CESifo Working Papers
www.cesifo.org/wp

Optimal Taxation with Current and Future Cohorts

Hans Fehr
Fabian Kindermann

CESIFO WORKING PAPER NO. 3973
CATEGORY 1: PUBLIC FINANCE
OCTOBER 2012

CESifo *Working Paper No. 3973*

Optimal Taxation with Current and Future Cohorts

Abstract

This note demonstrates that optimal tax calculations in overlapping generations models should not be based exclusively on long-run welfare changes. As the latter represent a mix of efficiency and intergenerational redistribution effects, they typically favor policies which redistribute towards future cohorts. Taking the recent study of Conesa et al. (2009) as an example, we explicitly consider short- and long-run welfare effects and isolate the aggregate efficiency consequences of a tax reform. Based on this aggregate efficiency measure, we find a much lower capital income tax rate and a significantly less progressive labor income tax schedule than Conesa et al. (2009) to be optimal. As we demonstrate, the optimality of capital income taxation is explained by the low interest elasticity of precautionary savings compared to that of life-cycle savings.

JEL-Code: C680, H210, D910.

Keywords: stochastic OLG model, precautionary savings, intragenerational risk sharing and redistribution.

Hans Fehr
Department of Economics
University of Würzburg
Sanderring 2
Germany – 97070 Würzburg
hans.fehr@uni-wuerzburg.de

Fabian Kindermann
Department of Economics
University of Würzburg
Sanderring 2
Germany – 97070 Würzburg
kindermann.fabian@uni-wuerzburg.de

September 2012
We thank participants at university seminars in Dortmund, Bonn and Groningen and Andras Simonovits for many useful comments and discussions. Financial support from the German Research Foundation (FE 377/ 5-2) is gratefully acknowledged.

1 Introduction

Optimal income taxation is one of the oldest, most controversial and most policy relevant topics in public finance. The discussion combines both the debate about the tax base, e.g. whether capital income should be taxed or not, and the question of the optimal tax schedule, especially its progressivity. Recent surveys by Mankiw et al. (2009) as well as Diamond and Saez (2011) in the *Journal of Economic Perspectives* document the ongoing debate and explain the logic underlying the seemingly contradictory results. As the distribution of individual abilities, the life-cycle income process, and individual preferences are the major determinants of the optimal income tax schedule, drawing robust conclusions from quantitative analyses is not a simple task.

Nevertheless, analyzing the optimal tax structure in numerical studies has a long tradition in the literature. The majority of papers are testing the sensitivity of results with respect to certain model assumptions. In contrast, we focus on what is an adequate measure of optimality. Our special interest lies with models of overlapping generations with households facing both borrowing constraints and uncertainty about future labor earnings. Studies in this framework include İmrohoroğlu (1998), Conesa and Krueger (2006) or Conesa et al. (2009) which all aim to quantitatively characterize the optimal capital and/ or labor income tax. The latter of these studies finds that the optimal capital income tax rate is significantly positive at 36 percent and the optimal progressive labor income tax combines a flat tax of 23 percent and a deduction of $ 7,200. A number of recent studies have extended the benchmark model of Conesa at al. (2009) in various directions in order to test the sensitivity of their results. Nakajima (2010) incorporates a housing asset and shows that in this case the optimal capital income tax rate is close to zero. Kitao (2010) demonstrates that it may be optimal to reduce the capital income tax rate when labor supply rises. Fukushima (2010) highlights the optimality of age- and history-dependent income taxes while Kumru and Piggott (2012) analyze the implications of means-tested pension benefits for optimal capital income taxation.

All studies discussed so far focus on steady states and derive the optimal tax system as the one that maximizes long-run welfare. Fukushima (2010) shows that this corresponds to maximizing a utilitarian social welfare function that places equal weights on all cohorts. However, focussing exclusively on long-run welfare and neglecting transitional cohorts is a very arbitrary welfare concept. It may also be misleading in economic terms, since welfare effects arising from efficiency and intergenerational redistribution can not be isolated. Typically, optimal tax schedules derived from long-run welfare maximization come at large costs for transitional cohorts which are not taken into account. To overcome this issue we therefore offer an alternative quantitative assessment of optimality which explicitly accounts for the welfare consequences of all generations, i.e. current and future, and isolates the effects of income redistribution across cohorts from aggregate efficiency.

As in the previous studies mentioned above, we take the model and calibration of Conesa et al. (2009) as a benchmark and start from the same initial equilibrium. Yet our simulation approach differs in that we not only compute long-run equilibrium results of a policy change, but also the transition path and the welfare effects of all current and future cohorts. We then derive the optimal policy schedule with three different welfare concepts. The first one simply calculates the present value of welfare changes along the transition path. The other two concepts follow Huang et al. (1997), Nishiyama and Smetters (2005, 2007) as well as Fehr and Habermann (2008) in isolating pure efficiency effects from intergenerational redistribution by means of lump-sum transfers. Taking pure efficiency as measure of optimality, we still find a positive tax on capital income, but this rate is much lower at 14 percent.

In addition, the optimal income tax schedule turns out to be significantly less progressive with a marginal rate of 17 percent and a lump-sum tax of \$ 712. The optimality of capital income taxation in our model is explained by the low interest elasticity of precautionary savings compared to life-cycle savings. As a consequence, preferences and/ or policies that increase the fraction of life-cycle savings in total savings will reduce the optimal capital income tax rate and vice versa.

The remainder of the paper is arranged as follows: the next section describes our model and its calibration. Section 3 presents simulation results, section 4 concludes.

2 The model economy

The description of our simulation model's structure and calibration follows closely that of Conesa et al. (2009).

2.1 Demographics

The model economy is populated by J overlapping generations. At any discrete point t in time a new generation is born, the mass of which grows at rate n. Agents survive from age j to age $j+1$ with probability ψ_j, where $\psi_J = 0$. Since we abstract from annuity markets, individuals may leave accidental bequests Tr_t that are distributed in a lump-sum manner across the currently alive. Agents retire at age j_r and start to receive social security payments SS_t, which are financed by proportional payroll taxes at rate $\tau_{SS,t}$ that are paid up to an income threshold $\bar{y}$. In the following, we omit the time index t for simplicity reasons wherever possible.

2.2 Endowments and preferences

Individuals enter the economy with zero assets $a_1 = 0$ and are not allowed to run into debt throughout their whole life, i.e. $a_j \geq 0$. During their working phase, they supply part of their maximum time endowment of one unit per period as labor to the market. The remainder of time is consumed as leisure.

Households are heterogeneous along three dimensions that affect their labor productivity. First, average labor productivity ϵ_j varies with age, governing the average wage of a cohort. Second, households are born with permanent differences in productivity, standing in for differences in education and innate abilities. We consider two ability types α_1 and α_2 with equal mass. Finally, workers of same age and ability face idiosyncratic shocks $\eta \in \mathcal{E}$ with respect to their individual labor productivity. The stochastic process for labor productivity status is identical and independent across agents and follows a finite-state Markov chain with stationary transitions over time, i.e.

$$\Pr(\eta' \in \mathcal{E} | \eta) = Q(\eta, \mathcal{E}). \tag{1}$$

Since Q consists of strictly positive entries only, there exists a unique, strictly positive, invariant distribution associated with Q denoted by Π. All individuals start their life with average stochastic productivity $\bar{\eta} = \sum_\eta \eta \Pi(\eta)$, where $\bar{\eta} \in \mathcal{E}$ and $\Pi(\eta)$ is the probability of η under the stationary distribution. Different realizations of the stochastic process then give rise to cross-sectional productivity distributions that become more dispersed as a cohort ages.

At any given time households are characterized by (a, η, i, j), where a are current holdings of assets, η is the stochastic labor productivity status, i is ability type, and j is age. A household of type (a, η, i, j) working l hours commands pre-tax labor income $y = \epsilon_j \alpha_i \eta l w_t$, where w_t is the wage per efficiency unit of labor. Let $\Phi_t(a, \eta, i, j)$ denote the measure of agents of type (a, η, i, j) at date t.

Preferences over consumption c_j and and leisure $1 - l_j$ are assumed to be representable by a time-separable utility function of the form

$$W(c, 1-l) = E\left\{\sum_{j=1}^{J} \beta^{j-1} u(c_j, 1-l_j)\right\},$$

where β is the time discount factor. Expectations are taken with respect to the stochastic processes governing idiosyncratic labor productivity and mortality. Due to additive separability, we can formulate the individual optimization problem recursively:

$$v_t(a, \eta, i, j) = \max_{c,l,a'} \left\{u(c, 1-l) + \beta\psi_j \int v_{t+1}(a', \eta', i, j+1) Q(\eta, \mathrm{d}\eta')\right\}$$

The dynamic budget constraint reads

$$(1+\tau_{c,t})c + a' = [1 + r_t(1-\tau_{k,t})](a + Tr_t) + y + SS_t - \tau_{SS,t}\min\{y, \bar{y}\} - T_t(y_{\text{tax}}),$$

where savings a' and consumption expenditure (including consumption taxes) are financed out of current assets and inheritances (including capital income net of capital taxes at rate τ_k), gross income from labor y or pensions SS_t net of payroll taxes and income taxes according to the tax schedule $T_t(\cdot)$ in period t. y_{tax} is taxable income, see below.

2.3 Technology

We let the production technology be represented by a Cobb-Douglas production function. The aggregate resource constraint is given by

$$C_t + K_{t+1} - (1-\delta)K_t + G_t \leq K_t^{\alpha} N_t^{1-\alpha}, \tag{2}$$

where K_t, C_t, G_t and N_t measure the aggregate capital stock, aggregate private and public consumption and aggregate labor input (in efficiency units) in period t, and α defines the capital share. The depreciation rate for physical capital is denoted by δ.

2.4 Government policy

In each period t, the government engages in three activities: it spends resources, levies taxes and runs a social security system. The social security system collects contributions up to a maximum labor income level $\bar{y}$ from working households and pays benefits SS_t to retirees, independent of their earnings history. The payroll tax rate $\tau_{SS,t}$ is used to balance the budget of the system. The social security system is exogenously given and not subject to the optimization of the policymaker.

Furthermore the government faces an exogenously given consumption path $\{G_t\}_{t=1}^{\infty}$. It finances expenditure by means of a proportional tax $\tau_{c,t}$ on private consumption, taxes on capital and labor income $\tau_{k,t}$ and $T_t(\cdot)$ and public debt B_{t+1}. The government's budget constraint therefore reads

$$G_t + (1+r_t)B_t = \tau_{c,t}C_t + \int [\tau_k y_r + T_t(y_{\text{tax}})]\Phi_t(\mathrm{d}a \times \mathrm{d}\eta \times \mathrm{d}i \times \mathrm{d}j) + (1+n)B_{t+1}. \tag{3}$$

The consumption tax rate is exogenously given, while the income tax schedule $T_t(\cdot)$ balances the budget. Note that we assume the income tax schedule to be invariant along the transition, i.e. it only closes the intertemporal budget. The temporary budget is balanced by debt, where we assume an initial debt level of 0.

Taxable labor income consists of labor earnings net of employer contributions to the pension system $y_l = y - 0.5\tau_{SS,t}\min\{y, \bar{y}\}$. Capital income is fully taxable $y_r = r(a + Tr_t)$. In the initial equilibrium, henceforth denoted by $t=0$, the government taxes the sum of labor and capital income $y_{\text{tax}} = y_l + y_r$ according to the schedule $T_0(\cdot)$. There is no additional capital tax, i.e. $\tau_{k,0} = 0$. The policymaker changes the tax schedule once and for all in the reform period $t = 1$. From that moment capital income is taxed at constant rate $\tau_{k,t}$ and labor income according to the schedule $T_t(\cdot)$, i.e. $y_{\text{tax}} = y_l$. The optimal tax structure ($\tau_{k,t}$ and $T_t(\cdot)$) is the one that maximizes aggregate efficiency as defined below.

2.5 Functional forms and calibration

We use the very same calibration as Conesa et al. (2009) in their benchmark scenario. Households are born at age 20 (model age 1), retire at age 65 (model age $j_r = 46$) and may reach a maximum age of 100 years (model age $J = 81$). Population grows at an annual rate of $n = 0.011$, conditional survival probabilities are taken from Bell and Miller (2002). We assume standard Cobb-Douglas preferences

$$u(c, 1-l) = \frac{(c^\gamma(1-l)^{1-\gamma})^{1-\sigma}}{1-\sigma}, \tag{4}$$

where γ is a share parameter and σ determines the risk aversion of the household. We set $\sigma = 4$, $\beta = 1.001$ and $\gamma = 0.377$ in order for the capital-output ratio to be 2.7 and the share of hours worked in total time endowment to amount to 0.33.

We take Hansen's (1993) age-productivity profile $\{\epsilon_j\}_{j=1}^{j_r}$. Abilities α_1 and α_2 are specified so as to match the cross-sectional variance of household labor income at age 22 reported in Støresletten et al. (2004). We assume the idiosyncratic part of the wage process to be a seven-state discretized version of an AR(1) process with persistence parameter ρ and unconditional variance σ_η^2. Our choice of these two parameters targets the cross-sectional household age-earnings variance profile reported in Støresletten et al. (2004).

Both a capital share parameter of $\alpha = 0.36$ and a depreciation rate of $\delta = 0.08$ guarantee a realistic investment-output ratio. The payroll tax rate $\tau_{SS,t}$ is 12.4 percent and the maximum labor income level $\bar{y}$ amounts to 2.5 times the average income. The social security benefit level is endogenous in the initial equilibrium and balances the pension budget. We keep it constant in the reform periods. Government spending G accounts for 17 percent of GDP and remains constant per capita in all future periods. Consequently, the ratio G/Y would decline, if output increased in consequence of a change in tax policy. We initially abstract from public debt and choose a consumption tax rate $\tau_{c,t}$ of 5 percent. Finally, the tax function is given by

$$T(y_{\text{tax}}) = T(y_{\text{tax}}, \kappa_0, \kappa_1, \kappa_2) = \kappa_0 \left[y_{\text{tax}} - (y_{\text{tax}}^{-\kappa_1} + \kappa_2)^{-1/\kappa_1}\right], \tag{5}$$

where κ_i are parameters. This functional form proposed by Gouveia and Strauss (1994) is typically employed in the quantitative finance literature. κ_0 controls the level of the top marginal tax rate and κ_1 determines the progressivity of the tax code. We yet extend out functional choice set in two dimensions. We therefore let

$$T(y_{\text{tax}}) = \begin{cases} \kappa_0 \cdot y_{\text{tax}} + \kappa_2 & \text{for } \kappa_1 = 0 \\ \kappa_0 \ \max[y_{\text{tax}} - \kappa_2, 0] & \text{for } \kappa_1 = \infty, \end{cases} \tag{6}$$

i.e. in the first case we have a proportional tax paired with a lump-sum tax of κ_2 and in the case of $\kappa_1 = \infty$ a flat tax with a deduction of κ_2. In order to approximate the existing U.S. income tax system we set $\kappa_0 = 0.258$ and $\kappa_1 = 0.768$ in the initial equilibrium and adjust κ_2 to balance the budget.

2.6 Welfare and efficiency calculation

We refer to current generations as generations having already been economically active in the initial equilibrium. Future generations are generations that enter the labor force in or after the reform year. We distinguish generations according to the year of their labor market entry t. Consequently, the generation that just entered the labor force in the reform year is indexed $t=1$, the generation aged $j=2$ in the reform year is index $t=0$, the generation aged $j=3$ with $t=-1$, etc.

Given the specific form of the utility function, the welfare consequences of switching from the initial allocation $(c_0, 1-l_0)$ to a new allocation $(c_*, 1-l_*)$ for a specific current or future cohort t can be computed from

$$CEV_t = \left[\frac{W(c_*, 1-l_*)}{W(c_0, 1-l_0)}\right]^{\frac{1}{\gamma(1-\sigma)}} - 1,$$

where $W(c, 1-l)$ is expected lifetime utility. CEV_t is the percentage change in consumption at all ages and all states of the world, which makes an individual in the initial allocation as well off as in the new allocation. We can compare all current generations in the reform year $t=1$ and all future cohorts along the transition path with their respective counterparts in the initial equilibrium, since their individual state variables are identical.

In order to derive an aggregate welfare measure for a specific policy, we compute the present value of welfare changes for all existing and future cohorts along the transition path, i.e.

$$SW = \sum_{t=1-J}^{\infty} \frac{R_t}{(1+n)^{1-t}} \cdot \frac{W_t(c_*, 1-l_*) - W_t(c_0, 1-l_0)}{\lambda_t}, \quad \text{where} \quad R_t = \frac{\prod_{k=t+1}^{1}(1+r_k)}{\prod_{k=2}^{t}(1+r_k)} \tag{7}$$

defines the discount factor and λ_t denotes the marginal utility of income of generation t. In the tables below, this sum is converted into an annuity and reported as a percentage of initial aggregate consumption.

Assessing aggregate efficiency consequences is far more difficult. As mentioned above this includes separating efficiency from intergenerational redistribution. We chose two different approaches to fulfill this task. The first follows Huang et al. (1997) and compute the present value of additional wealth required to make all individuals along the transition path (i.e. current and future cohorts) indifferent to remaining under the initial system. Next, we derive an annuity from this stock measure which is payed out to any future generation. As final result all current generations face a welfare level equal to the one they experienced in the initial equilibrium.[1] All future generations are at the same utility level W^*. We call this level compensated expected utility and use it to calculate the

[1] Their CEV^c therefore is zero.

compensated relative consumption change CEV^c. A positive CEV^c indicates a Pareto improvement (after lump-sum compensation), a reform inducing a negative CEV^c is Pareto inferior. Consequently, we may interpret CEV^c as a measure of efficiency.

Compensating transfers induce behavioral reactions. Hence, factor markets and the government budget are not in equilibrium anymore when we calculate compensated welfare changes. Our second approach overcomes this partial equilibrium (p.e.) framework and takes into account all general equilibrium (g.e.) repercussions.[2] Compensating transfers are computed in exactly the same way as in the partial equilibrium framework. Yet, transfers between generations can only be processed through the asset market and therefore trigger price reactions. [3]

3 Simulation results

We now want to study the optimal tax structure in our model. Following Conesa et al. (2009) we allow for a flexible labor income tax code and restrict capital income taxes to be proportional, i.e.

$$T(y_l, \kappa_0, \kappa_1, \kappa_2) + \tau_k y_r.$$

Thus the government optimizes four parameters $(\kappa_0, \kappa_1, \kappa_2, \tau_k)$, where κ_2 is determined by budget balance. Given a specific parameter choice we quantify the macroeconomic, welfare and efficiency effects along the transition path and in the new long-run equilibrium. The optimal tax structure then is the one that maximizes our efficiency measure. We finally test the sensitivity of our results with respect to the openness of the economy, individual risk aversion and assumptions about initial government debt.

3.1 The optimal tax system

Long-run welfare comparisons Taking long-run welfare as measure of optimality, Conesa et al. (2009) find a capital income tax rate τ_k of 36 percent as well as a labor income tax schedule with marginal rate of 23 percent ($\kappa_0 = 0.23$) and a deduction of about \$7,200 (which corresponds to $\kappa_1 \approx 7$ and $\kappa_2 = 34711$) to be the optimal choice of the policy maker. The first column of Table 1 reports the resulting long-run changes in aggregate variables.[4] Conesa et al. (2009) restrict their parameter choice set to finite values for κ_1. The flat tax case (i.e. $\kappa_1 = \infty$) is only considered in the sensitivity analysis of Table 6 (p. 47) yet without welfare calculations. However, this special case turns out to be the optimal one in terms of long-run welfare maximization. The resulting tax schedule and long-run macro- and welfare effects are shown in the second column of Table 1. A higher tax rate on capital income and a lower marginal tax rate on labor income induce individuals to work more and save less compared to the Conesa et al. (2009) scenario. The long-run gain in equivalent consumption increases from 1.31 to 1.48 percent.

---

2 Note that this implies running the complete model twice: once without and once with compensating transfers.

3 The computation of compensating transfers in general equilibrium dates back to the work of Auerbach and Kotlikoff (1987, 62f.) and has recently been applied by Nishiyama and Smetters (2005, 2007) as well as Fehr and Habermann (2008) in similar stochastic frameworks.

4 The figures correspond to the ones in Table 2 (p. 36) of Conesa et al. (2009). Slight differences in values arise from a different computational method, see Appendix A.

Table 1: Optimal tax schemes: Long-run welfare vs. aggregate efficiency

	Long-run welfare		Aggregate efficiency	
	Conesa et al. (2009)	optimal scheme	base case	optimal scheme
τ_k	0.36	0.43	0.43	0.14
κ_0	0.23	0.20	0.20	0.17
κ_1	7	∞	∞	0
κ_2	34711	12108	12195	712
Average hours worked	-0.66	0.69	0.72	5.84
Total labor supply N	-0.18	1.18	1.19	5.04
Capital stock K	-6.50	-8.16	-8.02	11.14
Government debt to GDP (in %)	0.00	0.00	-0.72	2.98
Output Y	-2.50	-2.29	-2.23	7.20
Aggregate consumption C	-1.45	-0.34	-0.30	7.59
Long run CEV	**1.31**	**1.48**	**1.54**	-0.66
SW			-1.89	0.75
CEV^c (p.e.)			-6.29	2.15
CEV^c (g.e.)			-1.66	**1.07**
Political support (in %)			33.63	72.38

All macro figures are reported as changes in percent of the initial equilibrium values.
Welfare figures are reported as percentage of initial aggregate (SW) or household consumption.

Transitional dynamics Up to now, transitional dynamics were completely neglected. In the third column of Table 1 we consequently simulate the complete transition path for the tax function maximizing long-run welfare. The optimal marginal tax rates on capital and labor income resulting from this exercise turn out to be identical to those of the previous simulation. The major difference is that we adjust κ_2 only once in the reform year $t = 1$ and keep it constant afterwards. Since public debt then balances the periodic government budget, reported long-run macro and welfare effects slightly differ.[5] The cohort welfare effects of this policy reform are depicted in the left panel of Figure 1.[6] Most importantly, due to capital income tax rates increasing significantly in year 1 of the transition, we find long run welfare gains to come along with dramatic welfare losses for initial cohorts who have already accumulated some assets. Summing up across cohorts, these initial losses dominate future welfare gains. Consequently, the present value of all welfare changes is negative and equivalent to an annual reduction in consumption of almost 2 percent. The proposed tax schedule also significantly reduces aggregate efficiency. In partial equilibrium future cohorts experience a compensated welfare loss equivalent to a drop in consumption of more than 6 percent. In general equilibrium this loss reduces to 1.66 percent, which is due to factor price adjustment damping individual reactions and therefore compensating transfers.[7] The left panel of Figure 1 also visualizes the difference between long-run welfare gains and aggregate efficiency. Shifting from labor to capital income taxation

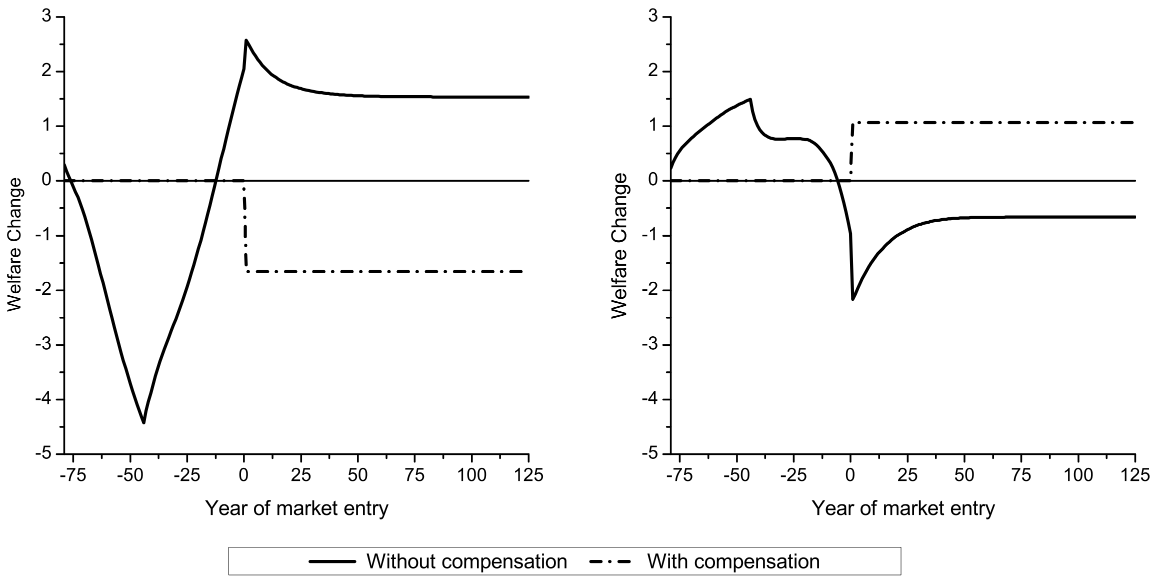
5 Alternatively, we could also keep public debt unaltered and adjust κ_2 in every period to balance the budget. That would yield identical long run values in columns two and three. The difference in efficiency numbers are negligibly small.

6 Note that we report average welfare changes for current cohorts.

7 Transfers to current cohorts amount to 33.5 percent of GDP in partial equilibrium and only 17.6 in general equilibrium.

induces dramatic welfare losses for older current generations, while younger current and future cohorts benefit from reduced tax burdens. The compensation mechanism eliminates intergenerational redistribution and reveals the aggregate efficiency loss. All initial cohorts, which experience rising burdens from capital income taxes, receive lump-sum transfers which bring them back to the welfare level of the initial equilibrium. Younger and future cohorts have to finance these transfers by means of lump-sum taxes, so that they end up at a lower welfare level. Given the dramatic welfare losses for current cohorts, it is not surprising that the considered reform is quite unpopular. The last row of Table 1 reveals that only one third of the population in the reform year experiences welfare gains and therefore would vote in favor of such a reform.[8]

Figure 1: Intergenerational welfare effects: Base case vs. optimal scheme



Efficiency comparisons But what would an optimal scheme look like that takes aggregate efficiency as measure of optimality? The next column of Table 1 answers this question. It features both lower tax rates on capital and labor and comes along with a lump-sum tax of $ 712. This tax schedule induces individuals to work longer hours and save more. In consequence, long-run labor supply and capital stock increase by more than 5 and 11 percent, respectively. The right panel of Figure 1 shows that current cohorts benefit from this tax structure while tax burdens on young and future generations increase. Long-run welfare therefore declines by an amount equivalent to 0.66 percent of consumption. Aggregating welfare changes across cohorts shows that initial welfare gains dominate future welfare losses. The present value of all welfare changes is now equivalent to an annual increase in consumption of 0.75 percent. Applying our compensation mechanism now yields an aggregate efficiency gain of more than 2 percent of consumption in partial equilibrium and of more than 1 percent in general equilibrium. Note that since the optimal scheme balances efficiency losses from behavioral distortions and benefits arising from loosened liquidity constraints as well as the provision of insurance against labor market risk, it does not completely rely on lump-sum taxation. Lump-sum taxes do not distort individual labor supply and savings, but enforce borrowing constraints at young

8 Of course, this political economy interpretation has to be taken with care. It is only valid if there is a one time vote with full commitment to the reform ever after.

ages and provides little to no insurance against income fluctuations. Not surprisingly, the optimal tax system receives political support from more than 72 percent of current households.

Our simulations reveal that the optimal capital income tax declines when the weight of transitional cohorts rises in the social welfare function. This is quite intuitive given the intergenerational welfare effects reported in Figure 1, but stands in stark contrast to the findings of Conesa et al. (2009, p. 45). Maximizing a utilitarian social welfare function for all cohorts alive in the initial steady state, they find an optimal capital income tax rate of 65 percent and conclude that the optimal tax system is "...more strongly geared towards capital income taxes" when transitional cohorts instead of only long run cohorts are taken into account. Since Conesa et al. (2009) do not report their social welfare function, it is not possible to reconstruct their finding. All we can say is that their claim is somewhat counter-intuitive from our perspective. To further support our results, we also derived the optimal tax scheme that maximizes the sum of welfare changes (SW). It turned out that it is very similar to the one reported in the last column of Table 1. If we discounted future cohorts even stronger than we do in SW, the optimal capital income tax rate would decline further.

3.2 Why to tax capital income?

But why is it optimal from an efficiency point of view to tax capital income in the present model? The literature offers two possible explanations. On the one side Cagetti (2001) and Bernheim (2002, p. 1199) point out the fact that the savings motive is important for the interest elasticity of savings. While precautionary savings are fairly inelastic, life-cycle savings for old age are very elastic. Consequently, if the fraction of precautionary savings in total savings is high enough, it will be optimal to tax capital income. On the other side, Erosa and Gervais (2002) demonstrate that even without precautionary savings it might be optimal to tax capital income in a life cycle model. Since leisure consumption increases over the life-cycle and the government cannot tax leisure directly, optimal consumption taxes should increase with individuals' age. When income and consumption taxes cannot be conditioned on age, a nonzero tax on capital income can (imperfectly) mimic such a system.

Certain income In order to clarify the importance of these two explanations for our model economy, we consider a version without income uncertainty. This is done by setting the variances of both persistent and transitory shocks to zero, i.e. $\alpha_1 = \alpha_2 = 1$ and $\sigma_\eta^2 = 0$. Consequently, there is only one representative individual in each cohort that features the average productivity profile ϵ_j over the life-cycle. We finally adjust κ_2 in order to keep constant income tax revenues as a fraction of GDP. In absence of a precautionary savings motive, the initial capital-output ratio declines and the interest rate increases to 6.05 percent. We now consider two policy scenarios. In the left column of Table 2, we allow the government to raise lump-sum taxes. Not surprisingly, the optimal tax system then is a pure lump-sum tax of $ 7270 per household. Such a tax eliminates all labor supply and savings distortions. The resulting benefits dominate the costs arising from higher liquidity constraints. People work and save much more than under the initial system, so that long-run labor supply, capital stock, output and consumption increase strongly. As a consequence, overall efficiency rises by 12.6 percent of aggregate resources.

In a second step we now would like to examine how the taxation of labor and capital income relate. We therefore compute for an exogenous set of marginal tax rates on labor income κ_0 the optimal tax rate on capital income τ_k, given that $\kappa_1 = 0$ and that a lump-sum tax κ_2 balances the government's

Table 2: Optimal tax schemes: Certain vs. uncertain income

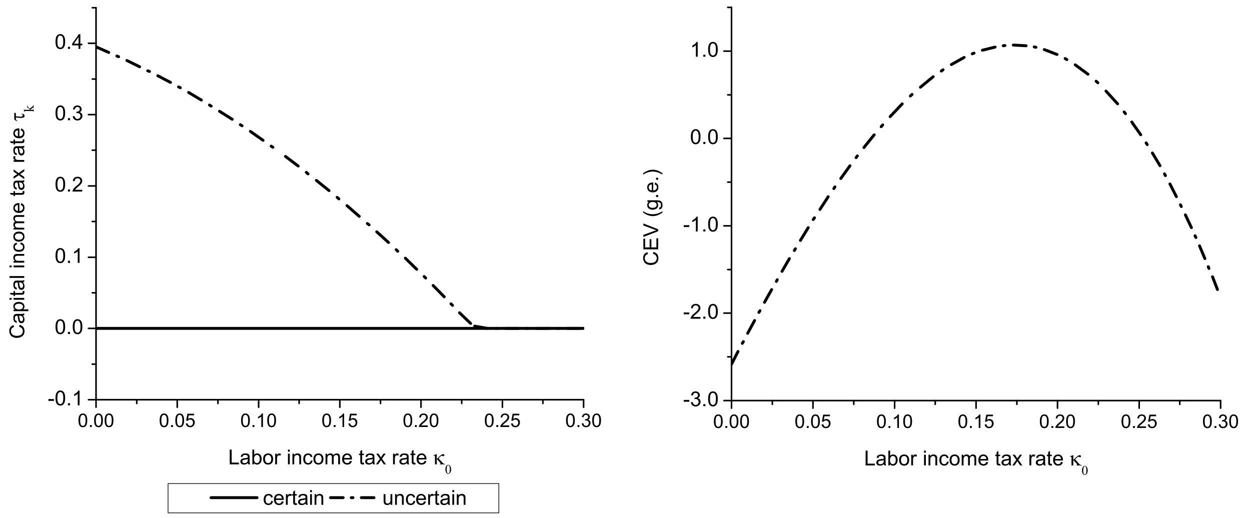
	Certain income		Uncertain income	
Lump-sum taxes	yes	no	yes	no
K_0/Y_0	2.50	2.50	2.72	2.72
r_0	6.05	6.05	4.89	4.89
τ_k	0.00	0.16	0.14	0.17
κ_0	0.00	0.19	0.17	0.19
κ_1	0	0	0	0
κ_2	7270	0	712	0
Average hours worked	17.74	4.04	5.84	3.90
Total labor supply N	17.77	4.07	5.04	3.72
Capital stock K	39.35	7.72	11.14	6.69
Government debt to GDP (in %)	5.19	2.68	2.98	2.14
Output Y	25.13	5.37	7.20	4.78
Aggregate consumption C	26.79	6.00	7.59	5.36
Long run CEV	9.90	2.08	-0.66	-0.57
SW	7.40	2.68	0.75	0.70
CEV^c (p.e.)	13.88	6.50	2.15	2.10
CEV^c (g.e.)	**12.61**	**4.09**	**1.07**	**1.00**
Political support (in %)	72.22	100.00	72.38	74.94

All macro figures are reported as changes in percent of the initial equilibrium values.
Welfare figures are reported as percentage of initial aggregate (SW) or household consumption.

budget. The solid line in the left panel of Figure 2 shows the resulting tax rates for values of κ_0 between 0.00 and 0.30. Not surprisingly, when lump-sum tax instruments are available, it is always optimal to not tax capital income. As the interest elasticity of old-age savings is very high, it is more favorable to avoid savings distortions rather than taxing leisure consumption in late life. This decision is independent of the marginal tax rate on labor income. But what if the government was not allowed to raise lump-sum taxes? The answer is given in the second column of Table 2. In this case a combination of positive labor and capital income taxes is the optimal choice of the government. Although savings are very elastic, capital income should be taxed on efficiency grounds, since such a tax indirectly burdens leisure consumption in late life. This is the argument of Erosa and Gervais (2002), which yet is only valid in a certain income world when lump-sum taxation is forbidden.

Uncertain income Now we turn back to the case of uncertain income and again derive optimal capital income tax rates for different values of κ_0 under the assumption that lump-sum taxes κ_2 balance the budget. The result is depicted as dashed line in the left panel of Figure 2. The picture looks quite different in this case meaning that we find the optimal capital income tax rate to now be strictly positive for many values of κ_0 although lump-sum taxation is available. The reason for this observation is, as discussed by Cagetti (2001) and Bernheim (2002, p. 1199), that the interest elasticity of precautionary savings is much lower than that of life-cycle savings. Therefore it is optimal to tax interest income in order to burden leisure consumption of the elderly. Furthermore in a model with uncertain income, the taxation of capital income also indirectly fulfills a redistributive role. Since it is those who are lucky in the labor market who tend to save the most, assets (at least imperfectly) mirror

Figure 2: Optimal tax rate combinations



the earnings history of an individual. Therefore, a capital income tax redistributes from high to low income individuals. As a consequence, when the marginal tax rate on labor income increases and therefore the labor income tax comprises more redistributive elements, the optimal capital income tax declines. In order to derive the optimal overall tax schedule, the right panel of Figure 2 shows the efficiency numbers that result from the respective tax schedules in the left panel. The efficiency curve obviously peaks for $\kappa_0 = 0.17$ and $\tau_r = 0.14$, i.e. our benchmark case which is reported once more in the third column of Table 2.

The last column of Table 2 complements the discussion of this subsection by assuming that lump-sum taxes are again not available. The optimal (proportional) capital and labor income tax rates then increase to 17 and 19 percent, respectively. This dampens labor supply and capital accumulation during the transition and results in a significantly lower long-run output. Not surprisingly, the elimination of lump-sum taxation slightly reduces the aggregate efficiency gain compared to the optimal tax schedule in the third column. Because of the missing lump-sum tax instrument, the optimal capital income tax rate lies far above the value of roughly 0.10 reported for $\kappa_0 = 0.19$ in the left panel of Figure 2. The reason is again the same as in the case with certainty and no lump-sum taxation.

Summing up the quantitative results of this section, we conclude that the consideration of transitional cohorts in the social welfare function still generates a positive optimal capital income tax, yet at a much lower rate as the one found by Conesa et al. (2009), who only consider long-run cohorts. The central reason for the positive capital income tax rate is the low elasticity of precautionary savings compared to life-cycle savings.

3.3 Sensitivity analysis

In this section we test the sensitivity of our results with respect to assumptions about the openness of the economy, individual risk aversion and initial government debt. Conesa et al. (2009) already provide an extensive discussion about variations in labor supply elasticities. They state that lower elasticities imply higher taxes on labor and lower taxes on capital income. In the limiting case with fixed labor supply, it is optimal to tax labor income at 100% and distribute the resulting revenue in a

Table 3: Sensitivity analysis I

	Optimal scheme	Smopec Reference	Smopec Optimal	Risk aversion $\nu = 0$	Risk aversion $\nu = 8$
$K_0/\ Y_0$	2.72	2.72	2.72	2.65	2.80
r_0	4.89	4.89	4.89	5.28	4.55
τ_k	0.14	0.14	0.06	0.00	0.24
κ_0	0.17	0.17	0.14	0.08	0.22
κ_1	0	0	0	0	5
κ_2	712	700	2145	4429	36415
Average hours worked	5.84	4.09	6.30	15.40	0.24
Total labor supply N	5.04	3.38	4.53	11.82	1.03
Private assets A	11.14	20.88	38.97	31.93	1.17
Government debt to GDP (in %)	2.98	1.04	-0.66	8.00	0.78
Output Y	7.20	3.38	4.53	17.74	1.00
Aggregate consumption C	7.59	7.49	12.16	18.14	1.34
Long run CEV	-0.66	-0.52	-0.88	3.08	0.77
SW	0.75	0.97	1.02	2.63	0.00
CEV^c (p.e.)	2.15	3.00	3.14	5.56	0.16
CEV^c (g.e.)	**1.07**	**1.97**	**2.07**	**4.77**	**0.42**
Political support (in %)	72.38	70.30	56.41	48.03	53.95

All macro figures are reported as changes in percent of the initial equilibrium values.
Welfare figures are reported as percentage of initial aggregate (SW) or household consumption.

lump-sum fashion across individuals. Capital income then stays untaxed. This completely eliminates borrowing constraints and labor income risk without triggering distortions on individual behavior. When we reduced the intertemporal elasticity of substitution, on the other hand, the optimal tax rate on capital income would increase.

The openness of the economy The left column of Table 3 again reports the results for the optimal scheme developed in the previous sections. The next column shows results from simulating a small open economy with the very same initial tax schedule. In a small open economy factor prices remain constant and the national capital market is balanced by cross-border capital flows. Without factor price repercussions behavioral reactions are much stronger. The elasticities of the tax base therefore increase. Consequently, when doing the same reform as in the closed economy setting, i.e. reduce marginal tax rates on both labor and capital income, the resulting efficiency gains are significantly higher.[9] In addition, aggregate efficiency could be even increased by reducing marginal tax rates on capital and labor income further while at the same time increasing lump-sum taxes, see the third column in Table 3.

[9] Note that in this case we can compare aggregate efficiency effects since they refer to the same initial equilibrium. However, when capital intensities and therefore interest rates differ initially, it is not possible to compare aggregate efficiency effects.

Individual risk aversion The last two columns on the right side of Table 3 show the optimal income tax regime when we alter the degree of risk aversion. We isolate the elasticity of intertemporal substitution from relative risk aversion by rewriting the preference structure following Epstein and Zin (1991) as

$$v_t(a,\eta,i,j) = \max_{c,a',l}\ u(c,1-l) + \frac{\beta\psi_j}{1-\sigma}\left\{\int \left[(1-\sigma)v_{t+1}(a',\eta',i,j+1)\right]^{\frac{1-\nu}{1-\sigma}} Q(\eta,\mathrm{d}\eta')\right\}^{\frac{1-\sigma}{1-\nu}}$$

where ν denotes the risk aversion parameter. Setting $\nu = 4.0$ we are back at the benchmark calibration.[10] Like in the previous section we adjust κ_2 in order to achieve the same income tax revenues in relation to GDP.[11] When individuals become less risk averse the fraction of precautionary savings in total savings decreases. This causes the interest rate elasticity of total savings to rise, which in turn leads to lower optimal marginal tax rates on capital income. As shown in the fourth column of Table 3, with risk-neutral preferences capital income should not be taxed at all. In addition, less risk averse individuals attach less value to insurance provision, which causes the marginal tax rate on labor income to decline. Yet, the optimal κ_0 is still positive, which is due to liquidity constraints binding to strongly under a full lump-sum tax regime. On the other hand, when individuals become more risk averse, the savings elasticity decreases. In consequence, optimal marginal tax rates on capital income are much higher than in the benchmark scenario. The last column of Table 3 reveals a risk aversion parameter of 8 to imply marginal tax rates of 24 percent on capital and 22 on labor income, respectively. Note that this is the only preference combination for which we obtain a tax schedule of the original Gouveia and Strauss (1994) form.[12]

Government debt Up to now government debt played hardly any role. It was only adjusted to balance the annual budget while the intertemporal budget was always in balance. In Table 4 we therefore compare optimal tax systems when the initial equilibrium features different levels of public debt (or assets). We again set κ_2 so as to guarantee identical income tax revenues. Public expenditure also needs to be adjusted, as part of the tax revenue needs to be spend on interest payments on existing debt. A positive amount of public debt crowds out part of the capital stock. This causes interest rates to rise which again has two effects on household behavior. First, the fraction of old-age savings in total savings increases, so that the interest elasticity of savings is higher with a higher level of public debt. Second, wanting to save more, households will less frequently run into the problem of liquidity constraints.[13] As a results, both optimal marginal tax rates on capital and labor income decrease the higher the government is indebted. Again this result is in stark contrast to the findings of Conesa et al. (2009, p.46), who argue that rising interest rates increase optimal capital income taxes. Of course, the different results are due to the different welfare criteria which are applied to asses

[10] With Epstein and Zin (1991) preferences altering the risk aversion parameter ν differs from altering the intertemporal elasticity of substitution σ. While the former values different realizations of income *within* a period, the latter values the realization of income *across* periods.

[11] For our sensitivity analysis, we also tried out model versions in which we additionally adjusted the time discount factor β in order to yield the same initial capital to output ratio and interest rates. The differences to the presented results were negligible.

[12] A value of $\nu = 8$ seems quite unreasonable for macroeconomists but it is not uncommon in the finance literature, see Cecchetti et al. (2000).

[13] Note that in a model version in which we adapt the time discount factor to yield the same initial capital-output ratio the reasoning about the effects of government debt would be exactly the same. Yet, higher life-cycle savings would result from a higher discount factor.

Table 4: Sensitivity analysis II

	Initial government debt (B_0/ Y_0)			
	1.00	0.50	0.00	-0.50
$K_0/\ Y_0$	2.55	2.63	2.72	2.84
r_0	5.79	5.37	4.89	4.33
τ_k	0.08	0.11	0.14	0.17
κ_0	0.14	0.16	0.17	0.19
κ_1	0	0	0	0
κ_2	1548	945	712	270
Average hours worked	8.21	6.51	5.84	4.31
Total labor supply N	6.92	5.60	5.04	3.77
Capital stock K	16.28	13.07	11.14	7.87
Government debt to GDP (in %)	103.94	53.92	2.98	-47.05
Output Y	10.20	8.23	7.20	5.23
Aggregate consumption C	9.95	8.32	7.59	5.71
Long run CEV	-1.29	-0.97	-0.66	-0.56
SW	1.14	1.07	0.75	0.48
CEV^c (p.e.)	3.78	3.02	2.15	1.30
CEV^c (g.e.)	**1.95**	**1.15**	**1.07**	**0.56**
Political support (in %)	66.70	72.19	72.38	73.35

All macro figures are reported as changes in percent of the initial equilibrium values.
Welfare figures are reported as percentage of initial aggregate (SW) or household consumption.

economic optimality. Higher public debt increases the burden of future cohorts. Consequently, it is optimal to neutralize (at least partly) this intergenerational redistribution via higher capital income taxes. Taking into account transitional cohorts and maximizing economic efficiency leads to exactly the opposite conclusion.

4 Discussion and conclusion

Our analysis reveals that it is quite misleading to exclusively focus on long-run welfare consequences of tax reforms. Neglecting transitional effects biases results towards capital income taxation burdening initial cohorts who are not represented in the social welfare function. In addition, suggested policy reforms would hardly receive any political support. If transitional cohorts are taken into account, it is still optimal to tax capital income, but the optimal capital income tax rate reduces from 43 to 14 percent in the closed economy and to 6 percent in the small open economy. The driving factor behind the taxation of capital income is the low interest elasticity of precautionary savings compared to life-cycle savings. When the fraction of the former in total savings decreases, e.g. due to lower risk aversion or higher interest rates, the overall savings elasticity rises and it might be optimal to desist completely from capital income taxation. If households feature higher risk aversion or initial interest rates decline, the optimal capital tax rate increases and the optimal labor tax schedule turns out to be more progressive.

Our results have to be interpreted with care, since they also depend on restrictions of the considered tax system and various other assumptions about the economy and individual behavior. We have assumed that the optimal tax scheme is set once at time zero and then maintained into perpetuity. Of course, a full analysis would search for the best time-indexed sequences of labor and capital income taxes across the transition that maximizes aggregate efficiency. However, this type of analysis is challenging in the overlapping generation economy because it quickly runs up against the curse of dimensionality. Furthermore, increasing uncertainty of the economic environment could result in a more progressive tax system. Uncertainty would for example rise, if unintended bequests were distributed in proportion to ability or realized income, if the pension system was less progressive or the income process more volatile. The progressivity of the optimal tax system will decline with rising labor supply elasticities. This includes considering household production or human capital formation.

A final remark refers to linkages between and within generations due to one- or two-sided altruism. While our approach does not directly incorporate any bequest motive, our compensation mechanism can be interpreted as a system of private transfers neutralizing all intra- and intergenerational redistribution effects. Effectively we are then in a Barro (1974) world, where successive generations are linked by an operative altruistic bequest motive. With this interpretation our approach includes both extremes, the standard overlapping generations model without intergenerational linkage and the infinitely lived agent model, in which cohorts are perfectly linked via bequest motives.

A Computational appendix

We use two distinct solution algorithms: one to solve the household problem and one to solve for macroeconomic quantities and prices.

A.1 Solving the household problem

We first have to discretize continuous elements of the state space (a, η, i, j), respectively the asset dimension. We therefore choose $\hat{\mathcal{A}} = \{\hat{a}^1, \ldots, \hat{a}^{n_A}\}$. We then solve the household problem by backward induction, iterating on the following steps:

1. Compute household decisions at maximum age J for any $(\hat{a}, \eta, i, J)$. Since households are not allowed to work anymore and they die for sure in the next period, they consume all remaining resources.

2. Find the solution to the household optimization problem for all possible $(\hat{a}, \eta, i, j)$ recursively using a line search method à la Powell, see Press et al. (2001, 406ff.). This algorithm requires a continuous function to optimize. We therefore use an interpolated version of $v_{t+1}(a, \eta, i, j+1)$. Having computed the data $v_{t+1}(\hat{a}, \eta, i, j+1)$ at any discrete asset grid point in the last iteration step, we can find a piecewise polynomial function $sp_{t+1,j+1}$ satisfying the interpolation conditions
$$sp_{t+1,j+1}(\hat{a}^k, \eta, i, j+1) = \int v_{t+1}(\hat{a}^k, \eta', i, j+1) Q(\eta, \mathrm{d}\eta') \tag{8}$$
for all $k = 1, \ldots, n_A$. We use the multidimensional spline interpolation algorithm described in Habermann and Kindermann (2007).

We choose $n_A = 25$. We also tried higher values, but the results didn't change.

A.2 The macroeconomic computational algorithm

We solve for quantities and prices using a Gauss-Seidel procedure in line with Auerbach and Kotlikoff (1987). Starting with a guess for quantities and government policy, we compute prices, optimal household decisions, and value functions. Next we obtain the distribution of households on the state space and new macroeconomic quantities. We then update the initial guesses. These steps are iterated until the initial guesses and the resulting values for quantities, prices and public policy have sufficiently converged.

A.3 Computational efficiency

Our algorithm turns out to be quite efficient. It differs from the one used in Conesa et al. (2009) in that we first apply a more efficient interpolation routine and second allow any choice dimension (consumption, leisure, and assets) to indeed be continuous. Minor differences in simulation results are the consequence. Solving for a long-run equilibrium in the original model of Conesa et al. (2009) takes about 10 to 15 minutes, depending on the calibration.[14] Our simulation approach obtains the

[14] We simulate our models on a regular PC with a Intel® Core™ i7-870 Processor with 2.93 GHz and 8M Cache.

same results within 4 to 6 seconds! Computing a complete transition path with 320 transition periods takes about 40 minutes time.

References

Auerbach, A.J. and L.J. Kotlikoff (1987): *Dynamic fiscal policy*, Cambridge University Press, Cambridge.

Barro, R.J. (1974): Are government bonds net wealth? *Journal of Political Economy 82*, 1095-1118.

Bell, F.C. and M.L. Miller (2002): Life tables for the United States social security area 1900-2100, Office of the Chief Actuary, Social Security Administration, Actuarial Study 116.

Bernheim, D. (2002): Taxation and saving, in: A. Auerbach and M. Feldstein (eds.), Handbook of Public Economics, Vol. 3, Amsterdam, 117-1250.

Cagetti, M. (2001): Interest elasticity in a life-cycle model with precautionary savings, *American Economic Review 91(2)*, 418-421.

Cecchetti, S.G., P.S. Lam and N.C. Mark (2000): Asset pricing with distorted beliefs: Are equity returns too good to be true? *American Economic Review 90(4)*, 787-805.

Conesa, J.C. and D. Krueger (2006): On the optimal progressivity of the income tax code, *Journal of Monetary Economics 53(7)*, 1425-1450.

Conesa, J.C. S. Kitao, and D. Krueger (2009): Taxing capital? Not a bad idea after all! *American Economic Review 99(1)*, 25-48.

Diamond, P. and E. Saez (2011): The case for a progressive tax: From basic research to policy recommendations, *Journal of Economic Perspectives 25(4)*, 165-190.

Epstein, L.G. and S.E. Zin (1991): Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis, *Journal of Political Economy 99*, 263-286.

Erosa, A. and M. Gervais (2002): Optimal taxation in life-cycle economies, *Journal of Economic Theory 105*, 338-369.

Fehr, H. and C. Habermann (2008): Risk sharing and efficiency implications of progressive pension arrangements, *Scandinavian Journal of Economics 110(2)*, 419-443.

Fukushima, K. (2011): Quantifying the welfare gains from flexible dynamic income tax systems, Discussion Paper Series 176, Hitotsubashi University.

Gouveia, M. and R.P. Strauss (1994): Effective federal individual income tax functions: An exploratory empirical analysis, *National Tax Journal 47(2)*, 317-339.

Habermann, C. and F. Kindermann (2007): Multidimensional spline interpolation: Theory and applications, *Computational Economics 30(2)*, 153-169.

Hansen, G.D. (1993): The cyclical and secular behavior of the labour input: Comparing efficiency units and hours worked, *Journal of Applied Econometrics 8(1)*, 71-80.

Huang, H., S. İmrohoroğlu, and T. J. Sargent (1997): Two computations to fund social security, *Macroeconomic Dynamics 1*, 7-44.

İmrohoroğlu, S. (1998): A quantitative analysis of capital income taxation, *International Economic Review 39(2)*, 307-328.

Kitao, S. (2010): Labor-dependent income taxation, *Journal of Monetary Economics 57(8)*, 959-974.

Kumru, C. and J. Piggott (2012): Optimal capital income taxation with means-tested benefits, CEPAR Working Paper 2012/ 13.

Mankiw, N.G., M. Weinzierl and D. Yagan (2009): Optimal taxation in theory and practice, *Journal of Economic Perspectives 23(4)*, 147-174.

Nakajima, M. (2010): Optimal capital income taxation with housing, Working Paper No. 10-11, Federal Reserve Bank of Philadelphia.

Nishiyama, S. and K. Smetters (2005): Consumption taxes and economic efficiency with idiosyncratic wage shocks, *Journal of Political Economy 113*, 1088-1115.

Nishiyama, S. and K. Smetters (2007): Does social security privatization produce efficiency gains? *Quarterly Journal of Economics xx*, 1677-1719.

Press, W.H., S.A. Teukolsky, W.T. Vetterling and B.P. Flannery (2001): *Numerical recipes in Fortran: The art of scientific computing*, Cambridge University Press, Cambridge.

Støresletten, K., C. I. Telmer and A. Yaron (2004): Consumption and risk sharing over the life cycle, *Journal of Monetary Economics 51(3)*, 609-633.