

Taxing capital along the transition - Not a bad idea after all?

Hans Fehr*

*University of Würzburg
CESifo and Netspar*

Fabian Kindermann

*University of Bonn
and Netspar*

September 2014

Abstract

This paper quantitatively characterizes optimal tax systems in a model of overlapping generations, when transitional cohorts are explicitly taken into account. We use the recent study of Conesa et al. (2009) as an example, but extend it by transitional dynamics. We furthermore develop a general and coherent way of aggregating welfare effects of different individuals and cohorts in the short- and long-run. Our welfare measure includes the case of a utilitarian social welfare function, yet is not limited to this perspective.

We show that the optimality of a high capital income tax rate along the transition crucially depends on the assumption of a utilitarian social welfare function. This objective of the policy maker comprises implicit redistributive objectives across and within cohorts. Based on pure economic efficiency and insurance effects, however, we find a zero capital income tax rate and a less progressive labor income tax schedule to be optimal. Such a tax system receives political support from initial cohorts. A high capital income tax regime on the other hand doesn't.

JEL Classifications: C68, H21, D91

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

*Corresponding author:

Address: Department of Economics, University of Würzburg, Sanderring 2, D-97070 Würzburg, Phone: +49 931 31 82972
Email: hans.fehr@uni-wuerzburg.de

1 Introduction

Analyzing optimal tax policy in numerical studies has a long tradition in the macroeconomic and public finance literature. While several types of models can be used to do so, the special interest of this paper lies in overlapping generations models with households facing both borrowing constraints and uncertainty about future labor earnings. In order to quantitatively characterize the optimal capital and/or labor income tax structure in such a setting, most studies such as İmrohoroğlu (1998), Conesa and Krueger (2006), Conesa, Kitao and Krueger (2009), Kitao (2010) or Gervais (2012) exclusively focus on long-run welfare changes as a measure of optimality. They thereby completely neglect transitional cohorts who are also affected by a tax reform. Peterman (2013), however, shows that the tax policy that maximizes long-run welfare can induce huge welfare costs for generations living in the reform year. Yet, accounting for the welfare effects of both short- and long-run cohorts in one aggregate welfare measure is difficult. It requires both to calculate the full transition path of a tax reform, which is computationally demanding, and to consistently aggregate the welfare changes of different generations and individuals. The latter involves putting specific welfare weights on different households in the reform year and discounting the welfare effects of future cohorts. Of course, choosing welfare weights and social discount rates is somewhat arbitrary and therefore numerical studies that account for transitional dynamics apply a variety of welfare measures in order to derive optimal tax structures. Krueger and Ludwig (2013), for example, use a "utilitarian" welfare criterion that aggregates the utility of all individuals living in the reform year. Heathcote, Storesletten and Violante (2014, p.38) compare the optimal tax structures resulting from four different social welfare functions, which differ with respect to the weights put on agents within a cohort. Alternatively, Nishiyama and Smetters (2005) or Fehr, Kindermann and Kallweit (2013) derive optimal policy under a welfare criterion that uses individual-specific lump-sum redistribution payments in the spirit of Auerbach and Kotlikoff (1987).

In the present paper we want to contribute to this strand of literature along two dimension: First, we show how optimal tax results change when we account for the welfare effects of transitional generations. In order to address this issue, we use the same model and calibration as the recent study by Conesa, Kitao and Krueger (2009), but extend it by transitional dynamics. Second, we develop a general and coherent way of aggregating welfare effects of different individuals and cohorts in the short- and long-run. Our welfare measure includes the case of a utilitarian social welfare function, yet is not limited to this perspective. From this general measure we develop three specific ways to aggregate welfare, which all result in a different perspective of the social planner: a utilitarian, a cohort based and an individual based measure. We argue that the first and second measure comprise redistribution objectives within and/or across cohorts, while the individual based measure is a measure of pure economic efficiency. We finally show that these three perspectives of the policy maker have very different implications for optimal tax policy along the transition.

Specifically, we find that the high capital income tax result discussed by Conesa et al. (2009) can only be maintained, when the planner takes a utilitarian perspective. Economically this result is due to the fact that higher capital income taxation redistributes from current middle-aged generations towards young and future cohorts. The latter usually have a higher marginal utility of consumption. A utilitarian social planner positively values redistribution from generations with low to those with high marginal utility, so that a high capital income tax rate is beneficial from this perspective. The individual based aggregate welfare measure, however, abstracts from any redistributive objectives across and within cohorts and focuses solely on economic efficiency and insurance effects. On pure economic efficiency grounds, capital should not be taxed at all when transitional generations are taken into account. Capital income taxes increase welfare of young and future generations due to

the mimicking of age dependent labor tax rates and an improved insurance of labor productivity risk. For middle-aged and older cohorts, however, most of their labor productivity risk is already revealed and the number of remaining working periods is much smaller. Hence, they only slightly benefit from age dependent taxation and insurance provision. On the other hand, imposing a high capital income tax comes along with a severe savings distortion for these cohorts, which results in a huge welfare loss. These welfare losses outweigh gains of the young and future generations. As we also show, the resulting optimal tax structure would receive political support from transitional cohorts. A high capital income tax on the other hand doesn't.

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 A quick summary of the simulation model

We use exactly the simulation model and calibration of Conesa et al. (2009), yet extend it by transitional dynamics. In the following we therefore just quickly sketch the model in its original form. When we present results from simulations with transition path in the next section, we will further elaborate on what our assumptions for the transition are.

2.1 Model structure

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 constant 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 that are distributed in a lump-sum manner across the currently alive. Agents retire mandatorily at age j_r and start to receive social security payments which are financed by proportional payroll taxes.

Individuals enter the economy with zero assets $a_1 = 0$ and are not allowed to run into debt throughout their whole life, i.e. $a \geq 0$. During their working phase, they supply part of their maximum time endowment of one unit per period as labor l to the market. The remainder of time is consumed as leisure. Individual labor productivity is due to two types of shocks: a persistent component α_i that is realized at the beginning of the life cycle and a transitory component η that evolves as an autoregressive process over time. Consequently an individual is characterized by the state vector

$$(a, \eta, i, j) \in \mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \mathcal{J},$$

with $\phi_t(a, \eta, i, j)$ being the measure of households on the state space at time t .

Preferences over allocations of consumption c and and leisure $1 - l$ are assumed to be representable by a time-separable expected 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. Due to additive separability, we can formulate the individual optimization problem recursively:

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

where $Q(\eta, d\eta')$ defines the conditional distribution of tomorrow's labor productivity. The dynamic budget constraint reads

$$(1 + \tau_c)c + a' = [1 + r_t(1 - \tau_k)](a + Tr_t) + y + SS_t - \tau_{SS} \min\{y, \bar{y}\} - T_t(y_{\text{tax}}),$$

with labor income $y = w_t \epsilon_j \alpha_i \eta l$ and Tr_t being accidental bequests. Beneath productivity shocks, the hourly wage depends on an age specific productivity profile ϵ_j and the wage rate for effective labor w_t at time t . SS_t is a pension payment and $\tau_{SS} \min\{y, \bar{y}\}$ the contribution to social security. Households have to pay proportional taxes on consumption expenditure and asset income. In addition, taxable income y_{tax} is taxed according to a (potentially progressive) tax schedule $T_t(\cdot)$, see below.

Production takes place under perfect competition with a Cobb-Douglas production function. Firms employ capital and labor from households and produce a single output good. Capital depreciates over time at constant rate δ .

The government runs two separate closed-budget systems: social security and the tax system. The social security system collects all payroll contributions and distributes them in a lump-sum manner across retirees, so that the budget is balanced in every period. The social security tax rate is constant over time and not subject to the optimization of the policymaker. The tax system collects taxes on consumption expenditure and income in order to finance public consumption G_t which is exogenously given and held fix per capita. While the consumption tax rate τ_c and the capital income tax rate τ_k are constant over time, the tax schedule $T_t(\cdot)$ may vary. Taxable labor income consists of labor earnings net of employer contributions to the pension system $y_l = y - 0.5\tau_{SS} \min\{y, \bar{y}\}$. 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 $T_0(\cdot)$. The capital income tax rate is zero, i.e. $\tau_k = 0$. In a reform scenario, the policy maker can adjust both the progressive tax schedule $T_t(\cdot)$ as well as the capital income tax rate. Note that taxable income in a reform only consists of labor income, i.e. $y_{\text{tax}} = y_l$, since capital will be taxed at a constant rate.

2.2 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 $n = 0.011$. We assume standard Cobb-Douglas preferences

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

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 to generate a capital-output ratio of 2.7 and a share of hours worked in total time endowment of 0.33. The resulting interest rate r_0 is 5%.

The payroll tax rate τ_{SS} is 12.4 percent and the contribution ceiling \bar{y} amounts to 2.5 times average income. Government spending G amounts to 17 percent of GDP in the initial equilibrium. The consumption tax rate τ_c is set at 5 percent. Finally, the tax function is given by

$$T(y_{\text{tax}}) = T(y_{\text{tax}}, \kappa_0, \kappa_1, \kappa_2) = \kappa_0 [y_{\text{tax}} - (y_{\text{tax}}^{-\kappa_1} + \kappa_2)^{-1/\kappa_1}], \quad (2)$$

with parameters κ_0 , κ_1 and κ_2 . κ_0 controls the level of the top marginal tax rate and κ_1 determines the progressivity of the tax code. κ_2 shifts the point at which the top marginal tax rate starts applying and is used to balance the tax system's budget. 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.

In the reform scenario, the policy maker can choose a new κ_0 and κ_1 , whereas κ_2 is again defined by budget balance. We yet extend the functional choice set in two dimensions so that

$$T(y_{\text{tax}}) = \begin{cases} \kappa_0 \cdot y_{\text{tax}} & \text{for } \kappa_1 = 0 \\ \kappa_0 \max[y_{\text{tax}} - \kappa_2, 0] & \text{for } \kappa_1 = \infty, \end{cases} \quad (3)$$

i.e. in the first case we have a purely proportional tax and in the case of $\kappa_1 = \infty$ a flat tax with a deduction of κ_2 .¹ Furthermore the policy maker can choose a capital income tax rate τ_k , so that an optimal policy in our model consists of four parameters in total $\Psi = (\kappa_0, \kappa_1, \kappa_2, \tau_k)$.

3 Simulation results

This section deals with optimal tax calculations. We will present our simulation results in several steps. The first is to follow Conesa et al. (2009) and base optimal tax calculations on long-run welfare considerations only. In a next step we focus on transitional dynamics. We start with a welfare function that focuses on short-run generations only, meaning generations that were already living in the initial equilibrium and that are hit by a tax reform at some stage during their life cycle. In order to aggregate short-run welfare effects we develop a class of welfare measures that also include the social welfare function of Conesa et al. (2009) and examine the impact of these different welfare measures on optimal tax outcomes. Finally we quantify optimal tax policy when all current and future generations that are affected by a tax reform are taken into account.

3.1 Optimal tax policy in the long-run

In this section we first want to replicate the results of Conesa et al. (2009). The thought experiment therefore is fairly simple. We just compute a new long-run equilibrium that results from changing the parameters of the tax schedule $(\kappa_0, \kappa_1, \kappa_2, \tau_k)$, keeping the structural and exogenous government parameters at initial equilibrium levels. 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 long-run allocation $(c^*, 1 - l^*)$ can be computed from

$$CEV = \left[\frac{W(c^*, 1 - l^*)}{W(c_0, 1 - l_0)} \right]^{\frac{1}{\gamma(1-\sigma)}} - 1,$$

where $W(c, 1 - l)$ is ex ante expected lifetime utility. CEV 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.

Pure long-run simulations Taking this long-run welfare criterion 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 roughly corresponds to $\kappa_1 \approx 7$ and $\kappa_2 = 30262$) to be the optimal choice of the policy maker. The first column of Table 1 reports the resulting long-run changes in aggregate variables.² Yet, Conesa et al. (2009) restrict their

¹ In the case of a proportional tax we use κ_0 to balance the government's budget. Note that the proportional tax is exactly the limiting case of (2) for $\kappa_1 \rightarrow 0$.

² The figures correspond to the ones in Table 2 of Conesa et al. (2009, p. 36). Differences in values arise from a different computational method, specifically from the fact that in our simulation approach labor supply is a continuous choice variable, see Appendix C. All appendices are available online.

parameter choice set to finite values for κ_1 . However, the special case of a flat tax ($\kappa_1 = \infty$) turns out to be the optimal one in terms of long-run welfare maximization. The resulting tax schedule as well as long-run macroeconomic 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 1.50 to 1.66 percent.

Table 1: Optimal tax schemes: Long-run welfare comparison

	Conesa et al. (2009)	optimal scheme	alternative closure
τ_k	0.36	0.43	0.43
κ_0	0.23	0.20	0.20
κ_1	7	∞	∞
κ_2	30262	11882	11975
Average hours worked	-0.63	0.67	0.71
Total labor supply N	-0.13	1.21	1.23
Capital stock K	-6.42	-8.14	-8.27
Government debt to GDP (in %)	0.00	0.00	-0.92
Output Y	-2.44	-2.26	-2.17
Aggregate consumption C	-1.41	-0.34	-0.27
Long-run CEV	1.50	1.66	1.75

Macro figures as changes in percent over initial equilibrium values.

Welfare figures as percentage of initial household consumption.

Transitional dynamics and closure rules The calculations above were based on pure long-run simulations, meaning that it was only necessary to compute an initial and a new long-run equilibrium. When we in addition calculate the transition path in between these two equilibria, assumptions about tax code adjustments and fiscal closure rules during the transition will come into play. First of all, we assume that the parameters that determine the shape of the tax code (κ_0, κ_1) as well as the tax rate on capital income τ_k are only changed once in the first year of the transition $t = 1$ and stay constant afterwards. Note that households in the initial equilibrium will not anticipate this change in tax rates and therefore will be surprised by the tax reform at some stage during their life-cycle. In analogue to the long-run simulations, κ_2 will be determined in order to balance the governments budget. We can imagine two ways in which the government might do this. The most straightforward way would be to adjust κ_2 in every period in order to clear the budget periodically. This would result in a sequence $\{\kappa_{2,1}, \kappa_{2,2}, \dots, \kappa_{2,\infty}\}$ of budget balancing kappas, where $\kappa_{2,\infty}$ will be the same as in the previous simulations. Obviously in this scenario, the new optimal long-run equilibrium would look exactly the same as in column 2 of Table 1. An alternative (and probably more realistic) way of clearing the government's budget would be to calculate a time-invariant κ_2 that satisfies

$$\sum_{t=1}^{\infty} G_t R_t = \sum_{t=1}^{\infty} \left[\tau_c C_t + \int [\tau_k y_r + T(y_{\text{tax}})] \phi_t (da \times d\eta \times di \times dj) \right] R_t,$$

meaning that it balances the intertemporal budget of the government. $R_t = \prod_{s=2}^t (1 + r_s)^{-1}$ thereby defines the intertemporal discount factor. Short-run fluctuations in tax revenue could then be bal-

anced by adjustments in government debt. Consequently, we would calculate the path for debt B_{t+1} according to the periodical budget constraint

$$G_t + (1 + r_t)B_t = \tau_c C_t + \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \mathcal{J}} [\tau_k y_r + T(y_{\text{tax}})] \phi_t(da \times d\eta \times di \times dj) + B_{t+1}.$$

When we assume such a fiscal closure rule, adjustments in government debt throughout the transition will certainly have an effect on the new long-run equilibrium. As the third column in Table 1 shows, these effects are however rather modest. When we assume debt to balance the government's budget in the short-run, the government will be able to build up some assets. The stock of assets amounts to roughly 1% of GDP in the long-run. The reason for this is that private assets decline throughout the transition due to the high tax rate on capital income. As a result, tax revenue is higher in the short than in the long-run. The effect on macroeconomic aggregates as well as long-run welfare are however small. In the following, we will therefore only simulate tax reforms in which government debt balances the government budget in the short-run and neglect the case of periodical adjustments of the tax code.

3.2 Optimal tax policy in the short-run

After studying optimal tax policy in the long-run, we now want to discuss how short-run generations are affected by certain tax reforms. Peterman (2013) already shows that the policy which is optimal for long-run cohorts hurts many generations in the short run. They have built up assets in the past which are now taxed more heavily. In contrast to Peterman (2013) we now want to study what the optimal tax policy would be, if only the short-run generations were taken into account. We therefore neglect the welfare effects of all newborn cohorts throughout the transition, but only focus on those generations that were already alive in the initial equilibrium. A synthesis of short-run and long-run welfare effects will follow in the next section. The problem at hand is that for any future generation there is a clear measure of welfare, namely *ex ante* expected lifetime utility as we used it in the previous section. For current generations the case is a little different, since part of their life cycle has already passed. Consequently, there is a distribution of households over the state space $\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}$ with (potentially) very different welfare effects. Yet, it is not so straightforward to come up with a useful and coherent way of aggregating these individual welfare effects to an aggregate welfare measure. In general there are almost infinite ways to characterize aggregate welfare for these cohorts. One typical approach in the literature that is applied e.g. by Conesa et al. (2009) or Heathcote et al. (2014) puts equal weight on the utilities of all agents currently alive and just weighs them according to their share in the total population, i.e.

$$\int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} v_1(a, \eta, j, i) \phi_1(da \times d\eta \times di \times dj).$$

By maximizing this social welfare function over the relevant policy space $\Psi = \{\kappa_0, \kappa_1, \kappa_2, \tau_k\}$ under a closed budget constraint, Conesa et al. (2009) claim that the capital income tax rate τ_k should be even higher than in the new long-run equilibrium.

3.2.1 An aggregate welfare measure for policy evaluation

In the following we derive a general and coherent aggregate welfare measure which will include the case of Conesa et al. (2009), but is not limited to their perspective. Lets assume \mathcal{X} was an arbitrary partition of the state space $\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}$, i.e. \mathcal{X} is a set of non-empty and disjoint subsets

of the state space, such that $\cup_{X \in \mathcal{X}} X = \mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}$.³ For each of the subsets X we define aggregate utility at time t as

$$\int_X v_t(a, \eta, i, j) \phi_t(da \times d\eta \times di \times dj).$$

As already mentioned above the key ingredient in our welfare measure will now be (hypothetical) lump-sum transfers. Specifically lets assume that each member of the subset X received a lump-sum transfer $T(X)$. We can then compute the aggregate utility level of the subset X depending on the amount of lump-sum transfer $T(X)$ as well as a specific tax system Ψ as

$$V_t[X, T(X), \Psi] = \int_X v_t(a + T(X), \eta, i, j) \phi_t(da \times d\eta \times di \times dj).$$

To determine our aggregate welfare measure we finally do the following thought experiment: We know that before a reform has taken place, subgroup X enjoyed aggregate utility $V_0(X, 0, \Psi_0)$ in the initial equilibrium, whereas after the reform their aggregate utility is $V_1(X, 0, \Psi_1)$.⁴ We now want to compute the amount of lump-sum transfer $T(X)$ we would have to give to each member of subgroup X in order to guarantee the same aggregate utility level after the reform as in the initial equilibrium. We can obviously do that by determining the solution to the equation

$$V_1(X, T(X), \Psi_1) = V_0(X, 0, \Psi_0). \quad (4)$$

The transfer level $T(X)$ then indicates the size of the aggregate welfare change of the subgroup X of the total population in dollar units.

In order to derive an aggregate measure of the welfare change for the total population from these transfers, we now just have to aggregate the negative of the individual transfer payments of all subgroups. Specifically we compute

$$\Delta V(\mathcal{X}, \Psi) = - \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} T(X) \phi_1(da \times d\eta \times di \times dj).$$

To finally get to a consumption equivalent variation interpretation, we just divide $\Delta V(\mathcal{X}, \Psi)$ by aggregate consumption in the initial equilibrium. We therefore can interpret

$$\varphi(\mathcal{X}, \Psi) = \frac{\Delta V(\mathcal{X}, \Psi)}{C_0}$$

as the aggregate welfare effect of a tax reform for the short-run cohorts measured in consumption units.

3.2.2 Some intuition

But does it make sense to think about welfare changes of a reform like this? And how does the choice of a partition \mathcal{X} affect aggregate welfare numbers? To shed light on these questions and give some intuition on the functionality of our welfare measure, we can use a simple first order approximation of the value function. Specifically, we can approximate the utility of a household characterized by

³ It is straightforward to generalize our welfare measure to partitions of a subset of the state space, therefore e.g. taking a Rawlsian perspective. Yet, to keep things as simple as possible, we leave this generalization to the interested reader.

⁴ Ψ_0 thereby characterizes the initial equilibrium tax system, while Ψ_1 is the reform tax system.

the state $(a, \eta, i, j) \in \mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}$ at time $t = 1$ after having received a lump-sum transfer amount T as

$$v_1(a + T, \eta, i, j) \approx v_1(a, \eta, i, j) + \frac{\partial v(a + T, \eta, i, j)}{\partial T} \cdot T = v_1(a, \eta, i, j) + u_c(c, 1 - l) \cdot T,$$

where the last identity is obtained from the envelope theorem. Now let's define

$$\lambda(X) = \frac{\int_X \phi_1(da \times d\eta \times di \times dj)}{\int_X u_c(c, 1 - l) \phi_1(da \times d\eta \times di \times dj)}$$

which is the inverse of average marginal utility of consumption of the respective subgroup X .⁵ We can also interpret $\lambda(X)$ as the average value of an additional unit of utility for this subgroup X in dollar units. With this definition as well as the above approximation, we quickly obtain

$$V_1[X, T(X), \Psi] \approx \int_X v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj) + \frac{T(X)}{\lambda(X)} \cdot \int_X \phi_1(da \times d\eta \times di \times dj). \quad (5)$$

Using this relation we find that the transfer $T(X)$ that solves equation (4) is approximately equal to

$$T(X) \approx -\lambda(X) \cdot \frac{\int_X \Delta v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj)}{\int_X \phi_1(da \times d\eta \times di \times dj)} \quad \text{with } \Delta v_1(\cdot) = v_1(\cdot) - v_0(\cdot).$$

Note that we used the fact that the distribution of individuals over the state space in period $t = 1$ is already determined in period $t = 0$ and therefore we have $\phi_0 = \phi_1$. Not surprisingly the amount of lump-sum transfer that is needed to bring the entire group X back to their initial equilibrium level in terms of aggregate utility is the average change in utility weighted with the respective price $\lambda(X)$. Since $\lambda(X) > 0$, $T(X)$ is negative in the case this subgroup's aggregate utility increases from the tax reform. Furthermore, the smaller $T(X)$, the larger is the average utility gain of subgroup X . Finally, we can approximate the aggregate welfare change as

$$\begin{aligned} \Delta V(\mathcal{X}, \Psi) &= - \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} T(X) \phi_1(da \times d\eta \times di \times dj) \\ &\approx \int_{X \in \mathcal{X}} \lambda(X) \cdot \int_X \Delta v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj) \times \phi_1(da \times d\eta \times di \times dj) \\ &= \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} \lambda(X) \cdot \Delta v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj). \end{aligned}$$

Our aggregate measure of welfare therefore approximately consists of the aggregated utility differences between initial equilibrium and reform year, weighted with some social welfare weights $\lambda(X)$. By choosing a certain partition of the state space, the welfare weights are endogenously determined as the average marginal utility of consumption of the respective subgroup.⁶

In the following we will consider three different partitions of the state space that lead to three different aggregate welfare functions.⁷ When we assume that the only set in \mathcal{X} is the whole state space

⁵ There are two things to note here. First a problem with the definition of λ might occur when $X = \{(a, \eta, i, j)\}$ and $\int_X \phi_1(da \times d\eta \times di \times dj) = 0$. Yet, if ϕ_1 is continuous around X then L'Hôpital's rule tells us that $\lim_{Y \rightarrow X} \lambda(Y) = \frac{1}{u_c(c, 1 - l)}$. Second, if ϕ_1 is not continuous or X is some "larger" set of measure zero, then $\lambda(X)$ is not properly defined. Since those sets do not play a role in our aggregate welfare measure anyway and we use the above approximation for purely illustrative reasons, we can ignore this problem and set e.g. $\lambda(X) = 1$ in that case.

⁶ Heathcote et al. (2014, p.36f.) apply a very similar idea in order to construct alternative welfare measures.

⁷ Appendix A shows in very analytical detail how to partition the state space in order to receive these welfare functions. Here we just rely on some intuitive descriptions.

itself, we obtain the *utilitarian welfare function*. The respective welfare weight λ^{util} is then the inverse of average marginal utility of consumption of the whole short-run population. The aggregate welfare change then becomes

$$\Delta V^{\text{util}}(\Psi) = \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} \lambda^{\text{util}} \cdot \Delta v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj).^8$$

In the second example we partition the state space cohort by cohort so that the cohort welfare weight $\lambda^{\text{coh}}(j)$ denotes the inverse of the average marginal utility of consumption of the considered cohort j . The corresponding *cohort based welfare function* is

$$\Delta V^{\text{coh}}(\Psi) = \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} \lambda^{\text{coh}}(j) \cdot \Delta v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj).$$

In the last example we consider a partition in which each individual forms its own subset so that the individual welfare weight is $\lambda^{\text{ind}}(a, \eta, i, j) = u_c(c, 1 - l)^{-1}$. The corresponding aggregate individual based welfare function then reads

$$\Delta V^{\text{ind}}(\Psi) = \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} \lambda^{\text{ind}}(a, \eta, i, j) \cdot \Delta v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj).$$

The three aggregate welfare measures introduced above differ with respect to the weights that are attached to the welfare changes $\Delta v_1(\cdot)$. Figure 1 shows the weights for each generation under the three different welfare measures.⁹ On the x-axis of the graph, we see the age of a generation in the reform period. Consequently, cohorts on the left end of the graph are the oldest, while those on the right end are the youngest. The utilitarian weight λ^{util} obviously is constant over cohorts. In order to understand the evolution of the cohort based welfare weights $\lambda^{\text{coh}}(j)$, we have to take a deeper look at the cohort-specific marginal utilities of consumption. Given the first order condition of the household

$$u_c(c, 1 - l) = \beta \psi_{j+1} (1 + r) E [u_c(c', 1 - l')],$$

the consumption path and therefore marginal utility will in general not be constant over the life cycle. Specifically, at young ages we have $\beta \psi_{j+1} (1 + r) > 1$, and therefore marginal utility of consumption will decline with age.¹⁰ Only when survival probabilities become small enough, marginal utility of consumption will increase again. The cohort based welfare weights neutralize such differences in marginal utility of consumption across cohorts by putting a higher weight on cohorts with a low (average) marginal utility. Consequently, in Figure 1 cohort based welfare weights initially increase with age as the average cohort consumption increases and marginal utility of consumption declines. At retirement, there is an upward jump discontinuity. The reason is that leisure consumption jumps up to a value of 1 at this stage therefore reducing the marginal utility of consumption significantly. Finally,

⁸ Note that the welfare change of each individual on the state space gets the same weight in the social welfare function. Consequently, maximizing $\Delta V^{\text{util}}(\Psi)$ is equivalent to maximizing

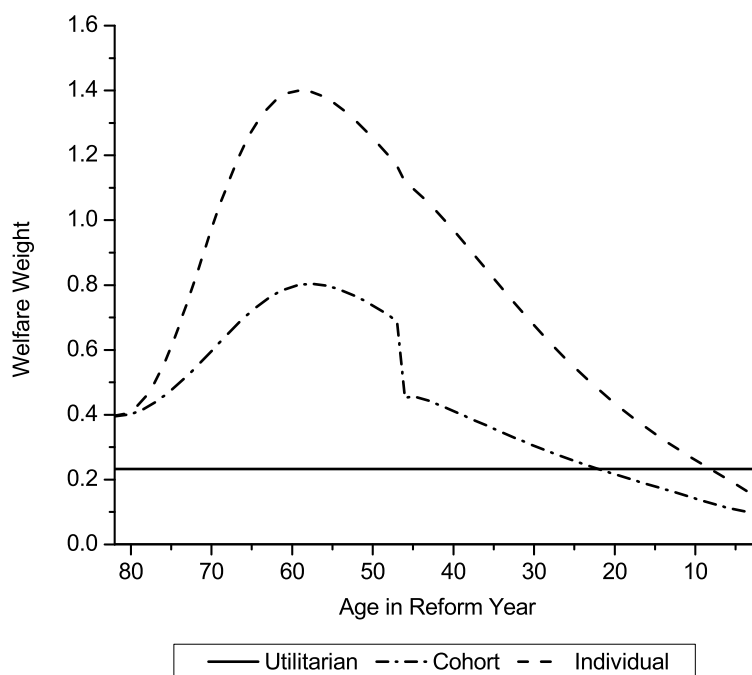
$$\int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj),$$

which is the same utilitarian welfare criterion as in Conesa et al. (2009).

⁹ In case of the individual measure we used the average welfare weight for each generation that we computed from the individual weights λ^{ind} as well as the distribution of households ϕ_1 .

¹⁰ Note that the above equality only holds for individuals that are currently not hit by the borrowing constraint. Yet individuals at the borrowing constraint (who are usually young) do also have a relatively high marginal utility of consumption so that it is very likely that their marginal utility of consumption will decline with age as well.

Figure 1: Welfare weights across generations



as survival probabilities decline towards the end of the life cycle, marginal utility of consumption increases again and the welfare weight of the cohort declines.

The same rationale applies to the individual welfare weights. Within a cohort, individuals with different wealth levels and productivity shocks have different marginal utilities of consumption. Individual welfare weights neutralize these differences in marginal utilities. Consequently, the weight increases with individual asset holdings, since richer households have a lower marginal utility of consumption. Appendix B shows that the difference in welfare weights within a cohort becomes larger and larger with increasing age, since wealth is much more important for individual consumption (in contrast to labor earnings). Figure 1 reveals that across cohorts individual weights have a similar shape as the cohort based weights. Yet, since the individual weight is an aggregate of inverse marginal utilities of consumption, it differs in size from the cohort weight.

But what is the economic implication of these different welfare weights? Since the utilitarian welfare weight does not at all neutralize changes in marginal utilities of consumption, it will certainly favor policies that increase consumption of young or very old individuals, since their marginal utility of consumption is relatively high. Alternatively one can say that the utilitarian measure includes an intergenerational redistributive objective away from middle aged cohorts towards young and very old generations. The cohort based measure, however, is neutral with respect to intergenerational redistribution. Yet, since an identical welfare weight is given to each member of a generation, the cohort based welfare measure (and therefore also the utilitarian one) still favors policies that redistribute from rich to poor households within a cohort. The individual welfare measure finally neutralizes all differences in marginal utilities of consumption within and across cohorts. Hence, it is neutral with respect to both inter- and intragenerational redistribution. The only reason why this measure would then favor a redistributive tax policy would be that it provides insurance against productivity risk in future periods.

Table 2 summarize the impact of different features of tax policy on our three different welfare measures. Certainly, everything that improves economic efficiency – e.g. reduced labor supply or savings

distortions or improved insurance provision against labor productivity risk – will positively influence aggregate welfare for all three welfare measures. Yet, redistribution across cohorts¹¹ will only have a positive aggregate welfare impact in the utilitarian measure. Furthermore, intragenerational redistribution will be beneficial under both the utilitarian and the cohort based welfare measure. Since the individual welfare measure is neutral with respect to both inter- and intragenerational redistribution, we can interpret it as a measure of pure economic efficiency. The other two measures combine economic efficiency with some implicit redistributive objectives across and/or within cohorts.

Table 2: Impact of tax policy on different welfare measures

	Utilitarian	Cohort based	Individual
Economic efficiency	+	+	+
Redistribution within a cohort	+	+	0
Redistribution across cohorts	+	0	0

There is one final thing to note here. The sum of individual welfare changes between the initial equilibrium and the reform period depends heavily on the initial equilibrium resource allocation. This is because asset holdings and therefore the distribution of individuals over the state space in $t = 1$ is already determined through individual behavior in $t = 0$. When applying utilitarian welfare weights, the optimal tax system consequently is determined not only by the physical environment and market structure of the economy but also by the status quo tax system. Using the cohort and especially the individual specific welfare weights, we account for differences in marginal utilities of individuals that may be caused by initial equilibrium government policies. On the one hand one might then argue that we depart from the Bergson-Samuelson tradition in welfare economics in that the weight given to a cohort or individual is not structural. On the other hand by doing so, we neutralize the effects of initial equilibrium policies which is perfectly in line with the Lucas critique.

3.2.3 Consequences for optimal tax policy

With the knowledge about our welfare measures at hand, we can now determine optimal tax policies in terms of short-run aggregate welfare under our three different welfare measures. Table 3 shows the respective results.¹² Not surprisingly, the tax system that maximizes the utilitarian welfare criterion favors a large tax rate on capital income of $\tau_k = 0.49$, similarly to what has been found by Conesa et al. (2009). This tax mainly hurts the middle-aged generations. Younger generations will however benefit from such a tax system through lower tax rates on labor income and a higher basic allowance. Yet, the other two welfare measures report an aggregate welfare loss for this high capital income tax rate. In order to understand why, we have to look at the welfare effects of different cohorts and how they enter the aggregate welfare criterion. In order to do so, we decompose the aggregate welfare numbers into cohort contributions by calculating

$$\frac{1}{C_0} \cdot \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{j\}} \lambda(X) \cdot \Delta v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di)$$

¹¹ By redistribution we always mean redistribution from individuals with low to individuals with high marginal utility of consumption.

¹² Be reminded that we compute transfer payments by actually solving (4) through numerical simulation and not by the approximation that was used for pure illustration.

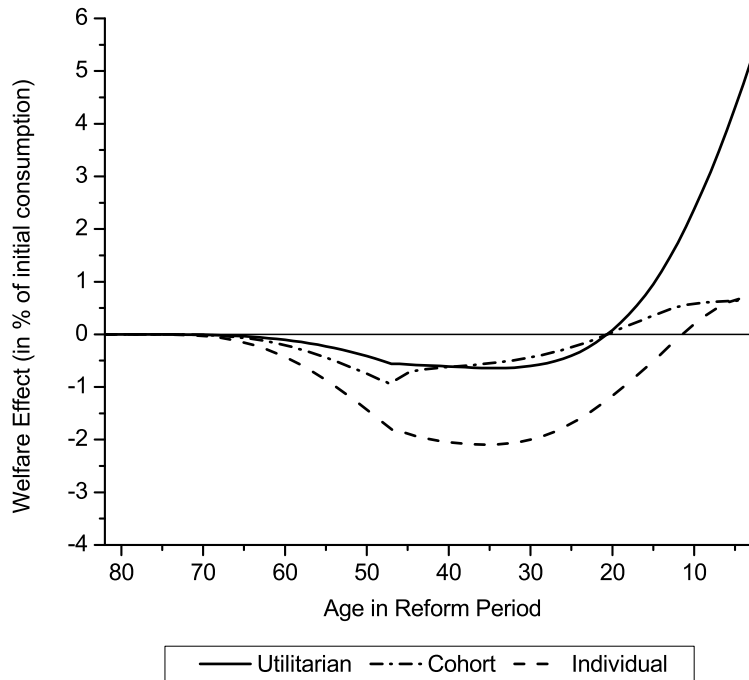
Table 3: Optimal tax schemes: Short-run welfare comparison

	Utilitarian	Cohort based	Individual
τ_k	0.49	0.21	0.00
κ_0	0.18	0.20	0.23
κ_1	∞	∞	0
κ_2	12853	3020	-
Utilitarian	10.45	-2.21	-17.04
Cohort based	-12.46	0.61	-2.62
Individual	-67.38	12.20	40.94
Political support	36.04	60.49	65.66

Welfare figures as percentage of initial aggregate consumption

for all three different welfare measures. Figure 2 shows the respective results. The aggregate welfare numbers reported in the first column of Table 3 can be obtained from this graph by just adding up the cohort specific welfare numbers.

Figure 2: Welfare contributions of each cohort ($\kappa_0 = 0.18, \tau_k = 0.49$)



First of all we see that the welfare changes of very old generations do basically not play a role in all of the three aggregate welfare measures. This is because (i) due to survival risk there is only a fairly small amount of individuals that reach this age and (ii) at very old age households have already decumulated all their assets and solely rely on pension payments that are not taxed under the income tax scheme. The solid line in the graph represents the cohort specific welfare effects under the utilitarian welfare measure. Since the welfare weight is constant across generations, this

line basically shows the change in aggregate utility of each generation weighted by the generation's size. As argued above, a tax reform with high capital income tax rates and low labor tax rates favors the younger generations and burdens the middle-aged. Yet, since young generations have a much larger marginal utility of consumption, redistributing towards these cohorts is beneficial under the utilitarian welfare criterion. Therefore the positive contribution to aggregate welfare of the young cohorts is remarkably large.

This changes significantly when we look at the cohort based welfare measure. In this measure intergenerational redistribution is completely neutral in terms of aggregate welfare. Consequently, that part of the positive welfare contribution from young cohorts that was solely due to the fact that they have a high marginal utility of consumption disappears. Yet, young cohorts still have a positive welfare effect, since the same argumentation as for the long-run cohorts applies here. Consequently, young households benefit from capital income taxation mimicking an age dependent labor tax scheme and from the increased insurance provision against productivity risk, see Conesa et al. (2009). Note that the intersection with the abscissa, i.e. the point at which welfare effects turn negative, needs to be the same for both welfare measures.

When we finally turn to the welfare effects under the individual welfare measure, we find that they turn much more negative especially for middle aged cohorts. The reason for this is fairly simple. Both the utilitarian and the cohort based welfare measure favor redistribution within a cohort from the rich towards the poor. A positive capital income tax does exactly achieve this goal. Yet, when we control for different marginal utilities within a cohort – or put differently allow for individual specific lump-sum transfers to undo intragenerational redistribution – this redistributive goal is eliminated. What remains is a huge distortion on savings of the middle-aged generations induced by a large tax on capital income, which has a clear negative effect on economic efficiency and therefore aggregate welfare. For the young cohorts, who do not have any significant wealth holdings yet, the same reasoning as under the cohort based measure applies. In consequence, for these cohorts the welfare effects are fairly similar for the cohort and the individual measure.

Summing up, the large welfare gains resulting from high capital income taxes under the utilitarian welfare criterion are mostly due to intergenerational redistribution. The reform redistributes from middle-aged cohorts towards young generations. Since the latter have the higher marginal utility of consumption, aggregate utilitarian welfare increases. Under the cohort based measure intergenerational redistribution is welfare neutral. Yet, capital income taxation still redistributes a lot within cohorts. Consequently, aggregate welfare is still higher under this scheme than under the individual welfare criterion. Since this last measure of welfare is purely efficiency based, we can conclude that high capital income taxes in the short-run are bad in terms of aggregate efficiency.

Before proceeding with the optimal tax systems under the other welfare measures, we want to analyze whether implementing a tax reform with high tax rates on capital income would be politically feasible. In order to do so, we calculate a simple measure of political support. Specifically, we assume that each individual in the initial equilibrium was to vote one time for a new tax system which would then immediately be implemented. We determine their voting behavior by comparing utility before and after the tax reform and assume a household votes in favor of the reform, if he receives a utility gain. Consequently, the share of households that would vote in favor of a tax reform can be calculated from

$$\frac{\int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} 1 \left[\Delta v_1(a, \eta, i, j) \geq 0 \right] \phi_0(da \times d\eta \times di \times dj)}{\int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} \phi_0(da \times d\eta \times di \times dj)}$$

The optimal tax reform under the utilitarian welfare measure does not receive a lot of political sup-

port. Since political support is based on an individual welfare comparison and the individual welfare measure reports a clear negative welfare effect, this outcome seems fairly natural. Clearly, imposing a huge capital income tax rate favors the few young generations with high marginal utility of consumption, but as we have seen in Figure 2, almost any other generation loses.

Finally, we also want to compute optimal tax schemes under the other welfare criteria. When optimizing the cohort based welfare criterion (see second column of Table 3), we find a much lower capital income tax rate to be optimal than under the utilitarian criterion. Note that the capital income tax rate is still significantly positive and the marginal tax rate on labor income increases compared to the previous scenario. This is due to the fact that the cohort based criterion still favors intragenerational redistribution, which can be achieved by both positive tax rates on capital income and a more progressive tax schedule. Since in this case the individual based welfare measure reports a positive value, it is not very surprising that the reform option receives much more political support. In fact, a majority of households would vote for such a tax reform proposal.

Last but not least, in terms of pure efficiency a zero capital income tax turns out to be the optimal choice in the short-run. Furthermore, the optimal marginal labor income tax rate slightly increases again to 0.23 so that the labor income tax schedule is perfectly proportional. The reason for this is straightforward. For the young generations progressive labor income taxes paired with high capital income taxes are beneficial due to the mimicking of an age dependent income tax system and an increased insurance provision against labor productivity risk. For the middle-aged generations, however, a large part of labor productivity risk is already revealed and the time to retirement is fairly short. Consequently, these generations don't benefit from the positive aspects of a high capital income tax, but only are left over with a severe savings distortion for the remainder of their life. In terms of aggregate efficiency, these negative effects for older cohorts clearly overcompensate the positive effects for young cohorts and therefore a zero capital income tax is to be favored in the short-run. Appendix B shows the cohort contributions to aggregate welfare for this tax reform under the three welfare measures.

Summing up, a high capital income tax rate of $\tau_k = 0.49$ might be optimal in the short-run, but only if one's objective is to redistribute away from generations with low to those with high marginal utilities of consumption. When the redistribution objective is only to redistribute within generations, the capital income tax rate should be significantly lower at $\tau_k = 0.21$, but the progressivity of the labor income tax schedule should be higher. Finally, on pure economic efficiency grounds, capital income should not be taxed at all in the short-run. The reason is that the efficiency benefits of positive capital income taxes and high tax progressivity can only be enjoyed by a few young generations. For the middle-aged and old generations, however, higher capital income taxes only come along with severe savings distortions.

3.3 Optimal taxation with current and future cohorts

Having learned about optimal tax policy in the short and the long-run, one might finally ask what an optimal tax schedule looks like that takes into account welfare effects of all generations. The question coming up then is how to compare welfare effects across generations. For the short-run generations we have previously seen that there are ample ways of aggregating individual welfare effects. For future cohorts – i.e. those cohorts who enter the economy along the transition path – we can apply the same method. Yet, to stay consistent with our long-run analysis, we will always evaluate welfare of future cohorts from an ex ante perspective meaning behind the Rawlsian veil of

ignorance.¹³ Hence, the relevant state space that includes all generations affected by a reform is given by

$$\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\} \cup \{1, 2, 3, \dots\}.$$

While the first part of the state space comprises all current generations, i.e. those who have been economically active already in the initial equilibrium and were surprised by a tax reform at some point during their life cycle. The second part contains all future cohorts, i.e. those who enter the economy at some point along the transition path. In analogy to the previous analysis we can derive an aggregate welfare measure

$$\begin{aligned} \Delta V(\mathcal{X}, \Psi) = & \int_{\mathcal{A} \times \mathcal{E} \times \mathcal{I} \times \{2, \dots, J\}} \lambda(X) \cdot \Delta v_1(a, \eta, i, j) \phi_1(da \times d\eta \times di \times dj) \\ & + \sum_{t=1}^{\infty} \theta^{t-1} \cdot \lambda(X) \cdot \int_{\mathcal{E} \times \mathcal{I}} \Delta v_t(0, \eta, i, 1) \phi_t(\{0\} \times d\eta \times di \times \{1\}), \end{aligned}$$

where θ defines the discount factor for the welfare change of future cohorts, see Appendix A for computational details. In principle we could use any value for the discount factor θ . Setting $\theta = 0$ we would only focus on the welfare changes of current cohorts, while for $\theta \rightarrow (1+n)^{-1}$, only the welfare effects of the long-run generations would be accounted for, see Fukushima (2011). The results of these choices have already been analyzed in the previous sections. In the following, we will set $\theta = \frac{1}{1+r_0}$, with r_0 being the capital market interest rate in the initial equilibrium. This choice of discount factor has mainly two advantages: First, it lies somewhere in between the two extremes described above and second, it does not depend on the specific type of reform we are doing. Note that since welfare of future generations is evaluated from an ex ante perspective, under the cohort and individual welfare approach there will be only one particular welfare weight $\lambda(X) = \lambda_t$ for each future generation t .

In order to again obtain a consumption based interpretation of the measure, we follow Huang et al. (1997) and turn $\Delta V(\mathcal{X}, \Psi)$ into an annuity that pays out in every period of the transition as well as in the new long-run equilibrium. We relate this to aggregate consumption in the initial equilibrium. Hence, we have

$$\varphi^{\text{total}}(\mathcal{X}, \Psi) = \frac{\Delta V(\mathcal{X}, \Psi)}{C_0} \cdot \frac{r_0 - n}{1 + r_0}.$$

The interpretation of φ^{total} is fairly straightforward. A value of 1% for φ^{total} means that the aggregate welfare effect of a tax reform is equivalent to a 1% increase in consumption in every period along the transition and in the new long-run equilibrium.

Table 4 shows the optimal tax systems that result when we maximize aggregate welfare effects of all cohorts. We thereby again use the three different aggregate welfare measures which were discussed in the previous subsection but now include current and future cohorts, see Appendix A. The results follow the same logic as in pure short-run analysis in Table 3. Again, under the utilitarian welfare criterion capital income should be taxed at a high rate, since high capital income taxes redistribute from the short-run middle-aged cohorts to the young and very old generations. A high capital income tax rate is furthermore beneficial for future cohorts, so that we find the optimal tax rate to be 0.49. In case of the cohort based welfare measure a tax rate of 0.21 was optimal in the short-run while in the long-run the optimum was at 0.43. The reported value of 0.34 in Table 4 lies somewhat in between these two. Finally, in terms of pure economic efficiency, the huge negative distortions on the currently

¹³ An alternative would be to evaluate welfare of future generations ex post, i.e. after they got to know their realization of the fixed effect i .

Table 4: Optimal tax schemes: Current and future cohorts

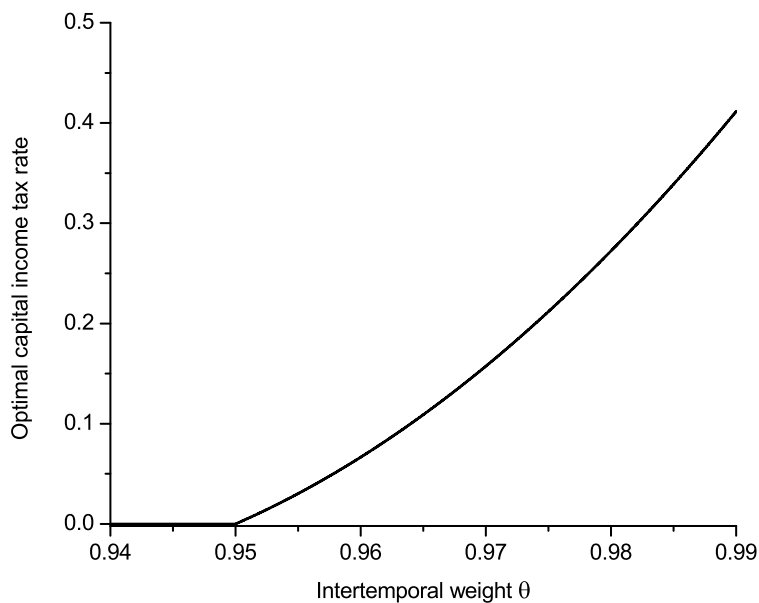
	Utilitarian	Cohort based	Individual
τ_k	0.49	0.34	0.01
κ_0	0.20	0.20	0.23
κ_1	∞	∞	0
κ_2	14980	8124	-
Utilitarian	1.17	0.75	-1.61
Cohort based	-0.08	0.23	-0.87
Individual	-2.40	-0.60	0.72
Political support	32.85	43.21	65.93

Welfare figures as percentage of initial aggregate consumption

living middle-aged generations imposed by a positive tax rate on capital can not be outweighed by future welfare gains. In consequence the optimal capital income tax rate is almost zero, even when future generations are taken into account. Appendix B disentangles the welfare effects by cohort.

The question that remains is how the optimality of very low capital income taxes under the individual welfare measure depends on the discount factor for future generations θ . As already mentioned above using a discount factor of $\theta = 0$ we would completely ignore the welfare effects of future generations while with a discount factor of $\theta \rightarrow (1+n)^{-1}$ we would not take into account the welfare effects of short-run generations. Figure 3 shows the optimal capital income tax rate as a function of the intertemporal discount factor θ . Not surprisingly, the tax rate increases with a higher θ as future generations get more weight in the social welfare function. When θ approaches $(1+n)^{-1} \approx 0.99$, the optimal capital tax rate again is close to the long-run tax optimum. At a discount factor of $\theta \approx 0.95$, the optimal capital income tax rate hits zero and stays there for lower values of θ .

Figure 3: Optimal capital tax rate for different discount factors θ



We can therefore conclude that from an efficiency perspective, capital income should not be taxed at high rates. A capital tax may increase long-run welfare, yet it imposes massive burdens on current generations in form of a severe savings distortion, which outweigh future welfare gains. Only when future generations get very high weight in the social welfare function, taxing capital income might be a valuable option for the policy maker.

4 Discussion and conclusion

In this paper we characterized optimal income tax policy when the short-run welfare effects of tax reforms are explicitly taken into account. We therefore developed a general and coherent way to measure welfare effects in the short-run. Our welfare measure includes the case of a utilitarian social welfare function, yet is not limited to this perspective. Another advantage of our approach is that it directly shows the social welfare weights we put on different subgroups of the population and therefore makes it easy to analyze the perspective a social planner takes when determining tax policy along such a welfare criterion.

We found that the only reason to tax capital income in the short-run is redistribution across or within generations. If the objective of a social planner is to redistribute away from generations with low marginal utility of consumption towards high marginal utility – as is the case with a utilitarian welfare criterion – a high capital income tax rate of 49% is the optimal choice. High capital income tax rates basically redistribute from middle-aged towards younger generations. Since the latter have the higher marginal utility of consumption, such a reform is beneficial under a utilitarian welfare function. When we shut down the intergenerational redistributive objective, but still favor intragenerational redistribution, the optimal capital income tax rate is much smaller at about 21%. Finally, on a pure economic efficiency basis, capital income should not be taxed at all in the short-run. Capital income taxes increase welfare of young and future generations due to the mimicking of age dependent labor tax rates and an improved insurance of labor productivity risk. For middle-aged and older cohorts, however, most of their labor productivity risk is already revealed and the number of remaining working periods is much smaller. Hence, they only slightly benefit from age dependent taxation and insurance provision. On the other hand, imposing a high capital income tax comes along with a severe savings distortion for these cohorts, which results in a huge welfare loss. These welfare losses outweigh gains of the few young generations in the short-run.

When the welfare effects of all generations along the transition path and in the new long-run equilibrium are taken into account, the optimal capital income tax rate is still almost zero from an economic efficiency point of view. In terms of the progressivity of the labor income tax code, we find that the marginal tax rate of labor income is always around 20 percent, regardless of the type of welfare measure we apply. The only thing that varies under the different optimal tax system is the basic allowance. While under a high capital income tax regime the basic allowance of the labor tax code may be as large as USD 15 000, under a no capital income tax regime the government would levy perfectly proportional labor income taxes.

Our lump-sum transfer schemes are a suitable tool to measure the pure economic efficiency effects of a reform, but they can not directly be thought of as an implementable reform option themselves. There are mainly two reasons for that: First, we compute these transfers ignoring any effects they may have on the aggregate economy. To overcome this issue we would have to implement a Lump-Sum Redistribution Authority (LSRA) that was first proposed by Auerbach and Kotlikoff (1987). Nishiyama and Smetter (2005) apply this technique to models with idiosyncratic productivity risk. Their compensation scheme for the short-run cohorts is identical to our "individual" welfare ap-

proach, i.e. they calculate for each individual currently alive a transfer that brings them back to their initial equilibrium utility level. The main difference between their approach and ours then is that the LSRA operates in general equilibrium, meaning that it issues debt on the capital market. Consequently the relevant discount rates for future cohorts are derived directly from the market interest rates $\{r_1, r_2, \dots\}$ along the transition. Hence they are endogenous and depend on the actual reform that is implemented. Fehr and Kindermann (2012) show how the optimal tax scheme changes when aggregate welfare is measured by means of a LSRA. Their results are qualitatively and quantitatively very similar to the results from the individual welfare measure presented above.

Second, the individual welfare measure relies on the assumption that individual productivity is perfectly observable. If the government was to implement such a transfer scheme, information constraints would certainly come into play. In such a situation a transfer scheme along the lines of the cohort measure would most likely be more suitable. Huang et al. (1998) propose such a scheme for the government to buy out a social security reform. Specifically when abolishing social security they assume that the government pays a lump-sum transfer to each generation that guarantees the mean agent of a cohort to be indifferent between the regime with and the one without social security. The resources that are needed to finance these transfers are collected by issuing debt on the capital market. This so-called "entitlement debt" will then have to be financed by future generations. Of course, we could interpret our simulation results in a similar way. For example, the first column in Table 3 reveals that, if the government wanted to implement the optimal tax scheme under the utilitarian welfare criterion ($\tau_k = 0.49$) while at the same making the mean agent of each cohort indifferent, the required entitlement debt would amount to about 12.46 percent of aggregate consumption (which equals roughly 7 percent of aggregate output). If the optimal cohort based tax system is implemented, the government could build up some wealth from transfers of initial cohorts.

Of course, 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 sequence 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, in a transitional analysis the status quo tax schedule plays an important role. The tax system used by Conesa et al. (2009) in the initial equilibrium is however just a rough approximation of reality, especially when it comes to the treatment of different sources of capital income. A more careful analysis of capital income taxation in the US would require to distinguish between dividend, capital gains and corporate profit taxes. Finally, 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.

References

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

- Fehr, H., M. Kallweit and F. Kindermann (2013): Should pensions be progressive? *European Economic Review* 63, 94-116.
- Fehr, H. and F. Kindermann (2012): Optimal taxation with current and future cohorts, CESifo Working Paper No. 3973.
- Fukushima, K. (2011): Quantifying the welfare gains from flexible dynamic income tax systems, Discussion Paper Series 176, Hitotsubashi University.
- Gervais, M. (2012): On the optimality of age-dependent taxes and the progressive U.S. tax system, *Journal of Economic Dynamics & Control* 36, 682-691.
- Heathcote, J., K. Storesletten and G. Violante (2014): Optimal tax progressivity: An analytical framework, Staff Report 496, Federal Reserve Bank of Minneapolis.
- 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.
- Krueger, D. and A. Ludwig (2013): Optimal progressive labor income taxation and education subsidies when education decisions and intergenerational transfers are endogenous, *American Economic Review* 103(3), 496-501.
- Nishiyama, S. and K. Smetters (2005): Consumption taxes and economic efficiency with idiosyncratic wage shocks, *Journal of Political Economy* 113, 1088-1115.
- Peterman, W.B. (2013): Determining the motives for a positive optimal tax on capital, *Journal of Economic Dynamics & Control* 37, 265-295.