(1-t)E[\theta_2]. \tag{14}$$

Hence, individuals get stronger incentives to supply more labor if the tax rate is lower or if labor income is less risky (lower π_i). Larger labor market risk, as indicated by a larger π_i, acts as an implicit tax on labor supply, since risk averse individuals reduce their labor effort if the latter raises their exposure to skill shocks.

Indirect expected utility of the household can be written as a function V over the policy variables (T, t, R):

$$V(T,t,R) \equiv u_1(\hat{c}_1) - v_1(\hat{l}_1) + \beta E\left[u_2(\hat{c}_2) - v_2(\hat{l}_2)\right], \tag{15}$$

where the hats indicate the optimized values for consumption and labor, which follow from solving the three first-order conditions (5), (9), and (10), and the household budget constraints (2) and (3) for c_1, c_2, l_1, l_2, and a. Note that we have suppressed the skill shocks for notational simplicity. We will continue to do so in the remainder of the paper.

The derivatives of indirect utility with respect to the policy instruments follow from applying the envelope theorem (see Appendix A):

$$\frac{\partial V}{\partial T} = \eta, \tag{16}$$

$$\frac{\partial V}{\partial t} = -\eta\left((1-\xi_1)l_1 + \frac{(1-\xi_2)E[\theta_2 l_2]}{R}\right), \tag{17}$$

$$\frac{\partial V}{\partial R} = \eta\frac{((1-\xi_1)(1-t)l_1 - c_1 + T)}{R}, \tag{18}$$

where $\eta(c_2) \equiv u_1'(c_1) = \beta RE[u_2']E[u_2'(c_2)]$ is the marginal utility of private income, and ξ_1 and ξ_2 are the insurance characteristics of first and second-period labor incomes:

$$\xi_1 \equiv -\frac{cov[u_2', \theta_1 l_1]}{E[u_2']E[\theta_1 l_1]} \geq 0, \tag{19}$$

$$\xi_2 \equiv -\frac{cov[u_2', \theta_2 l_2]}{E[u_2']E[\theta_2 l_2]} \geq 0. \tag{20}$$

The insurance characteristic ξ_i gives the marginal welfare loss of income risk in period i as a fraction of income in period i. In particular, $(1-\xi_i)E[\theta_i l_i]$ is the certainty equivalent of risky labor income $\theta_i l_i$.

To solve for the optimal tax structure below, we employ the risk-adjusted Slutsky equations. To that end, we define the expenditure function $X(t,R,V)$ as the minimum level of non-labor income T required to attain expected indirect utility V. $X(.)$ can be obtained from setting $X(t,R,V) \equiv T$ for the optimal level of indirect utility V as given in Eq. (15). The compensated demand functions are then defined as

$$l_i^c(t,R,V) \equiv l_i(t,R,X(t,R,V)), \tag{21}$$

$$c_i^c(t,R,V) \equiv c_i(t,R,X(t,R,V)), \tag{22}$$

where the superscript c denotes a compensated change. By totally differentiating the compensated demand functions for given V, we

obtain the following risk-adjusted Slutsky equations for l_1, l_2, and c_1 with respect to t and R (see Appendix B):

$$\frac{\partial l_1}{\partial t} = \frac{\partial l_1^c}{\partial t} - \left((1-\xi_1)l_1 + \frac{(1-\xi_2)E[\theta_2 l_2]}{R}\right)\frac{\partial l_1}{\partial T}, \tag{23}$$

$$\frac{\partial l_2}{\partial t} = \frac{\partial l_2^c}{\partial t} - \left((1-\xi_1)l_1 + \frac{(1-\xi_2)E[\theta_2 l_2]}{R}\right)\frac{\partial l_2}{\partial T}, \tag{24}$$

$$\frac{\partial c_1}{\partial t} = \frac{\partial c_1^c}{\partial t} - \left((1-\xi_1)l_1 + \frac{(1-\xi_2)E[\theta_2 l_2]}{R}\right)\frac{\partial c_1}{\partial T}, \tag{25}$$

$$\frac{\partial l_1}{\partial R} = \frac{\partial l_1^c}{\partial R} + \frac{((1-\xi_1)(1-t)l_1 - c_1 + T)}{R}\frac{\partial l_1}{\partial T}, \tag{26}$$

$$\frac{\partial l_2}{\partial R} = \frac{\partial l_2^c}{\partial R} + \frac{((1-\xi_1)(1-t)l_1 - c_1 + T)}{R}\frac{\partial l_2}{\partial T}, \tag{27}$$

$$\frac{\partial c_1}{\partial R} = \frac{\partial c_1^c}{\partial R} + \frac{((1-\xi_1)(1-t)l_1 - c_1 + T)}{R}\frac{\partial c_1}{\partial T}. \tag{28}$$

2.2. Government

We assume a benevolent government, which has full commitment. We abstract from a government-revenue requirement without loss of generality. The government optimally provides social insurance by choosing policy instruments T, t, and R, such that expected indirect utility $V(T,t,R)$ of the household is maximized.

By the law of large numbers, individual idiosyncratic risks cancel in the aggregate. The government is able to borrow in a perfect capital market at real interest rate r. Hence, we find that the intertemporal government budget constraint is given by

$$(1+r)tl_1 + tE[\theta_2 l_2] + (1+r-R)[(1-t)l_1 - c_1 + T] = (1+r)T. \tag{29}$$

All labor incomes are deterministic at the aggregate level. However, this neither implies that the expectations operators have on second-period labor income and on the second-period skill shock can be separated, nor that $E[\theta_2]$ can be normalized to one. The reason is that skill shocks θ_i may not be independent over time. If there is a correlation between both skill shocks, second-period income will depend on the realization of the first-period shock θ_1 and the second-period shock θ_2. As a result we have $E[\theta_2 l_2(\theta_1)] \neq E[\theta_2]E[l_2(\theta_1)]$. Only if skill shocks are independent, i.e., if $cov[\theta_1,\theta_2] = 0$, we obtain $E[\theta_2 l_2(\theta_1)] = E[\theta_2]E[l_2(\theta_1)]$.

3. Optimal taxation

The Lagrangian for maximizing social welfare is given by

$$\max_{\{T,t,R\}} \mathcal{L} \equiv V(T,t,R) + \lambda[tl_1(1+r) + tE[\theta_2 l_2]] \\ + \lambda[(1+r-R)((1-t)l_1 - c_1 + T) - (1+r)T], \tag{30}$$

where λ is the deterministic shadow value of public resources.

The first-order conditions for an optimum are[12]

$$\frac{\partial \mathcal{L}}{\partial T} = \beta RE[u_2'] - \lambda R + \lambda(tR + \tau r)\frac{\partial l_1}{\partial T} + \lambda t E\left[\theta_2 \frac{\partial l_2}{\partial T}\right] - \lambda \tau r\frac{\partial c_1}{\partial T} = 0, \tag{31}$$

$$\frac{\partial \mathcal{L}}{\partial t} = -\beta E[u_2']((1-\xi_1)Rl_1 + (1-\xi_2)E[\theta_2 l_2]) + \lambda(Rl_1 + E[\theta_2 l_2]) \\ + \lambda(tR + \tau r)\frac{\partial l_1}{\partial t} + \lambda t E\left[\theta_2 \frac{\partial l_2}{\partial t}\right] - \lambda \tau r\frac{\partial c_1}{\partial t} = 0, \tag{32}$$

---

[12] As is standard in the optimal tax literature, we assume that these necessary first-order conditions are also sufficient to describe the optimum allocation, i.e., the second-order conditions for the government program are fulfilled.

$$\frac{\partial \mathcal{L}}{\partial R} = \beta E[u'_2]((1-\xi_1)(1-t)l_1 - c_1 + T) - \lambda((1-t)l_1 - c_1 + T)$$
$$+ \lambda(tR + \tau r)\frac{\partial l_1}{\partial R} + \lambda t E\left[\theta_2 \frac{\partial l_2}{\partial R}\right] - \lambda \tau r \frac{\partial c_1}{\partial R} = 0. \tag{33}$$

From the first-order condition for the lump-sum transfer in Eq. (31) follows that the expected social value of transferring one euro to the household (b) should be equal to its resource cost (unity):

$$b \equiv \frac{\beta E[u'_2]}{\lambda} + \frac{(tR + \tau r)}{R}\frac{\partial l_1}{\partial T} + \frac{t}{R}E\left[\theta_2 \frac{\partial l_2}{\partial T}\right] - \frac{\tau r}{R}\frac{\partial c_1}{\partial T} = 1. \tag{34}$$

Given that there is no revenue-requirement, all tax revenues are recycled in the form of transfers. Thus, the transfer T will always be positive as long as tax rates on labor and capital incomes are positive. Generally, it is not feasible to draw inferences about the size of the transfer, as this depends on the total level of taxation of labor and capital incomes. The analysis below does not permit us to compare tax levels across different (sub-)cases. We will not return to this discussion in the following sections, as the first-order condition for the demogrant always remains the same.

The first-order condition for the labor tax rate in Eq. (32) can be rewritten by substituting the risk-adjusted Slutsky equations for $\frac{\partial l_1}{\partial t}$, $\frac{\partial l_2}{\partial t}$, and $\frac{\partial c_1}{\partial t}$ in Eqs. (23), (24) and (25), using the definition for b in Eq. (34), and rearranging to find

$$\omega\xi_1 + (1-\omega)\xi_2 + \frac{t}{1-t}\left(\omega\varepsilon_{l_1 t} + (1-\omega)\varepsilon_{l_2 t}\right) + \frac{\tau r/R}{1-t}\left(\omega\varepsilon_{l_1 t} - \gamma\varepsilon_{c_1 t}\right) = 0, \tag{35}$$

where $\varepsilon_{l_1 t} \equiv \frac{\partial l_1^c}{\partial t}\frac{1-t}{l_1}$, $\varepsilon_{l_2 t} \equiv E\left[\theta_2 \frac{\partial l_2^c}{\partial t}\right]\frac{1-t}{E[\theta_2 l_2]}$, and $\varepsilon_{c_1 t} \equiv \frac{\partial c_1^c}{\partial t}\frac{1-t}{c_1}$ designate the compensated labor-tax elasticities of first-period labor income, expected second-period labor income, and first-period consumption, respectively. $\omega \equiv \frac{Rl_1}{Rl_1 + E[\theta_2 l_2]}$ is the share of first-period labor income in expected total labor income. $\gamma \equiv \frac{Rc_1}{Rl_1 + E[\theta_2 l_2]}$ is the share of first-period consumption in expected total labor income.

Similarly, we can simplify the first-order condition for the capital tax in Eq. (33) by substituting the risk-adjusted Slutsky equations for $\frac{\partial l_1}{\partial R}$, $\frac{\partial l_2}{\partial R}$, and $\frac{\partial c_1}{\partial R}$ (see Eqs. (26) to (28)), using the definition for b in Eq. (34), and rearranging to find

$$-\omega\xi_1 + \frac{t}{1-t}\left(\omega\varepsilon_{l_1 R} + (1-\omega)\varepsilon_{l_2 R}\right) + \frac{\tau r/R}{1-t}\left(\omega\varepsilon_{l_1 R} - \gamma\varepsilon_{c_1 R}\right) = 0, \tag{36}$$

where $\varepsilon_{l_1 R} \equiv \frac{\partial l_1^c}{\partial R}\frac{R}{l_1}$, $\varepsilon_{l_2 R} \equiv E\left[\theta_2 \frac{\partial l_2^c}{\partial R}\right]\frac{R}{E[\theta_2 l_2]}$, $\varepsilon_{c_1 R} \equiv \frac{\partial c_1^c}{\partial R}\frac{R}{c_1}$ denote the compensated elasticities of first-period labor income, expected second-period labor income, and first-period consumption with respect to the interest factor, respectively.

In Appendix C, we formally derive all the behavioral elasticities, which we have signed under three parameter restrictions, see also Table 1. Our parameter restrictions ensure that the elasticities qualitatively have the same signs as the comparative statics results of the model in the absence of income risk. Moreover, our restrictions ensure that the signs of the elasticities are empirically warranted. The imposed parameter restrictions are summarized in the last row of Table 1.

First, $\varepsilon_{c_2 R} > 0$ holds independently of any assumption on parameters. Hence, a larger net return on saving boosts second period consumption. Moreover, $\varepsilon_{c_1 R} < 0$, since we assume $\delta < 0$ so that the standard substitution

Table 1
Summary of elasticities.

Elasticities	
$\varepsilon_{c_1 t} \equiv -\frac{\epsilon}{\Delta} < 0$	$\varepsilon_{c_1 R} \equiv \frac{\delta}{\Delta} < 0$
$\varepsilon_{c_2 t} \equiv -\frac{\rho_1}{\rho_2}\frac{\epsilon}{\Delta} < 0$	$\varepsilon_{c_2 R} \equiv \frac{1}{\rho_2} + \frac{\rho_1}{\rho_2}\frac{\delta}{\Delta} > 0$
$\varepsilon_{l_1 t} \equiv -\varepsilon_1\left[1 - \sum_1 \frac{\rho_1 \epsilon}{\Delta}\right] < 0$	$\varepsilon_{l_1 R} \equiv -\varepsilon_1 \sum_1\left[1 + \frac{\rho_1 \delta}{\Delta} - \frac{1}{\Sigma_1}\right] > 0$
$\varepsilon_{l_2 t} \equiv -\varepsilon_2\left[1 - \sum_2 \frac{\rho_1 \epsilon}{\Delta}\right] < 0$	$\varepsilon_{l_2 R} \equiv -\varepsilon_2 \sum_2\left[1 + \frac{\rho_1 \delta}{\Delta}\right] < 0$

Definitions
$\rho_i \equiv -\frac{E[u''_i(c_i)]E[c_i]}{E[u'_i(c_i)]} > 0$: global relative risk aversion in consumption in period i
$\varepsilon_i \equiv \left[\frac{E[v''_i(l_i)]E[\theta_i l_i]}{E[v'_i(l_i)]E[\theta_i]}\right]^{-1} > 0$: compensated labor supply elasticity in period i
$\pi'_i \equiv -\frac{cov[u''_2, \theta_i]}{E[u''_2]E[\theta_i]} > 0$: 'prudence-based' risk premium in period i
$\sum_i \equiv \frac{1 - \pi'_i}{1 - \pi_i} \geq 0$: 'elasticity of residual risk aversion' in period i
$\Delta \equiv \frac{\gamma + (1-\gamma)\frac{\rho_1}{\rho_2}}{(1-t)} + (1-\pi_1)\omega\sum_1\varepsilon_1\rho_1 + (1-\pi_2)(1-\omega)\sum_2\varepsilon_2\rho_1 > 0$
$\varepsilon \equiv (1-\pi_1)\omega\varepsilon_1 + (1-\pi_2)(1-\omega)\varepsilon_2 > 0$
$\delta \equiv -\frac{(1-\gamma)/\rho_2}{(1-t)} + (1-\pi_1)\omega\varepsilon_1(1-\sum_1) - (1-\pi_2)(1-\omega)\varepsilon_2\sum_2$

Parameter restrictions
i) $\delta < 0$, ii) $\sum_1 \approx \sum_2$, iii) $\pi'_1 > \pi_1 \Leftrightarrow \sum_1 < 1$

effect in saving dominates the insurance effect of taxes on saving. The insurance effect stems from the fact that taxes on saving help to reduce the exposure to first-period labor market shocks by reducing the variance in saving. Bernheim (2002) surveys many empirical studies estimating the interest elasticity of saving. Empirical findings indicate that the compensated interest elasticity of saving is indeed positive.

Second, $\varepsilon_{l_1 t} < 0$ and $\varepsilon_{l_2 t} < 0$. Under wage risk, the elasticities of labor supply with respect to the labor tax are generally ambiguous. By reducing the variance in earnings, a higher tax reduces the risk-exposure of individuals to adverse labor market shocks so that labor supply is ceteris paribus stimulated (see also Menezes and Wang, 2005). The change in exposure to labor market risk is captured by the 'elasticity of residual risk aversion' $\sum_i \equiv \frac{1 - \pi'_i}{1 - \pi_i}$, $\pi'_i \equiv -\frac{cov[u''_i, \theta_i]}{E[u''_i]E[\theta_i]}$. This elasticity measures the percentage change in the certainty equivalent of wages with respect to a one-percent change in expected wages in period i.[13] However, the standard, negative substitution effect of higher taxes on labor supply pulls in the opposite direction. We assume that $\Sigma_1 \approx \Sigma_2$ so that the substitution effects in labor supply dominate the insurance effects. Blundell and MaCurdy (1999) and Meghir and Philips (2010) survey a great deal of empirical studies demonstrating that the compensated wage elasticities of labor supply are positive. If one would like to interpret the second-period labor supply as retirement, Gruber and Wise (1999) demonstrate that the retirement age falls if the implicit tax on continued work increases.

Third, $\varepsilon_{c_1 t} < 0$ and $\varepsilon_{c_2 t} < 0$. These are unambiguous. The intuition is that a higher labor tax lowers the price of leisure and induces substitution away from consumption toward leisure.

Fourth, $\varepsilon_{l_1 R} > 0$ and $\varepsilon_{l_2 R} < 0$. A higher financial return R induces individuals to have relatively more consumption and leisure in the second-period and less consumption and leisure in the first period. Due to intertemporal substitution in leisure, labor supply in the first

---

[13] Σ_i can be compared to the 'coefficient of residual income progression', which is the elasticity of after-tax income with respect to before-tax income, see, e.g., Musgrave and Musgrave (1976).

period increases and labor supply in the second period decreases. In addition, there are wealth effects on labor supply in both periods due to intertemporal substitution effects in consumption. Intuitively, a lower (higher) first-period (second-period) level of consumption raises (lowers) marginal utility of consumption in the first (second) period. Consequently, in the first period the marginal willingness to pay for leisure, i.e., the marginal rate of substitution between leisure and consumption, decreases and labor supply expands. Similarly, in the second period the marginal willingness to pay for leisure increases, so that labor supply diminishes. Thus, intertemporal substitution effects in both leisure and consumption increase first-period labor supply and decrease second-period labor supply. Moreover, in case of $\varepsilon_{l_1 R}$, the interest rate also has a direct, positive effect on the effective first-period wage rate by increasing its net present value in terms of second period consumption, which is the numéraire commodity. While $\varepsilon_{l_2 R} < 0$ can be signed independently of any assumption on parameters, $\varepsilon_{l_1 R}$ can turn ambiguous under risk. If $\delta < 0$, a sufficient condition for $\varepsilon_{l_1 R} > 0$ is that the 'elasticity of residual risk aversion' in the first period should be smaller than one, i.e., $\Sigma_1 \equiv \frac{1 - \pi_1'}{1 - \pi_1} < 1$, which is equivalent to assuming $\pi_1' > \pi_1$. This restriction is harmless when the bivariate distribution of skill shocks is normal and should also hold more generally under mild conditions (see Appendix B). Not many studies directly estimate the interest elasticity of labor supply. Nevertheless, Pirttilä and Suoniemi (2010) and Gordon and Kopczuk (2011) demonstrate that average labor supply falls with larger capital incomes. This suggests that the income-weighted elasticity $\omega \varepsilon_{l_1 R} + (1 - \omega) \varepsilon_{l_2 R}$ is negative. If one interprets second-period labor supply as the retirement decision, the evidence in Gruber and Wise (1999) supports the (unambiguous) result that retirement falls substantially if individuals accumulate larger pension wealth ($\varepsilon_{l_2 R} < 0$).

To gain intuition for the optimal tax structure, we will first discuss two special cases before turning to the interpretation of the complete model. In the first case, we assume that first-period labor supply is exogenous and that there is no first-period labor income risk. We label this the 'working-for-retirement' model, as we could interpret second-period labor supply as the (intensive) retirement decision. This structure of the model corresponds to the setting analyzed in Cremer and Gahvari (1995a,b), and is similar to Diamond and Mirrlees (1978, 1986). This case corresponds to Cremer and Gahvari (1995a,b) because of the particular sequencing of household choices.[14] In particular, the tax base of the pre-committed good in Cremer and Gahvari (1995a,b) is deterministic. In our intertemporal setting, the pre-committed good corresponds to first-period consumption. As a result, the saving base is deterministic given that first-period labor income is exogenous. In this case, subsidizing first-period consumption is equivalent to taxing savings.

In the second case, we assume that the second-period labor supply is exogenous and there is no second-period labor-income risk. This model is denoted as the 'saving-for-retirement' model and extends the deterministic analyses in Ordover and Phelps (1979) and Atkinson and Sandmo (1980) to allow for income risk. Although first-period consumption is still deterministic, savings are not, since the earnings shock occurs after first-period consumption and leisure choices have been made. Consequently, in contrast to the working-for-retirement model (and Cremer and Gahvari, 1995a,b) the savings base is now stochastic. Therefore, subsidizing first-period consumption is no longer equivalent to taxing savings. Our particular sequencing of the risk realization and household decisions also distinguishes our paper from the new dynamic public finance literature, where savings in each period are determined after the realization of risk in that period, and not

before. See, for example, Golosov et al. (2003), Kocherlakota (2005), Albanesi and Sleet (2006), and Golosov and Tsyvinski (2006).

4. Working-for-retirement: exogenous first-period leisure

In case first-period labor supply is exogenous (but can in principle be positive, i.e., $\omega > 0$), and if there is no income risk in the first period, we have $\varepsilon_{l_1 t} = \varepsilon_{l_1 R} = \xi_1 = 0$. Labor supply can in this case also be interpreted as the retirement decision. We find from Eqs. (35) and (36) the following first-order conditions for the optimal labor and capital-income tax:

$$(1 - \omega)\xi_2 = -\left(\frac{t}{1-t}\right)(1 - \omega)\varepsilon_{l_2 t} + \left(\frac{\tau r/R}{1-t}\right)\gamma \varepsilon_{c_1 t}, \tag{37}$$

$$0 = -\left(\frac{t}{1-t}\right)(1 - \omega)\varepsilon_{l_2 R} + \left(\frac{\tau r/R}{1-t}\right)\gamma \varepsilon_{c_1 R}. \tag{38}$$

Expression (37) demonstrates that the labor tax is set in such a way that the marginal benefits in terms of larger social insurance $(1 - \omega)\xi_2$ are equated to the net marginal dead weight costs of doing so. The net costs consist of two effects. First, a higher labor tax distorts labor supply more heavily as indicated by $-\frac{t}{1-t}(1 - \omega)\varepsilon_{l_2 t} > 0$. Second, provided that capital income is taxed, and households thus tend to consume too much in the first period, a higher labor tax reduces these intertemporal distortions in consumption, as can be seen from $\frac{\tau r/R}{(1-t)}\gamma \varepsilon_{c_1 t} < 0$.

The intuition for Eq. (38) is simpler. Taxes on savings are used for efficiency reasons only, since the capital-tax base is deterministic. Therefore, capital taxes do not reduce the variance in risky labor earnings. Thus, capital-income taxes have no insurance benefits, and the insurance characteristic ξ_2 is absent in Eq. (38). The only role of the tax on saving is to mitigate the distortions on labor supply. The first term on the right-hand side gives the benefits of smaller labor supply distortions ($-\frac{t}{1-t}(1 - \omega)\varepsilon_{l_2 R} > 0$). A larger capital tax boosts second-period labor supply, since a capital tax generates a wealth effect on second-period labor supply due to intertemporal substitution effects in consumption. Note that there is no direct intertemporal substitution in leisure demand with leisure being chosen in one period only. The second term represents the costs of a saving tax in terms of a distorted pattern of consumption over the life cycle $\left(\frac{\tau r/R}{1-t}\gamma \varepsilon_{c_1 R} < 0\right)$.

From the last equation follows the optimal dual tax structure (hats denote the optimized values)[15]:

$$\frac{\hat{\tau} r}{R} = \frac{(1 - \omega)}{\gamma} \frac{\varepsilon_{l_2 R}}{\varepsilon_{c_1 R}} \hat{t} > 0. \tag{39}$$

Eq. (39) demonstrates that a dual income tax with both positive taxes on capital income and labor income is optimal as long as the labor tax is used for insurance ($t > 0$). Below we will show that this is indeed the case. By boosting labor supply the capital tax alleviates the labor-tax distortions associated with insuring labor income risks. Savings and second-period labor supply are substitutes. Therefore, taxing savings helps to reduce distortions in labor supply. The stronger the complementarity between first-period consumption and second-period labor supply, the larger is $\varepsilon_{l_2 R}$, and the higher should be the capital tax. If

---

[14] Eqs. (18a) and (18b) in Cremer and Gahvari (1995a) correspond to our Eqs. (37) and (38).

[15] We borrowed the 'dual-income tax' terminology from the Nordic countries, since we want to emphasize the separate taxation of capital and labor incomes. However, in the Nordic countries the dual-income tax is usually referred to as a flat-rate capital-income tax combined with a progressive-rate labor-income tax, where the tax rate in the lowest tax bracket of the labor-income tax corresponds to the capital-income tax rate.

the distortions in saving are larger, $\varepsilon_{c_1 R}$ increases, and optimal capital-income taxes should be set at lower levels. If more consumption is allocated toward the second-period of the life cycle, γ is smaller and capital taxes are less distortionary. Hence, optimal capital taxes can be higher. Similarly, if relatively more labor income is earned in the second period, $(1-\omega)$ is larger and the larger are the efficiency gains of taxes on capital income. Optimal capital taxes would only be zero when savings and labor supply would not be substitutes ($\varepsilon_{l_2 R}=0$), capital-income taxes would be infinitely distortionary ($\varepsilon_{c_1 R}=\infty$), or second-period labor income would be zero ($\omega=1$). None of these conditions is fulfilled with standard preferences.

By using the optimal dual income tax we can obtain the following expression for the optimal labor tax at the optimal capital tax:

$$\frac{\hat{t}}{1-\hat{t}} = \frac{\xi_2}{-\varepsilon_{l_2 t} + \varepsilon_{c_1 t}\dfrac{\varepsilon_{l_2 R}}{\varepsilon_{c_1 R}}} > 0. \tag{40}$$

The expression for the optimal labor tax illuminates the trade-off between insurance (numerator) and incentives (denominator). The optimal labor tax increases with the insurance characteristic of labor income. The more risky is second-period labor income, the larger is ξ_2, and the larger are the social gains from insurance. The optimal labor tax decreases with the compensated tax elasticity of labor supply. The higher is the elasticity $\varepsilon_{l_2 t} < 0$ in absolute value, the more labor supply responds to taxation, and the lower should be the optimal labor tax rate. From the denominator in the expression for social insurance follows that capital taxes allow for more social insurance – ceteris paribus ξ_2 – if labor income is a stronger substitute for savings, i.e., when $\dfrac{\varepsilon_{l_2 R}}{\varepsilon_{c_1 R}} > 0$ is larger. By taxing capital income, the government reduces labor-tax distortions in social insurance, and optimal labor taxes can be set higher accordingly. When the government would not be interested in providing social insurance ($\xi_2=0$) both the labor and capital tax would be zero.[16]

The capital tax is optimally employed irrespective of the preference structure of the households. In particular, the elasticities are not zero even when preferences are separable and sub-utility over consumption is homothetic, cf. the elasticities in Table 1. These are the standard conditions to obtain zero optimal capital-income taxes (no commodity tax differentiation) in deterministic models with linear instruments (cf. Sandmo, 1974; Atkinson and Stiglitz, 1976; Deaton, 1979; Atkinson and Sandmo, 1980). Hence, the Atkinson–Stiglitz no commodity tax-differentiation result breaks down under risk, as has been demonstrated before by Cremer and Gahvari (1995a).

Our analysis replicates the findings in Cremer and Gahvari (1995a), but sheds a different light on their explanation. This also affects the interpretation of optimal non-linear policies in Cremer and Gahvari (1995b). Cremer and Gahvari (1995a) cast their model in terms of optimal commodity taxes rather than labor income and capital-income taxes. They argue that commodity taxes should optimally be differentiated. In particular, the tax on the 'pre-committed' commodity (c_1) should be lower than that on the 'post-committed' commodity (c_2). This finding corresponds to our result of the desirability of capital-income taxes.

Cremer and Gahvari (1995a,b) argue that commodity tax differentiation is optimal to reduce 'underconsumption' and 'overconsumption' of pre- and post-committed goods — relative to a first-best rule with perfect insurance markets. In our reading of Cremer and Gahvari (1995a,b), it is implicitly assumed that under- and overconsumption

cause externalities that need to be internalized by adjusting tax policy. In our setting, the argument would then be that there is (precautionary) oversaving. Following the interpretation of Cremer and Gahvari (1995a,b), the government would thus like to correct this oversaving by levying a tax on saving.

We think that this explanation needs to be revised. The chosen terminology 'overconsumption' (and 'underconsumption,' respectively) only refers to a first-best situation. However, in the second-best setting that both Cremer and Gahvari (1995a,b) and we analyze, this under- or overconsumption needs no corrective government action. The reason is that individuals optimally reduce their risk exposure through self-insurance in the form of precautionary saving. Government intervention to reduce or increase saving only upsets the optimal private exposure to labor-market risk, since taxes on saving themselves do not reduce income risk as the tax base is deterministic. Levying a saving tax (and rebating the revenue in the form of transfers) would therefore not reduce the exposure of households to income risk, while at the same time it would create (larger) distortions in the saving decision. Consequently, such a policy cannot be welfare improving. The reason why commodity tax differentiation is optimal is that such a policy alleviates labor-supply distortions in social insurance caused by the labor tax. Hence, it allows for more social insurance in Cremer and Gahvari (1995a,b) and in our model. Indeed, Lemma 1 in Cremer and Gahvari (1995a) implies complementarity between (second-period) labor supply and first-period consumption, like in our model.

Finally, if one interprets labor supply as the retirement decision, our results indicate that (retirement) savings should optimally be taxed as long as the labor tax directly distorts the retirement decision. Consequently, in an optimal social insurance scheme it is not desirable to have actuarially neutral pension saving schemes, i.e., a zero net tax on pension saving. Moreover, if the aim is to raise the effective retirement age, this could be indirectly achieved by increasing the tax burden on (pension) savings.

We summarize the findings of this section in the next proposition.

Proposition 1. Exogenous first-period leisure

The optimal capital tax is positive and the Atkinson–Stiglitz theorem breaks down. The capital tax is not used for social insurance, but only to off-set distortions on second-period labor supply. The optimal capital tax increases if the capital tax boosts second-period labor supply more and if it features fewer distortions in saving.

5. Saving-for-retirement: exogenous second-period leisure

Our second special case is concerned with exogenous and non-stochastic second-period labor income: $\varepsilon_{l_2 t}=\varepsilon_{l_2 R}=\xi_2=0$. In this case, we assume that households choose leisure only when young, and save in order to finance their retirement consumption. Still, they can have some exogenous labor supply when old (i.e., $\omega<1$). From Eqs. (35) and (36), the expressions for the optimal taxes on labor income and saving are in this case given by

$$\omega\xi_1 = -\left(\frac{t+\tau r/R}{1-t}\right)\omega\varepsilon_{l_1 t} + \left(\frac{\tau r/R}{1-t}\right)\gamma\varepsilon_{c_1 t}, \tag{41}$$

$$\omega\xi_1 = \left(\frac{t+\tau r/R}{1-t}\right)\omega\varepsilon_{l_1 R} - \left(\frac{\tau r/R}{1-t}\right)\gamma\varepsilon_{c_1 R}. \tag{42}$$

Eq. (41) is the optimum condition of the labor tax where the effective marginal insurance benefits ($\omega\xi_1>0$), are equated with the marginal efficiency costs of the labor tax. The net marginal costs of employing a larger labor tax consist of two elements. First, increasing the labor tax results in larger labor market distortions that are

---

[16] If capital taxation was not available, the optimal labor tax in Eq. (40) would collapse to the standard trade-off between insurance and direct distortions in labor supply, as sketched in Eaton and Rosen (1980).

represented by $-\left(\frac{t+\tau r/R}{1-t}\right)\omega\varepsilon_{l_1 t}>0$. Second, intertemporal distortions will be smaller when the labor tax increases as indicated by $\left(\frac{\tau r/R}{1-t}\right)\gamma\varepsilon_{c_1 t}<0$. Intuitively, the labor tax reduces first-period consumption demand, and this alleviates the distortions on first-period consumption from a positive capital-income tax. When the capital-income tax is zero, only the labor tax determines the distortions in labor supply.

Eq. (42) is the optimum condition for the capital-income tax. The marginal insurance benefits ($\omega\xi_1>0$) are equal to the marginal efficiency costs of the capital-income tax. In contrast to the previous case (see Section 4), the capital-income tax now features insurance benefits, since savings are stochastic. Indeed, a larger variance in first-period income shocks gives a larger variance in savings, since individuals with lower first-period labor income save less. The costs of employing the capital tax for social insurance are two-fold. First, a larger capital-income tax entails larger intertemporal distortions in consumption as indicated by $-\left(\frac{\tau r/R}{1-t}\right)\gamma\varepsilon_{c_1 R}>0$. This term was also present before. Second, a larger capital-income tax exacerbates the labor tax distortions by acting as an implicit tax on labor supply as can be seen from $\left(\frac{t+\tau r/R}{1-t}\right)\omega\varepsilon_{l_1 R}>0$. Intuitively, the capital tax reduces first-period labor supply, since intertemporal substitution in consumption provokes a wealth effect on leisure demand in period one. The capital tax also affects labor supply via the tax wedge on labor. In particular, the capital tax changes the relative price of the first-period labor supply in terms of second-period consumption. The capital tax did not feature in the tax wedge on labor in the previous model, because the capital tax does not affect the relative price of the second-period labor supply in terms of second-period consumption. Again, there is no direct intertemporal substitution in leisure, since individuals consume leisure only in the first-period.

Compared to the previous model, the cross-elasticity of labor supply with respect to the net interest rate has switched in sign. A larger capital tax lowers the net return on saving and raises first-period consumption relative to second-period consumption. As a result, individuals would like to substitute first-period consumption for first-period leisure and first-period labor supply falls. Consequently, saving and first-period labor supply are complements. Capital-income taxes therefore do not reduce labor market distortions, but exacerbate them. Indeed, reducing labor market distortions ceteris paribus requires subsidies on capital income rather than taxes.

The insurance characteristic is identical in the expressions for both the labor and the capital tax. Hence, insuring income through either labor or capital taxes provides the same distributional benefits. The reason is that the marginal propensity to save out of first-period labor income is equal to one, given that the first-period consumption and labor supply choices are committed before the earnings shock is realized. Consequently, a tax on saving is equivalent to a tax on labor income in terms of reducing the variance in earnings. Thus, whether labor income should be taxed at a higher rate than capital income depends only on whether the marginal costs of employing labor taxes are lower than the marginal costs of employing capital-income taxes. Therefore, an optimal policy equalizes the marginal excess burdens of labor and capital taxes.

We obtain the optimal Ramsey rule for the dual income tax structure by subtracting Eqs. (41) and (42) to find

$$\left(\frac{\hat{\tau} r/\hat{R}}{1-\hat{t}}\right)\gamma\left(\varepsilon_{c_1 t}+\varepsilon_{c_1 R}\right)=\left(\frac{\hat{t}+\hat{\tau} r/\hat{R}}{1-\hat{t}}\right)\omega\left(\varepsilon_{l_1 t}+\varepsilon_{l_1 R}\right). \tag{43}$$

Our Ramsey rule is intuitively the same as the optimal dual income tax in deterministic Ramsey models with saving for retirement (see, e.g., Atkinson and Sandmo, 1980, Eq. (32)), but now in case of

providing optimal social insurance, rather than raising an exogenous amount of tax revenue with distorting tax instruments.[17]

The left-hand side represents the total marginal welfare costs of distorting savings for income insurance. The welfare costs of distorting savings increase with the tax wedge on capital income $\frac{\hat{\tau} r/\hat{R}}{1-\hat{t}}$, and the total elasticity of first-period consumption $\gamma(\varepsilon_{c_1 t}+\varepsilon_{c_1 R})<0$, which measures the behavioral response of the savings base with respect to both tax instruments combined. Both elasticities are negative. A higher capital tax distorts saving by boosting first-period consumption. Additionally, a higher labor tax counters the saving distortion by reducing first-period consumption. $\varepsilon_{c_1 t}+\varepsilon_{c_1 R}$ gives the combined effect of a lower capital tax while simultaneously increasing the labor tax so as to keep income insurance constant. Thus, simultaneously raising the labor tax and lowering the capital tax (higher R), result in a lower distortion on consumption.

Similarly, the right-hand side gives the total marginal welfare cost of distorting labor supply. The cost of distorting labor supply increases with the net tax wedge on labor supply $\frac{\hat{t}+\hat{\tau} r/\hat{R}}{1-\hat{t}}$, and the total elasticity of the labor tax base $\omega(\varepsilon_{l_1 t}+\varepsilon_{l_1 R})$ with respect to the two policy instruments. At first sight, the tax-base elasticity appears ambiguous. On the one hand, an increase in labor taxation will decrease labor supply: $\varepsilon_{l_1 t}<0$. On the other hand, an increase in the net interest rate boosts labor supply: $\varepsilon_{l_1 R}>0$. By substituting the elasticities (see Table 1), we find that the net effect is always negative: $\varepsilon_{l_1 t}+\varepsilon_{l_1 R}=-\frac{\varepsilon_1\Sigma_1}{\Delta}\frac{\gamma}{1-t}<0$ (for $\varepsilon_2=0$). Hence, simultaneously raising the labor tax and lowering the capital tax (higher R), while keeping insurance constant, result in a larger distortion on labor supply.

Accordingly, both tax wedges in Eq. (43) have the same sign at the optimum. Distortions in first-period labor supply by a non-zero total tax wedge on labor supply should be equal to the distortions in saving by a non-zero tax wedge on saving. Therefore, capital income is optimally taxed (subsidized) at a positive rate $\hat{\tau} r/\hat{R}>0$ (<0) if labor income is taxed (subsidized) on a net basis, i.e., if $\frac{\hat{t}+\hat{\tau} r/\hat{R}}{1-\hat{t}}>0$ (<0). Below we demonstrate that the net tax on labor is always positive so that capital income should always be taxed. Intuitively, starting from a situation without taxes on capital income, introducing a small tax on capital income, while lowering the labor tax at the same time, would produce no change in insurance benefits, since both instruments have identical insurance gains. Also, starting from a zero capital tax, the introduction of a small capital tax would only generate second-order intertemporal distortions in consumption. However, it would allow for a first-order reduction in labor-supply distortions by lowering the labor tax. Thus, taxing capital income helps to achieve the same insurance at lower efficiency costs.

From this discussion follows that the intuition from the working-for-retirement model does not apply to the saving-for-retirement model. The government may like to provide a subsidy on saving, rather than a tax on saving, so as to reduce labor-supply distortions. However, a saving subsidy would raise the exposure of households to income risk, since savings are stochastic. The increase in labor taxes needed to maintain the same level of social insurance increases labor supply distortions so much that this would more than off-set the positive effect of the capital subsidy on labor supply. Hence, it is never optimal to subsidize savings in the saving-for-retirement model.

The Atkinson–Stiglitz uniform commodity-tax result can again never be obtained as long as standard utility functions are adopted. In particular, zero taxation of capital income would require either that savings are infinitely elastic, or, (first-period) labor supply is completely inelastic. In these knife-edge cases the capital tax is either infinitely distortionary or the labor tax is completely non-distortionary. Consequently, in stark contrast to the deterministic Ramsey models, positive taxation of capital income is unambiguously part of the optimal tax policy under income risk.

By substituting Eq. (43) into the reduced first order conditions (41) and (42), and rearranging and collecting terms, we find that the total net tax on labor (i.e., the direct tax on labor plus the implicit tax on labor due to the capital tax) satisfies

$$\frac{\hat{t} + \hat{\tau} r/\hat{R}}{1 - \hat{t}} = \frac{\xi_1 + \varepsilon_{c_1 t} \frac{\xi_1}{\varepsilon_{c_1 R}}}{-\varepsilon_{l_1 t} + \varepsilon_{c_1 t} \frac{\varepsilon_{l_1 R}}{\varepsilon_{c_1 R}}} > 0. \tag{44}$$

The optimal net tax on labor is positive by substituting the elasticities from Table 1. Eq. (44) gives the standard trade-off between social insurance (numerator) and distortions (denominator) and proves that the capital-income tax is optimally positive, cf. Eq. (43).

The denominator represents the net distortions of taxing labor income, which decrease the optimal tax wedge on labor income. In particular, distortions of social insurance increase with the tax elasticity of labor supply $-\varepsilon_{l_1 t} > 0$. Like before, the second term in the denominator, $\varepsilon_{c_1 t} \frac{\varepsilon_{l_1 R}}{\varepsilon_{c_1 R}} > 0$, captures the interaction between labor supply and saving. The stronger the substitutability between first-period consumption and first-period labor supply, the larger (in absolute value) is $\frac{\varepsilon_{l_1 R}}{\varepsilon_{c_1 R}} < 0$. Thus, if capital taxes are higher, labor taxes should be lower as they exacerbate the distortions of the capital tax on labor supply. The interaction term is smaller if the cross-elasticity of consumption with respect to the labor tax ($\varepsilon_{c_1 t} < 0$) is smaller (in absolute value). In that case, a higher labor tax does not exacerbate labor supply distortions a lot.

The term in the numerator contains the standard, direct insurance gain of labor taxes $\xi_1 > 0$. In addition, there is also an indirect insurance gain of labor taxes, since $\varepsilon_{c_1 t} \frac{\xi_1}{\varepsilon_{c_1 R}} > 0$. Intuitively, the labor tax reduces the first-period consumption $\varepsilon_{c_1 t} < 0$, and thereby reduces the distortions of the capital tax on consumption choices. As a result, the trade-off between insurance and distortions of employing capital taxes improves, as indicated by the term $\frac{\xi_1}{\varepsilon_{c_1 R}} < 0$. Therefore, the optimal wedge on labor should increase as well.

By using the optimal tax wedge on labor (Eq. (44)) in the optimal dual tax structure in Eq. (43), we obtain the optimal capital tax rate

$$\frac{\hat{\tau} r/\hat{R}}{1 - \hat{t}} = \frac{\omega}{\gamma} \left(\frac{\xi_1 + \varepsilon_{l_1 R} \frac{\xi_1}{\varepsilon_{l_1 t}}}{-\varepsilon_{c_1 R} + \varepsilon_{l_1 R} \frac{\varepsilon_{c_1 t}}{\varepsilon_{l_1 t}}} \right) > 0. \tag{45}$$

Upon substitution of the relevant elasticities from Table 1, we can derive that the optimal capital tax is indeed unambiguously positive and increases with the desire to insure risk in first-period income ξ_1. As before, a lower γ implies that more consumption is allocated toward the second period of the life cycle, so that capital taxes are less distortionary. Hence, optimal capital taxes can be higher. Furthermore, the total insurance effect is increasing in risky labor income. Consequently, a higher share ω of the first-period labor income in total labor income calls for a higher capital tax rate.

The denominator in brackets represents the welfare cost of the capital tax. Welfare losses of capital-income taxes increase in the

elasticity of consumption with respect to the interest rate ($-\varepsilon_{c,R} > 0$). Capital taxes exacerbate the distortions of the labor tax on labor supply, so that the efficiency losses in saving increase further, cf. $\varepsilon_{l_1 t} \frac{\varepsilon_{c_1 t}}{\varepsilon_{l_1 t}} > 0$.

The first term in the numerator, ξ_1, designates the direct insurance gain of capital taxes, whereas $\varepsilon_{l_1 R} \frac{\xi_1}{\varepsilon_{l_1 t}} < 0$ represents again the indirect insurance effect. Capital taxes should be higher if this provides a lot of distributional benefits. However, by lowering the first-period labor supply ($\varepsilon_{l,R} > 0$), capital-income taxes worsen the insurance-incentives trade-off of the labor tax, which is captured by $\frac{\xi_1}{\varepsilon_{l_1 t}} < 0$. As a result, capital-income taxes reduce the attractiveness of using labor-income taxes to insure income risks, and should be lowered accordingly.

We can eliminate $\frac{\tau r}{R}$ from the optimal wedge on labor supply in Eq. (44) to find the optimal labor tax

$$\frac{\hat{t}}{1 - \hat{t}} = \left(\frac{1 + \frac{\varepsilon_{c_1 t}}{\varepsilon_{c_1 R}} - \frac{\frac{\omega}{\gamma} \left(\varepsilon_{l_1 t} + \varepsilon_{l_1 R} \right)}{\varepsilon_{c_1 R}}}{-\varepsilon_{l_1 t} + \varepsilon_{c_1 t} \frac{\varepsilon_{l_1 R}}{\varepsilon_{c_1 R}}} \right) \xi_1. \tag{46}$$

The labor tax is generally ambiguous in sign. The reason is that the capital tax is part of the labor wedge. If the optimal capital tax becomes larger, a negative labor tax $t < 0$ might be necessary in order maintain the optimal net tax wedge on labor ($\frac{t + \tau r}{1 - t} > 0$). The condition for optimally positive labor taxes is $\omega(\varepsilon_{c_1 t} + \varepsilon_{l_1 R}) > \gamma(\varepsilon_{c_1 t} + \varepsilon_{c_1 R})$. This condition will be fulfilled if the first-period labor supply is sufficiently inelastic (ε_1):

$$1 - \frac{\gamma}{\omega} \left(\frac{\varepsilon_{c_1 t} + \varepsilon_{c_1 R}}{\varepsilon_{l_1 t} + \varepsilon_{l_1 R}} \right) = 1 - (1 - t)(1 - \pi_1) - \frac{1 - \gamma}{\rho_2 \Sigma_1 \varepsilon_1 \omega} < 0, \tag{47}$$

We assume that this condition holds and that the labor tax is optimally positive. Note, however, that the sign of the capital tax does not depend on this assumption, so the result that capital income should optimally be taxed remains unchanged.

To conclude, subsidies on saving could boost labor supply of the young workers in the saving-for-retirement model. However, this is not an optimal policy. A negative capital tax raises the exposure to labor income risk. Hence, a rise in the labor tax is needed to maintain the same level of insurance. Intuitively, keeping the level of income insurance constant implies that the labor tax needs to increase as the capital tax is lowered. However, a negative capital tax combined with a higher labor tax so as to keep the level of social insurance constant generates larger distortions. The reason is that the rise in the labor tax more than off-sets the positive impact of the saving subsidy on labor supply. Consequently, capital income (i.e., retirement income in this context) should not be subsidized, but taxed so as to provide social insurance at the lowest social cost. Therefore, these results suggest that policies to subsidize retirement plans are questionable, because the distortions associated with a rise in the tax burden to finance the tax subsidies outweigh their beneficial effects on labor supply.

We summarize the results of this section in the following proposition.

Proposition 2. Exogenous second-period leisure

The optimal capital tax is positive and the Atkinson–Stiglitz theorem breaks down. The capital-income tax is equally effective as the labor tax in providing social insurance. The optimal capital tax increases with the welfare gains of social insurance, but decreases with distortions in saving and first-period labor supply.

6. General model

In the general model, in which labor supply in both periods is endogenous, not only the two mechanisms highlighted in the sub-models are

present (intertemporal wealth effects in labor supply and insurance of first-period labor-income risk), but a third factor will determine the desirability of capital-income taxes: intertemporal substitution effects in labor supply. This section discusses the model in full.

We obtain the following expression describing the optimal labor tax from rearranging Eq. (35):

$$\omega\xi_1 + (1-\omega)\xi_2 = -\left(\frac{t+\tau r/R}{1-t}\right)\omega\varepsilon_{l_1 t} - \left(\frac{t}{1-t}\right)(1-\omega)\varepsilon_{l_2 t} \qquad (48)$$
$$+ \left(\frac{\tau r/R}{1-t}\right)\gamma\varepsilon_{c_1 t}$$

The expression for the optimal labor tax equates the insurance gains of reducing risk in first- and second period incomes, $\omega\xi_1 + (1-\omega)\xi_2$, to the net marginal cost of doing so. The welfare costs of labor taxes are represented by three terms. The first two terms give the marginal excess burdens of labor taxes on first- and second period labor supply, respectively. Note that $-\omega\varepsilon_{l_1 t} > 0$, and $-(1-\omega)\varepsilon_{l_2 t} > 0$. The last term gives the reduction in the excess burden of a positive capital tax, since the labor tax partially off-sets the saving distortion by discouraging first-period consumption ($\gamma\varepsilon_{c1t} < 0$).

The optimal capital tax follows from rearranging Eq. (36):

$$\omega\xi_1 = \left(\frac{t+\tau r/R}{1-t}\right)\omega\varepsilon_{l_1 R} + \left(\frac{t}{1-t}\right)(1-\omega)\varepsilon_{l_2 R} - \left(\frac{\tau r/R}{1-t}\right)\gamma\varepsilon_{c_1 R}. \qquad (49)$$

In contrast to the labor-income tax, the capital tax can only be employed for insurance to reduce the risk of first-period incomes ($\omega\xi_1$), not second-period incomes ($(1-\omega)\xi_2$). The reason is that the second-period income shock occurs after savings have been made. Hence, taxing savings does not help to reduce the variance of incomes in the second period of the life cycle. The marginal insurance gains $\omega\xi_1$ should again be equal to the net marginal dead weight loss associated with more income insurance. In particular, a capital tax causes the standard saving distortion which is represented by $-\frac{\tau r/R}{1-t}\gamma\varepsilon_{c_1 R} > 0$. Moreover, the capital tax exacerbates the labor tax distortions on first-period labor supply, since $\omega\varepsilon_{l_1 R} > 0$. This is, first, due to wealth effects arising from intertemporal substitution in consumption. Second, in the general model with endogenous labor supply in both periods, capital taxes also generate direct intertemporal substitution effects on leisure demands so that first-period labor supply falls. Finally, the capital tax reduces distortions in the second-period labor supply, because $\frac{t}{1-t}(1-\omega)\varepsilon_{l_2 R} < 0$ for positive labor taxes. Wealth effects due to intertemporal substitution in consumption and intertemporal substitution in leisure both raise second-period labor supply.

By combining both equations, we obtain the optimal dual tax structure:

$$\left(\frac{\hat{\tau} r/R}{1-\hat{t}}\right)\gamma\left(\varepsilon_{c_1 t} + \varepsilon_{c_1 R}\right) = \left(\frac{\hat{t}+\hat{\tau} r/R}{1-\hat{t}}\right)\omega\left(\varepsilon_{l_1 t} + \varepsilon_{l_1 R}\right) \qquad (50)$$
$$+ \left(\frac{\hat{t}}{1-\hat{t}}\right)(1-\omega)\left(\varepsilon_{l_2 t} + \varepsilon_{l_2 R}\right) + (1-\omega)\xi_2.$$

The optimal capital tax is determined by four elements. The two elements in the first line are identical to the expression for the optimal capital tax of the previous section, see Eq. (43).

First, the optimal capital-income tax $\frac{\hat{\tau} r/R}{1-\hat{t}}$ is larger if the first-period consumption has a lower total elasticity with respect to the policy instruments and the income share of consumption today is lower (lower γ), so that $\gamma(\varepsilon_{c_1 t} + \varepsilon_{c_1 R}) < 0$ is lower in absolute value. Naturally, the capital tax distorts intertemporal consumption choices ($\varepsilon_{c_1 R} < 0$). However, the labor tax reduces the capital-tax distortions by reducing consumption in the first period ($\varepsilon_{c_1 t} < 0$) – provided

that the capital tax is positive. The net effect is negative, see also the previous section.

Second, the optimal capital tax increases if first-period labor is more heavily distorted, i.e., when $\left(\frac{\hat{t}+\hat{\tau}R}{1-\hat{t}}\right)\omega(\varepsilon_{l_1 t} + \varepsilon_{l_1 R})$ is higher in absolute value. The distortion is larger if individuals earn a relatively large fraction of their life-time income ω in the first period. The intuition for this term is identical to the model with only endogenous first-period labor supply. In particular, a capital subsidy could be employed to reduce the labor-tax distortion. However, the rise in labor taxes to maintain the same level of income insurance could more than off-set the positive effects of the capital subsidy on labor supply. In contrast to the previous section the net effect is no longer unambiguous, since $\varepsilon_{l_1 t} + \varepsilon_{l_1 R} = -\frac{\varepsilon_1 \Sigma_1}{\Delta}\left(\frac{\gamma}{1-t} - \rho_1(1-\pi_2)(1-\omega)\varepsilon_2\right) \gtrless 0$. Intuitively, in the current model with endogenous leisure in both periods, intertemporal substitution effects in the pattern of leisure demand over time provide an additional channel whereby capital-income taxes affect labor supply, besides the wealth effects generated by intertemporal substitution in consumption. In particular, a larger capital-income tax renders current leisure more attractive than future leisure. As a result, the capital tax raises the distortion on first-period labor supply even further, thereby reducing the desire to tax capital incomes. This intertemporal substitution effect in leisure is absent in the models with only one leisure demand decision.

Third, $\frac{t}{1-\hat{t}}(1-\omega)\left(\varepsilon_{l_2 t} + \varepsilon_{l_2 R}\right) < 0$ indicates the role of capital taxes to reduce the tax distortion on second-period labor supply. The combined elasticity is unambiguously signed: $\varepsilon_{l_2 t} + \varepsilon_{l_2 R} < 0$. A larger capital tax allows for a lower labor tax, so that labor tax distortions on second-period labor supply diminish. In addition, a capital tax boosts second-period labor supply through intertemporal substitution effects so that it alleviates the distortions of the labor tax on second-period labor supply even more. Accordingly, a positive capital tax ceteris paribus allows for more social insurance by reducing the distortions in second-period labor supply.

Fourth, the capital tax increases if labor taxes are less efficient in social insurance, thus, if $(1-\omega)\xi_2$ is lower, i.e., if second-period risk is relatively less important compared to first-period income risk (note that the previous three terms discussed so far are negative). Indeed, in the absence of second-period labor income risk (ξ_2), capital income is generally taxed at positive rates if intertemporal substitution of leisure is modest, and if labor supplies in both periods are taxed at net positive rates. Consequently, capital-income taxes alleviate labor-supply distortions in social insurance. However, if second-period labor income is substantially more risky than first-period income, ξ_2 is larger, and capital-income taxes loose their attractiveness as an insurance device. Therefore, capital-income taxes tend to be set lower.

We derive an explicit condition under which capital income should be taxed at positive rates even if capital taxes do not provide any insurance at all, i.e., if the saving base is deterministic. Capital-income taxes are then employed for efficiency reasons only. In that case, we can set $\xi_1 = 0$ in the expression for the optimal capital tax (Eq. (49)) to find

$$\frac{\hat{\tau} r}{\hat{R}} = \frac{-\bar{\varepsilon}_{lR}}{\varepsilon_{aR}}\hat{t}, \qquad (51)$$

where $\bar{\varepsilon}_{lR} \equiv \omega\varepsilon_{l_1 R} + (1-\omega)\varepsilon_{l_2 R}$ denotes the income-weighted average elasticity of total labor supply with respect to the interest factor, and $\varepsilon_{aR} \equiv \omega\varepsilon_{l_1 R} - \gamma\varepsilon_{c_1 R} > 0$ denotes the compensated interest rate elasticity of savings. ε_{aR} is unambiguously positive, because a higher net interest rate increases first-period labor supply ($\varepsilon_{l_1 R} > 0$) and it decreases the first-period consumption ($\varepsilon_{c_1 R} < 0$). Therefore, capital income is taxed if labor income is taxed ($t > 0$), and if the positive effect of capital-income taxes on second-period labor supply ($(1-\omega)\varepsilon_{l_2 R} < 0$) is not off-set by

the negative effect of capital-income taxes on first-period labor supply ($\omega\varepsilon_{l_1R} > 0$).

The net effect thus depends on the intertemporal substitution pattern in labor supply and the relative shares of labor earned in the first- and the second-period of the life cycle (ω). Theoretically, the sign of the capital tax is ambiguous. Using cross-sectional data, Pirttilä and Suoniemi (2010) and Gordon and Kopczuk (2011) estimate that labor supply falls if households have higher capital incomes. As long as populations are stationary, cross-sectional estimates are useful to sign the impact of capital income on life-time labor supply. These findings suggest that $\bar{\varepsilon}_{lR} < 0$, since average labor supply falls with higher capital incomes.[18] Additional evidence is provided by realistically calibrated life-cycle models in Erosa and Gervais (2002) and Conesa et al. (2009). These authors find that optimal capital taxes are generally positive for efficiency reasons only. In particular, the capital tax increases labor supply at later stages of the life cycle more than it decreases labor supply at earlier stages. This evidence also suggests that $\bar{\varepsilon}_{lR} < 0$.[19] Accordingly, capital income should optimally be taxed, even if capital-income taxes do not provide any insurance gains.

Returning to the general case where capital taxes also feature insurance gains, we find the optimal capital tax rate from solving Eq. (49) for the labor tax $\frac{t}{1-t}$ and substituting the resulting expression into Eq. (48). Naturally, this also gives the result that the optimal capital tax should remain positive as long as we maintain the assumption that $\bar{\varepsilon}_{lR} < 0$[20,21]:

$$\frac{\hat{\tau}r/\hat{R}}{1-\hat{t}} = \frac{\omega\xi_1 + \bar{\varepsilon}_{lR}\dfrac{[\omega\xi_1 + (1-\omega)\xi_2]}{\bar{\varepsilon}_{lt}}}{\varepsilon_{aR} - \bar{\varepsilon}_{lR}\dfrac{\varepsilon_{at}}{\bar{\varepsilon}_{lt}}} > 0, \tag{52}$$

where $\bar{\varepsilon}_{lt} \equiv \omega\varepsilon_{l_1t} + (1-\omega)\varepsilon_{l_2t} < 0$ denotes the income-weighted average elasticity of total labor supply with respect to the labor tax rate. $\varepsilon_{at} \equiv \omega\varepsilon_{l_1t} - \gamma\varepsilon_{c_1t}$ is the elasticity of saving with respect to the labor tax. If $\varepsilon_{at} > 0$ ($\varepsilon_{at} < 0$), saving increases (decreases) as a result of labor taxation. The sign of ε_{at} is ambiguous since labor taxation both reduces labor supply ($\omega\varepsilon_{l_1t} < 0$) and first-period consumption ($\gamma\varepsilon_{c_1t} < 0$).

The first term in the denominator represents the direct distortions in savings, i.e., intertemporal distortions in first-period consumption and in first-period labor supply, respectively. The larger are direct intertemporal distortions on consumption and leisure, the larger is $\varepsilon_{aR} > 0$, and the lower the optimal capital tax should be. Labor taxation mitigates distortions in savings, if labor taxes boost savings ($\varepsilon_{at} > 0$), but distort labor supply ($\bar{\varepsilon}_{lt} < 0$). This trade-off is represented by $\frac{\varepsilon_{at}}{\bar{\varepsilon}_{lt}}$. If $\varepsilon_{at} > 0$ and $\bar{\varepsilon}_{lR} < 0$, capital taxation boosts labor supply, and thereby alleviates the distortions of the labor tax on labor supply. Therefore, capital taxes should be set higher. If, instead, $\varepsilon_{at} < 0$ and $\bar{\varepsilon}_{lR} < 0$, a higher capital tax exacerbates the savings distortions of the labor tax by boosting life-time labor supply. Thus, $\bar{\varepsilon}_{lR}\frac{\varepsilon_{at}}{\bar{\varepsilon}_{lt}} < 0$, and capital taxation should decrease.

The numerator of Eq. (52) captures the insurance effects of capital taxes and consists of two parts. First, there is the direct insurance effect

$\omega\xi_1$. If taxing savings reduces the exposure to first-period income risk more, capital-income taxes should be higher. This is analogous to the explanation in the saving-for-retirement case in Section 5. Additionally, the indirect insurance effect is at work. In particular, if the capital tax boosts labor supply, $\bar{\varepsilon}_{lR} < 0$, the capital tax improves the insurance-incentives trade-off of the labor tax, since $\frac{[\omega\xi_1 + (1-\omega)\xi_2]}{\bar{\varepsilon}_{lt}} < 0$. As a result the labor tax becomes a more attractive instrument for social insurance, and the capital tax should optimally increase.

To derive the optimal labor tax, we insert Eq. (52) into Eq. (49) and collect terms in order to receive

$$\frac{t}{1-t} = \frac{\omega\xi_1 + (1-\omega)\xi_2 + \varepsilon_{at}\dfrac{\omega\xi_1}{\varepsilon_{aR}}}{-\bar{\varepsilon}_{lt} + \varepsilon_{at}\dfrac{\bar{\varepsilon}_{lR}}{\varepsilon_{aR}}}. \tag{53}$$

The denominator shows that the optimal labor tax falls if providing social insurance is more distortionary. The labor tax distorts labor supply as represented by the average labor supply elasticity ($\bar{\varepsilon}_{lt} < 0$). However, the labor tax is larger if the capital tax is helpful in reducing labor market distortions by indirectly boosting labor supply ($\bar{\varepsilon}_{lR} < 0$), and if the labor tax strengthens the complementarity effect $\frac{\bar{\varepsilon}_{lR}}{\varepsilon_{aR}} < 0$ by raising savings ($\varepsilon_{at} > 0$). Instead, if the labor tax reduces overall saving ($\varepsilon_{at} < 0$), it weakens the complementarity effect of capital taxation. Consequently, distortions from labor taxation will be exacerbated, and the labor tax should be set at a lower rate. The numerator reveals that the optimal labor tax increases with the desire to insure income risk in both periods ($\omega\xi_1 + (1-\omega)\xi_2$). Finally, there is the indirect insurance effect of the tax policy. If ε_{at} is positive, the labor tax improves the insurance-incentives trade-off of the capital tax. As a result, the optimal tax on labor income needs to be higher as a result. The reverse reasoning holds if $\varepsilon_{at} < 0$.

In case the government can employ age-dependent labor-income taxes, the optimal capital tax will unambiguously be positive.[22] The reason is that the government can effectively neutralize any effect of capital-income taxes on the intertemporal allocation of leisure by appropriately adjusting the structure of labor-income taxes if age-dependent labor-income taxes are available. Consequently, capital-income taxes are only employed to reduce distortions on second-period labor supply and to insure first-period labor-income risk, as in the sub-models we discussed previously. Hence, our results would be completely robust to introducing age-dependent labor-income taxation.

We conclude this section by summarizing our results in the final proposition.

Proposition 3. Leisure endogenous in both periods

The optimal capital tax is employed for both efficiency and insurance reasons and the Atkinson–Stiglitz theorem breaks down. The capital tax is optimally positive if life-time labor supply increases with taxation of capital income. The optimal capital tax increases with its insurance gains and its effectiveness to reduce life-time labor supply distortions. The capital-income tax decreases if saving distortions are larger.

7. Conclusions

This paper has demonstrated that capital income is generally taxed in a standard two-period life-cycle model with non-insurable risks in both periods of the life cycle. The Atkinson–Stiglitz theorem

---

[18] We note that these empirical findings hold for aggregates. For some sub-groups estimates may be imprecisely estimated or switch sign. However, since our model is concerned with aggregate outcomes, we believe that we draw the correct inference from these studies.

[19] Nishiyama and Smetters (2005) develop an OLG-model calibrated on US data. They demonstrate that a positive tax on capital is desirable for social insurance as well if there is uninsurable wage risk.

[20] Second-order conditions for the optimal tax problem ensure that the denominator of the optimal tax expression is positive.

[21] Note that the term exactly simplifies to the optimal capital tax rule (Eq. (45)), if second-period labor supply is inelastic ($\varepsilon_{l_2t} = \varepsilon_{l_2R} = 0$), and if there is no risk in the second period ($\xi_2 = 0$).

[22] An unpublished appendix demonstrates this formally, and is available upon request from the authors.

of non-differentiation of commodity taxes breaks down under risk. Intuitively, capital-income taxes boost second-period labor supply by making future leisure more costly. Taxing capital income thus alleviates distortions in second-period labor supply (or retirement). However, capital-income taxes reduce first-period labor supply, but this effect is off-set because capital-income taxes insure first-period labor-income risk. Indeed, optimal social insurance requires that distortions associated with insurance should be smoothed over labor income and saving bases. Capital-income taxes also entail intertemporal distortions in leisure demands, which tend to lower optimal capital-income taxes. However, as long as the increase in second-period labor supply dominates the reduction in first-period labor supply, life-time labor supply increases (which is the empirically relevant case), and optimal capital-income taxes should unambiguously be positive.

This paper employed linear policy instruments and confirmed results from the new dynamic public finance literature, where rich sets of non-linear instruments are analyzed. By directly implementing the optimal allocations with time-invariant linear tax instruments, and without record keeping, we have demonstrated that the basic results derived in the new dynamic public finance literature are robust to (very) large deviations from the informational requirements to implement time-dependent, non-linear policies.

In contrast to the previous literature, we have also demonstrated that capital taxes have a direct role in insuring labor-market risks, especially when labor risks are important in the early stages of the life cycle. In all new dynamic public finance papers we are aware of, savings in each period are chosen only after the skill shock is known to the individual. Hence, there are no unintended, risky savings so that taxing capital income features no social-insurance gains. We think that capital-income taxes could possibly feature social-insurance gains in these models under a different sequencing of household choices and realizations of earnings risk. However, it may also be the case that indirect instruments are not desirable for social insurance at all under optimal non-linear labor-income taxation, as is the case in deterministic models analyzing optimal commodity taxes under optimal non-linear taxation, cf. Atkinson and Stiglitz (1976), Mirrlees (1976) and Jacobs and Boadway (2012). Further research is needed to explore these issues.

This paper showed that capital-income taxes are especially desirable if they help to reduce tax distortions on labor supply that arise from social insurance. We think that this result is the mirror image of the finding that intertemporal wedges are optimal in the literature employing non-linear instruments. Intertemporal wedges relax the incentive constraints associated with social insurance as individuals with favorable skill shocks are less tempted to mimic individuals with unfavorable skill shocks. Future research could investigate this conjecture in more detail.

Our findings have large policy relevance for the debate on the desirability of capital-income taxes. The introduction started with the controversy in the economics literature on the desirability of capital-income taxation. This paper bolsters the recommendations by Banks and Diamond (2010) and Diamond and Saez (2012) by showing that taxing capital income is desirable when the earnings of households are risky. Moreover, it also has relevance for policy discussions on the tax treatment of pension savings and stimulating later retirement. We show that (retirement) saving should generally be taxed, and not subsidized. Consequently, actuarially fair retirement schemes are not optimal. Governments should therefore not try to subsidize retirement saving so as to reduce future public spending on state pensions or health care. By doing this, sustainability problems in public finances worsen rather than improve. As long as governments do not wish to sacrifice on social insurance, the government needs to raise the tax burden on working-age individuals, which results in larger labor market distortions and smaller tax bases. Moreover, subsidies on (pension) saving exacerbate labor supply distortions by strengthening the incentives to retire earlier. Hence, a policy of subsidies on (retirement) savings does not help to delay retirement either.

Appendix A. Derivatives of indirect utility

Using optimal second-period consumption, the indirect utility function (Eq. (15)) can be rewritten as

$$V(T,t,R) \equiv u_1(\hat{c}_1) - v_1(\hat{l}_1) + \beta E\left[u_2\left((1-t)\theta_2\hat{l}_2 + (1+(1-\tau)r)\hat{a}\right) - v_2(\hat{l}_2)\right]$$

$$= u_1(\hat{c}_1) - v_1(\hat{l}_1)$$
$$+ \beta E\left[u_2\left((1-t)\theta_2\hat{l}_2 + R\left[(1-t)\theta_1\hat{l}_1 + T - \hat{c}_1\right]\right) - v_2(\hat{l}_2)\right], \quad (54)$$

where $R \equiv 1 + (1-\tau)r$, and $\hat{a} = (1-t)\theta_1\hat{l}_1 + T - \hat{c}_1$ from Eq. (2). By applying the envelope theorem, the derivatives of indirect utility with respect to the policy instruments are equal to (cf. Eqs. (16)–(18) in the main text):

$$\frac{\partial V}{\partial T} = \beta R E[u'_2] = \eta, \quad (55)$$

$$\frac{\partial V}{\partial t} = -\beta E[u'_2\theta_2 l_2] - \beta R E[u'_2\theta_1 l_1] = -\beta(E[u'_2]E[\theta_2 l_2] + cov[u'_2, \theta_2 l_2])$$
$$- \beta R(E[u'_2]E[\theta_1 l_1] + cov[u'_2, \theta_1 l_1])$$
$$= \beta R E[u'_2]\left(\frac{(1-\xi_2)E[\theta_2 l_2]}{R} + (1-\xi_1)E[\theta_1 l_1]\right)$$
$$= -\eta\left((1-\xi_1)l_1 + \frac{(1-\xi_2)E[\theta_2 l_2]}{R}\right), \quad (56)$$

$$\frac{\partial V}{\partial R} = -\beta(1-t)E[u'_2\theta_1 l_1] + \beta E[u'_2](T-c_1)$$
$$= -\beta(1-t)(E[u'_2]E[\theta_1 l_1] + cov[u'_2, \theta_1 l_1]) + \beta E[u'_2](T-c_1)$$
$$= \beta R E[u'_2]\frac{(1-\xi_1)(1-t)E[\theta_1]l_1 - c_1 + T}{R}$$
$$= \eta\frac{(1-\xi_2)(1-T)l_1 - c_1 + T}{R}, \quad (57)$$

where we have used Steiner's Rule for two stochastic variables X and Y, $E[XY] = E[X]E[Y] + cov[X,Y]$, the definition of the insurance characteristic $\xi_i \equiv -\frac{cov[u'_2, \theta_i l_i]}{E[u'_2]E[\theta_i l_i]} \geq 0$ from Eqs. (19) and (20), $E[\theta_1] = 1$, and the fact that first-period labor supply l_1 is deterministic.

Appendix B. Slutsky equations under risk

Under risk, compensated changes in labor supply and first-period consumption can still be derived using the Slutsky equations. These follow from differentiating Eqs. (21) and (22), and rearranging:

$$\frac{\partial l_i}{\partial m} = \frac{\partial l_i^c}{\partial m} - \frac{\partial X(t,R,V)}{\partial m}\frac{\partial l_i}{\partial T}, \quad m = t, R, \quad i = 1, 2, \quad (58)$$

$$\frac{\partial c_1}{\partial m} = \frac{\partial c_1^c}{\partial m} - \frac{\partial X(t,R,V)}{\partial m}\frac{\partial c_1}{\partial T}, \quad m = t, R, \quad (59)$$

where $X(t,R,V) \equiv T$ is the expenditure function associated with the level of (expected) indirect utility V. Following the approach taken by Cremer and Gahvari (1995a, section 2.2), the critical point is to determine the correct income transfer T to keep expected utility constant. We totally differentiate the utility function (Eq. (1)) and the budget constraint (Eq. (3)) to find

$$dU = u'_1 dc_1 - v'_1 dl_1 + \beta E[u'_2 dc_2] - \beta E[v'_2 dl_2], \quad (60)$$

$$dc_2 = -Rdc_1 + R(1-t)\theta_1 dl_1 + (1-t)\theta_2 dl_2 + RdT - (R\theta_1 l_1 + \theta_2 l_2)dt$$
$$+ [(1-t)\theta_1 l_1 + T - c_1]dR. \quad (61)$$

T denotes exogenous lump-sum income. Substitution of Eq. (61) into Eq. (60) yields

$$dU = (u'_1 - \beta R E[u'_2])dc_1 + ((1-t)\beta R E[u'_2\theta_1] - v'_i)dl_1$$
$$+ \beta((1-t)E[u'_2\theta_2|\theta_1] - E[v'_2|\theta_1])dl_2$$
$$+ \beta R E[u'_2]dT - \beta E[u'_2(R\theta_1 l_1 + \theta_2 l_2)]dt$$
$$+ \beta((1-t)E[u'_2\theta_1 l_1] + E[u'_2](T-c_1))dR. \quad (62)$$

By applying the households' first-order conditions, the first two lines in Eq. (62) vanish and the expression reduces to

$$dU = \beta R E[u'_2]dT - \beta E[u'_2(R\theta_1 l_1 + \theta_2 l_2)]dt$$
$$+ \beta((1-t)E[u'_2\theta_1 l_1] + E[u'_2](T-c_1))dR$$
$$= \beta R E[u'_2]dT - \beta R E[u'_2]\left((1-\xi_1)l_1 + \frac{(1-\xi_2)E[\theta_2 l_2]}{R}\right)dt \quad (63)$$
$$+ \beta R E[u'_2]\frac{(1-\xi_1)(1-t)l_1 - c_1 + T}{R}dR,$$

where the second equality results from applying Steiner's Rule, the definition of the insurance characteristic, and $E[\theta_1]=0$. Keeping expected utility constant, $dU=0$, we find the necessary changes in lump-sum income T (i.e., the change in the expenditure function $X(t, R, V)$ to compensate for changes in the labor tax rate t or the net interest factor R):

$$\left.\frac{dT}{dt}\right|_{dU=0} = \frac{\partial X(t,R,V)}{\partial t} = (1-\xi_1)l_1 + \frac{(1-\xi_2)E[\theta_2 l_2]}{R}, \quad (64)$$

$$\left.\frac{dT}{dR}\right|_{dU=0} = \frac{\partial X(t,R,V)}{\partial R} = -\frac{(1-\xi_1)(1-t)l_1 - c_1 + T}{R}. \quad (65)$$

By inserting these expressions in the Slutsky Eqs. (58) and (59), respectively, we obtain the Slutsky Eqs. (23)–(28) in the main text.

Appendix C. Compensated elasticities under risk

To derive the compensated elasticities, we log-linearize the first-order conditions and the expected utility function, where we set the change in the latter to zero. Log-linearization provides a very powerful method to solve for the comparative statics of highly non-linear models, as the system of linearized equations is linear in its relative changes. Since an elasticity is just the ratio of the relative change of an endogenous variable with respect to the relative change of a policy variable, the coefficients in the solved linearized model are just the elasticities we are looking for. Deriving the elasticities from differentiating the equations of the model with respect to the policy variables is mathematically equivalent, but generally extremely cumbersome in highly non-linear models, and often so nontransparent that interpretation of the elasticities becomes impossible. We focus on the elasticities of expected consumption and labor supply in both periods with respect to deterministic (expected) changes in policies. Hence, we can employ the concept of global risk aversion (see, e.g., Varian, 1992, p. 380). We define global relative risk aversion in consumption as $\rho_i \equiv -\frac{E[u''_i(c_i)]E[c_i]}{E[u'_i(c_i)]} > 0$. $\varepsilon_i \equiv \left[\frac{E[v''_i(l_i)]E[\theta_i l_i]}{E[v'_i(l_i)]E[\theta_i]}\right]^{-1} > 0$ is a measure for the expected compensated labor supply elasticity in period $i=1,2$ The log-linearized utility function is given by

$$c_1 u'_1 \tilde{c}_1 - l_1 v'_1 \tilde{l}_1 + \beta E[c_2]E[u'_2]\tilde{c}_2 - \beta \frac{E[\theta_2 l_2]}{E[\theta_2]}E[v'_2]\tilde{l}_2 = 0, \quad (66)$$

where a tilde (˜) denotes a relative change, e.g., $\tilde{c}_i \equiv \frac{E[dc_i]}{E[c_i]}$ is the relative change in the expected value of c_i, and $\tilde{l}_i \equiv \frac{E[d(\theta_i l_i)]}{E[\theta_i l_i]}$ is the relative change in l_i, and where we used the fact that $E[d(\theta_2 l_2)] = E[\theta_2]E[dl_2]$], because we are evaluating the change for a given θ_1, and because l_2 is chosen before θ_2 realizes. Substituting the households' first-order conditions for labor supply and consumption in the linearized utility function, we find, after rearranging,

$$Rc_1\tilde{c}_1 + E[c_2]\tilde{c}_2 - (1-\pi_1)(1-t)RE[\theta_1]l_1\tilde{l}_1 - (1-\pi_2)(1-t)E[\theta_2 l_2]\tilde{l}_2 = 0. \quad (67)$$

Hence,

$$\gamma\tilde{c}_1 + (1-\gamma)\tilde{c}_2 - (1-\pi_1)(1-t)\omega\tilde{l}_1 - (1-\pi_2)(1-t)(1-\omega)\tilde{l}_2 = 0, \quad (68)$$

where we defined $\gamma \equiv \frac{Rc_1}{RE[\theta_1]l_1 + E[\theta_2 l_2]}$ and $(1-\gamma) \equiv \frac{E[c_2]}{RE[\theta_1]l_1 + E[\theta_2 l_2]}$ as the expected expenditure shares of consumption in both periods, and $\omega \equiv \frac{RE[\theta_1]l_1}{RE[\theta_1]l_1 + E[\theta_2 l_2]}$ and $1-\omega \equiv \frac{E[\theta_2 l_2]}{RE[\theta_1]l_1 + E[\theta_2 l_2]}$ as the expected share of labor income in period $i=1,2$ in total labor income (before taxes). Log-linearizing the first-order conditions (before introducing the π_i-terms in labor supply) yields

$$\frac{u''_1 c_1}{u'_1}\tilde{c}_1 = \tilde{R} + \frac{E[u''_2]E[c_2]}{E[u'_2]}\tilde{c}_2, \quad (69)$$

$$\frac{v''_1 l_1}{v'_1}\tilde{l}_1 = -\tilde{t} + \tilde{R} + \frac{E[u''_2\theta_1]E[c_2]}{E[u'_2\theta_1]}\tilde{c}_2, \quad (70)$$

$$\frac{E[v''_2]E[l_2]}{E[v'_2]}\frac{dl_2}{E[d(\theta_2 l_2)]}\frac{E[\theta_2 l_2]}{E[l_2]}\tilde{l}_2 = -\tilde{t} + \frac{E[u''_2\theta_2]E[c_2]}{E[u'_2\theta_2]}\tilde{c}_2. \quad (71)$$

Substituting the definitions of (global) relative risk aversion into Eq. (69) delivers

$$\tilde{c}_2 = \frac{\rho_1}{\rho_2}\tilde{c}_1 + \frac{1}{\rho_2}\tilde{R}. \quad (72)$$

Relying on Steiner's Rule for covariances, we find

$$E[u''_2\theta_i] = E[\theta_i]E[u''_2] + cov[u''_2, \theta_i] = \left(1-\pi'_i\right)E[\theta_i]E[u''_2], \quad (73)$$

$$E[u'_2\theta_i] = E[\theta_i]E[u'_2] + cov[u'_2, \theta_i] = (1-\pi_i)E[\theta_i]E[u'_2], \quad (74)$$

where $\pi'_i = -\frac{cov[u''_2, \theta_i]}{E[u''_2]E[\theta_i]} > 0$, as long as we assume non-increasing absolute risk aversion ($u'''_2 > 0$). Since the π_i' terms are normalized covariances, they are always smaller than or equal to one: $0 < \pi_i \le 1$. Substituting these expressions into Eqs. (70) and (71), we find

$$\frac{v''_1 l_1}{v'_1}\tilde{l}_1 + \tilde{t} - \tilde{R} = \frac{(1-\pi'_1)}{(1-\pi_1)}\frac{E[u''_2]E[c_2]}{E[u'_2]}\tilde{c}_2, \quad (75)$$

$$\frac{E[v''_2]E[l_2]}{E[v'_2]}\frac{dl_2}{E[d(\theta_2 l_2)]}\frac{E[\theta_2 l_2]}{E[l_2]}\tilde{l}_2 + \tilde{t} = \frac{(1-\pi'_2)}{(1-\pi_2)}\frac{E[u''_2]E[c_2]}{E[u'_2]}\tilde{c}_2. \quad (76)$$

Using the definitions of the labor supply elasticities and rearranging yields

$$\tilde{l}_1 = \varepsilon_1\left(-\tilde{t} + \tilde{R}\right) - \Sigma_1 \varepsilon_1 \rho_2 \tilde{c}_2, \tag{77}$$

$$\tilde{l}_2 = -\varepsilon_2 \tilde{t} - \Sigma_2 \varepsilon_2 \rho_2 \tilde{c}_2, \tag{78}$$

where $\Sigma_1 \equiv \frac{1-\pi'_1}{1-\pi_1} \geq 0$ and $\Sigma_2 \equiv \frac{1-\pi'_2}{1-\pi_2} \geq 0$, since $0 < \pi_i \leq 1$. Together with the linearized Euler consumption equation and the linearized utility function we have a linear system of four equations in four unknowns, which can be solved to find the elasticities. First, substitute the linearized Euler Eq. (72) in the other three linearized Eqs. (68), (77), and (78) to find

$$\tilde{l}_1 = -\varepsilon_1 \tilde{t} + \varepsilon_1(1-\Sigma_1)\tilde{R} - \Sigma_1 \varepsilon_1 \rho_1 \tilde{c}_1, \tag{79}$$

$$\tilde{l}_2 = -\varepsilon_2 \tilde{t} - \Sigma_2 \varepsilon_2 \tilde{R} - \Sigma_2 \varepsilon_2 \rho_1 \tilde{c}_1, \tag{80}$$

$$(1-t)\sum_{i=1}^{2}(1-\pi_i)\omega_i \tilde{l}_i = (1-\gamma)\frac{1}{\rho_2}\tilde{R} + \left[\gamma + (1-\gamma)\frac{\rho_1}{\rho_2}\right]\tilde{c}_1. \tag{81}$$

Use the first two equations to substitute for $\tilde{l}_1$ and $\tilde{l}_2$ in the last equation to find the solution of the model for $\tilde{c}_1$:

$$\begin{aligned} &-[(1-\pi_1)\omega\varepsilon_1 + (1-\pi_2)(1-\omega)\varepsilon_2]\tilde{t} \\ &+\left[-\frac{(1-\gamma)/\rho_2}{(1-t)} + (1-\pi_1)\omega\varepsilon_1(1-\Sigma_1) - (1-\pi_2)(1-\omega)\Sigma_2\varepsilon_2\right]\tilde{R} \\ &= \left[\frac{\gamma + (1-\gamma)\frac{\rho_1}{\rho_2}}{1-t} + (1-\pi_1)\omega\Sigma_1\varepsilon_1\rho_1 + (1-\pi_2)(1-\omega)\Sigma_2\varepsilon_2\rho_1\right]\tilde{c}_1. \end{aligned} \tag{82}$$

Using the last result in Eqs. (79), (80) and (72), we can write the solution of the complete model as

$$\tilde{c}_1 = -\frac{\epsilon}{\Delta}\tilde{t} + \frac{\delta}{\Delta}\tilde{R}, \tag{83}$$

$$\tilde{c}_2 = -\frac{\rho_1}{\rho_2}\frac{\epsilon}{\Delta}\tilde{t} + \left[\frac{1}{\rho_2} + \frac{\rho_1}{\rho_2}\frac{\delta}{\Delta}\right]\tilde{R}, \tag{84}$$

$$\tilde{l}_1 = -\varepsilon_1\left[1 - \Sigma_1\frac{\rho_1\epsilon}{\Delta}\right]\tilde{t} - \varepsilon_1\Sigma_1\left[1 + \frac{\rho_1\delta}{\Delta} - \frac{1}{\Sigma_1}\right]\tilde{R}, \tag{85}$$

$$\tilde{l}_2 = -\varepsilon_2\left[1 - \Sigma_2\frac{\rho_1\epsilon}{\Delta}\right]\tilde{t} - \varepsilon_2\Sigma_2\left[1 + \frac{\rho_1\delta}{\Delta}\right]\tilde{R}, \tag{86}$$

where

$$\Delta \equiv \frac{\gamma + (1-\gamma)\frac{\rho_1}{\rho_2}}{(1-t)} + (1-\pi_1)\omega\Sigma_1\varepsilon_1\rho_1 + (1-\pi_2)(1-\omega)\Sigma_2\varepsilon_2\rho_1 > 0, \tag{87}$$

$$\epsilon \equiv (1-\pi_1)\omega\varepsilon_1 + (1-\pi_2)(1-\omega)\varepsilon_2 > 0, \tag{88}$$

$$\delta \equiv -\frac{(1-\gamma)/\rho_2}{(1-t)} + (1-\pi_1)\omega\varepsilon_1(1-\Sigma_1) - (1-\pi_2)(1-\omega)\varepsilon_2\Sigma_2. \tag{89}$$

ϵ is a measure for the weighted labor-supply elasticity, where the certainty equivalent of each period's income is used as a weight. We can sign the elasticities as follows. First, the consumption elasticities with respect to the tax rate are unambiguously signed: $\varepsilon_{c_1 t} < 0$, $\varepsilon_{c_2 t} < 0$. Next, the elasticity of second period consumption with respect to the interest factor is unambiguous as well, $\varepsilon_{c_2 R} > 0$, because $1 + \frac{\rho_1\delta}{\Delta} = \frac{1}{\Delta}\left[\frac{\gamma}{1-t} + (1-\pi_1)\omega\varepsilon_1\rho_1\right] > 0$. Second, as long as we assume $\delta > 0$, the first-period consumption elasticity with respect to the interest factor will be negative, $\varepsilon_{c_1 R} < 0$ and standard saving behavior is obtained.

This assumption holds true if either there is no first-period income, if $\pi'_1 - \pi_1$ is sufficiently small, or if the labor tax rate t is sufficiently high. For $\delta < 0$, a higher net interest factor makes first-period consumption less attractive and second-period consumption more attractive. These signs of the elasticities would also be found in the absence of risk. Third, the elasticity of second-period labor supply with respect to the interest factor is unambiguously negative, i.e. $\varepsilon_{l_2 R} < 0$, since $1 + \frac{\rho_1\delta}{\Delta} = \frac{1}{\Delta}\left(\frac{\gamma}{(1-t)} + (1-\pi_1)\omega\varepsilon_1\rho_1\right) > 0$. Moreover, if δ is negative, then $0 < 1 + \frac{\rho_1\delta}{\Delta} < 1$. Consequently, the first-period labor supply elasticity with respect to the interest factor is negative, if $\Sigma_1 < 1$, since $1 + \frac{\rho_1\delta}{\Delta} - \frac{1}{\Sigma_1} < 0$ in that case. The latter assumption is equivalent to assuming $\pi'_1 > \pi_1$. This is a relatively weak requirement. For the special case of multivariate normally distributed skill shocks, it can be shown that this assumption is equivalent to require (global) absolute prudence being larger than (global) absolute risk aversion. The latter holds for most utility functions and should also carry over to uncertainty under mild conditions. Fourth, we assume that the substitution effect is dominant to obtain standard labor supply behavior, i.e. $\varepsilon_{l,t} < 0$. Thus, we impose $1 - \Sigma_i\frac{\rho_1\epsilon}{\Delta} > 0$. These assumptions imply $-\frac{1}{(1-\pi_2)(1-\omega)\varepsilon_2} < \frac{(\Sigma_2 - \Sigma_1)(1-t)}{\gamma\frac{1}{\rho_1} + (1-\gamma)\frac{1}{\rho_2}} < \frac{1}{(1-\pi_1)\omega\varepsilon_1}$. Therefore, a sufficient condition to ensure standard behavior of labor supply is that the difference between Σ_1 and Σ_2 is not too large (or that they are close to being equal) such that $\Sigma_1 \approx \Sigma_2$ holds.

References

Aiyagari, S.R., 1995. Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. Journal of Political Economy 103, 1158–1175.
Albanesi, S., Sleet, C., 2006. Dynamic optimal taxation with private information. Review of Economic Studies 73, 1–30.
Atkinson, A.B., Sandmo, A., 1980. Welfare implications of the taxation of savings. The Economic Journal 90, 529–549.
Atkinson, A.B., Stiglitz, J.E., 1976. The design of tax structure: direct versus indirect taxation. Journal of Public Economics 6, 55–75.
Banks, J., Diamond, P.A., 2010. The base for direct taxation. In: Mirrlees, J.A., et al. (Ed.), Dimensions of Tax Design: The Mirrlees Review. University Press, Oxford, pp. 548–648 (Ch. 6).
Bastani, S., Blomquist, S., Micheletto, L., 2010. The welfare gains of age-related optimal income taxation. Uppsala Center for Fiscal Studies Working Paper 2010:12, Uppsala.
Bernheim, B.D., 2002. Taxation and saving. In: Auerbach, A.J., Feldstein, M. (Eds.), Handbook of Public Economics, vol. 3. Elsevier, Amsterdam, pp. 1173–1249 (Ch. 18).
Blomquist, S., Christiansen, V., 2008. Taxation and heterogenous preferences. FinanzArchiv 64, 218–244.
Blomquist, S., Micheletto, L., 2008. Age-related optimal income taxation. The Scandinavian Journal of Economics 110, 45–71.
Blundell, R., MaCurdy, T., 1999. Labor supply: a review of alternative approaches. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3. Elsevier, Amsterdam, pp. 1559–1695 (Ch. 27).
Boadway, R., Pestieau, P., 2003. Indirect taxation and redistribution: the scope of the Atkinson–Stiglitz theorem. In: Arnott, R., et al. (Ed.), Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz. MIT Press, Cambridge, pp. 387–403 (Ch. 21).
Chamley, C., 1986. Optimal taxation of capital income in general equilibrium with infinite lives. Econometrica 54, 607–622.
Conesa, J.C., Kitao, S., Krueger, D., 2009. Taxing capital? Not a bad idea after all! American Economic Review 99, 25–48.
Cremer, H., Gahvari, F., 1995a. Uncertainty and optimal taxation: in defense of commodity taxes. Journal of Public Economics 56, 291–310.
Cremer, H., Gahvari, F., 1995b. Uncertainty, optimal taxation and the direct versus indirect tax controversy. The Economic Journal 105, 1165–1179.
Cremer, H., Gahvari, F., 1999a. Uncertainty, commitment and optimal taxation. Journal of Public Economic Theory 1, 51–70.
Cremer, H., Gahvari, F., 1999b. Uncertainty and optimal uniform commodity taxes. Southern Economic Journal 66, 156–167.
Cremer, H., Lozachmeur, J.-M., Pestieau, P., 2004. Social security, retirement age and optimal income taxation. Journal of Public Economics 88, 2259–2281.

Cremer, H., Lozachmeur, J.-M., Pestieau, P., 2008. Social security and retirement decision: a positive and normative approach. Journal of Economic Surveys 22, 213–233.

Da Costa, C., Werning, I., 2002. Commodity taxation and social insurance. MIT, mimeo.

Deaton, A., 1979. Optimally uniform commodity taxes. Economics Letters 2, 357–361.

Diamond, P.A., 2006. Comment on new dynamic public finance: a user's guide. NBER Macroeconomic Annual, pp. 365–379.

Diamond, P.A., Mirrlees, J.A., 1978. A model of social insurance with variable retirement. Journal of Public Economics 10, 295–336.

Diamond, P.A., Mirrlees, J.A., 1986. Payroll-tax financed social insurance with variable retirement. The Scandinavian Journal of Economics 88, 25–50.

Diamond, P.A., Saez, E., 2011. The case for a progressive tax: from basic research to policy recommendations. Journal of Economic Perspectives 25, 165–190.

Eaton, J., Rosen, H.S., 1980. Labor supply, uncertainty, and efficient taxation. Journal of Public Economics 14, 365–374.

Erosa, A., Gervais, M., 2002. Optimal taxation in life-cycle economies. Journal of Economic Theory 105, 338–369.

Fischer, S., 1980. Dynamic inconsistency, cooperation, and the benevolent dissembling government. Journal of Economic Dynamics and Control 2, 93–107.

Golosov, M., Tsyvinski, A., 2006. Designing optimal disability insurance. A case for asset testing. Journal of Political Economy 114, 257–279.

Golosov, M., Kocherlakota, N.R., Tsyvinski, A., 2003. Optimal indirect and capital taxation. Review of Economic Studies 70, 569–587.

Golosov, M., Tsyvinski, A., Werning, I., 2006. New dynamic public finance: a user's guide. NBER Macroeconomic Annual, pp. 317–363.

Gordon, R.H., Kopczuk, W., 2011. The choice of the personal income tax base. Public Economics Programme Discussion Papers No. 11, LSE, London.

Gruber, J., Wise, D.A., 1999. Social Security and the Retirement around the World. The University of Chicago Press, Chicago.

Jacobs, B., Boadway, R., 2012. Optimal linear commodity taxation under optimal nonlinear income taxation. Erasmus School of Economics and Queens University, mimeo.

Judd, K.L., 1985. Redistributive taxation in a simple perfect foresight model. Journal of Public Economics 28, 59–83.

Kocherlakota, N.R., 2005. Zero expected wealth taxes. A Mirrlees approach to dynamic optimal taxation. Econometrica 73, 1587–1621.

Kydland, F., Prescott, E.C., 1977. Rules rather than discretion: the inconsistency of optimal plans. Journal of Political Economy 85, 473–492.

Mankiw, N.G., Weinzierl, M., Yagan, D., 2009. Optimal taxation in theory and practice. Journal of Economic Perspectives 23, 147–174.

Meghir, C., Philips, D., 2010. Labour supply and taxes. In: Mirrlees, J.A. (Ed.), The Mirrlees Review. Dimensions of Tax Design. Oxford University Press, Oxford, pp. 202–274 (Ch. 3).

Meghir, C., Pistaferri, L., 2011. Earnings, consumption and life cycle choices. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 4B. Elsevier, Amsterdam, pp. 773–854 (Ch. 9).

Menezes, C.F., Wang, X.H., 2005. Duality and the Slutsky income and substitution effects of increases in wage rate uncertainty. Oxford Economic Papers 57, 545–557.

Mirrlees, J.A., Adam, S., Besley, T.J., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G.D., Poterba, J.M., 2011. Tax by Design: The Mirrlees Review. Oxford University Press, Oxford.

Musgrave, R., Musgrave, p., 1976. Public Finance in Theory and Practice, second ed. McGraw-Hill, New York.

Nishiyama, S., Smetters, K., 2005. Consumption taxes and economic efficiency with idiosyncratic wage shocks. Journal of Political Economy 113, 1088–1115.

Ordover, J.A., Phelps, E.S., 1979. The concept of optimal taxation in the overlapping-generations model of capital and wealth. Journal of Public Economics 12, 1–26. (reprint third edition 1949)

Pigou, A.C., 1928. A Study in Public Finance. MacMillan & Co. Ltd., London.

Pirttilä, J., Suoniemi, I., 2010. Public provision, commodity demand and hours of work: an empirical analysis. CESifo Working Paper No. 3000, CESifo, Munich.

Saez, E., 2002. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. Journal of Public Economics 83, 217–230.

Sandmo, A., 1974. A note on the structure of optimal taxation. American Economic Review 64, 701–706.

Sandmo, A., 1976. Optimal taxation: an introduction to the literature. Journal of Public Economics 6, 37–54.

Sinn, H.-W., 1996. Social insurance, incentives, and risk-taking. International Tax and Public Finance 3, 259–280.

Varian, H.R., 1992. Microeconomic Analysis, third ed. W.W. Norton & Company, New York.