Netspar

Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Optimal Ramsey Taxation with Social Security

Francesco Del Prato, Marco Francischello, Matteo Paradisi

DP 05/2022-015

Optimal Ramsey Taxation with Social Security*

Francesco Del Prato
KU Leuven, *IMT Lucca* and *CERGE-EI*

Marco Francischello
Imperial College Business School

Matteo Paradisi
Einaudi Institute for Economics and Finance

May 9, 2022

Abstract

We develop an OLG model with heterogeneous agents and aggregate uncertainty to study optimal Ramsey taxation when the government can use a credible set of social security instruments. Social security mitigates the income effect in optimal labor tax smoothing and, together with heterogeneity, adds new redistributive motives to both labor and capital taxes while crowding out others. We calibrate the model on three different economies: the US, Netherlands, and Italy. We argue that the three countries would experience heterogeneous gains, in redistributive and efficiency terms, by moving from the status-quo allocations to those prescribed by a utilitarian Ramsey planner. Our simulations show that retirement benefits in the current economies are higher than their Ramsey-optimal level while we argue that the use of funded social security schemes, neglected in current actual policies, could be welfare improving.

*We especially thank Cristiano Ricci for the research assistance with the numerical simulations. We also thank the participants to the Netspar IPW 2022 for various comments and suggestions. The realization of this article was possible thanks to the financial support of the Netspar comparative research grant 2021.

1 Introduction

Three well-established trends have recently characterized advanced economies. First, labor productivity growth has halved since the 1980s and particularly after the 2007-2009 global financial crisis (Dieppe, 2021). Second, both wealth and income inequalities have been increasing, most notably in the US (Chancel et al., 2021). Third, the population has dramatically aged: over the last forty years, the number of over-65 per 100 people of working age in OECD countries increased from 20 to 31 (OECD, 2019). These trends have heterogeneous impacts across generations and worker groups, pushing toward policy reassessments. For example, a well-designed pension system can help to reduce inequalities among retirees, while the generational policy imbalance observed in many western countries will cause the future generation to pay significantly higher taxes in the years ahead (Kotlikoff, 2009). This paper provides a new normative framework for studying how the optimal policy mix of taxes and social security instruments can help deal with these trends.

In our analysis, we rely on a neoclassical, general equilibrium Ramsey taxation model (Ramsey, 1927) augmented along three directions: *i)* we substitute the infinite-horizon households with a series of overlapping generations (OLG); *ii)* we add aggregate and idiosyncratic risks; *iii)* we include, among the policies, a rich set of social security instruments. The OLG structure breaks the Ricardian equivalence and lets us capture two relevant factors: first, the policies' incidence on a single generation and, therefore, the reduced scope for their perfect smoothing; second, the across-generations redistributive and risk-sharing concerns. Aggregate uncertainty allows us to study how the government responds to temporary shocks that impact asymmetrically on different generations, while through idiosyncratic shocks, we can investigate the need for redistribution within a single generation—in a similar way to equity concerns in static models. Finally, we add social security to the standard policy mix usually studied in Ramsey problems because it improves the credibility of our framework, being a relevant-yet-understudied government's leverage to impact all the trends we described earlier. This is why we find it crucial to model social security to mimic how many pension schemes work closely. In our model, workers receive a retirement transfer from the government made of two components: one is fixed (i.e., a minimum pension benefit), while the other is work-related (i.e., fees paid by workers during their working period, reevaluated at a specific replacement rate). The government finances this social security scheme by mixing two methods: one is taxing current payrolls (PAYGO or unfunded approach); the other is capitalizing taxes levied on the current generation of retirees while they were working (funded approach).

Through this framework, we theoretically characterize optimal distortionary labor and capital taxation and we comment their interactions with social security. The key question here is to understand how social security can be optimally used *together with* classic Ramsey policy instruments at the optimum.

First, we provide conditions for optimal labor taxation smoothing over time. In particular, we show that the government has preferences for smoothing across generations the deadweight loss of labor taxation, i.e., the net loss in welfare efficiency that follows the collection of an additional dollar of revenue through labor taxes. We formalize this intuition under a simplifying quasi-linear assumption,when such smoothing is independent of the households' types: the government fixes the same distortion pattern over time for each type so that such a pattern is decided only through preferences for intergenerational redistribution. We then introduce concavity in preferences and decompose the marginal effects of taxation in three elements: a mechanical, a distortionary, and an income effect. In doing this, we extend the smoothing formula proposed by Hipsman (2018) in three directions. First, we adjust the income effects to accommodate the mitigation brought in by the social security instruments, which reduce the response in labor supply that would follow a marginal increase in labor taxes. Second, since agents are heterogeneous in our setting, we derive formulas describing the optimal smoothing pattern of the marginal contribution to the deadweight loss of a given type. Third, we obtain additional motives for labor income taxation that stem from redistributional objectives. These new motives mark the distance between our formula and a perfect smoothing result.

We then turn to capital taxation, deriving an optimal capital taxes formula that we decompose in different terms, each representing an additional motive for taxing capital. We start showing that in the baseline case—homogeneous agents without social security—optimal taxes can be broken down into a budget, an aggregate resources, an intertemporal, and a hedging component, on the lines of Farhi (2010). Each accounts for a specific taxation motive that affects social welfare through one particular channel. In particular, the budget and the intertemporal components refer to individual contributions to capital taxation, while the other two relate to aggregate motives. Allowing for households' heterogeneity then introduces an additional redistributional component that adds a taxation motive to impose distortions to the more productive type for equity purposes. Moreover, we show that when the Ramsey planner has access to social security instruments and households are homogeneous, the aggregate terms of the decomposition disappear. The funded component of social security is indeed financed through a specific tax levied on income and capitalized on the market, thus acting as a mandatory saving from the point of view of the households. Therefore, this new instrument crowds out both the aggregate resources and hedging components. The very same does not hold when allowing for agents' heterogeneity. Actually, we argue that, compared to the baseline decomposition in the heterogeneous setting, all the terms remain unchanged, but a new funded social security term substitutes the hedging one. This new component shows that, at the Ramsey optimum, the change in the income tax base that follows a variation in the capital taxes must be taken into account by the planner, who now levies specific taxes to fund social security instruments. Thus, the planner uses this new income-related channel to smooth capital taxes when expecting adverse shocks to the economy.

Finally, we briefly discuss the insightful relationship between the pension replacement rate, i.e., the conversion rate from income to social security benefit the government fixes for the work-related component of the retirement transfer, and the debt issued by the government for financing expenses. We show that, in equilibrium, the optimal replacement rate needs to balance the present labor distortions with the future risk-free capitalization. Intuitively, a higher replacement rate calls for a larger labor supply in the present period while changing the bond's return in the next one. This effect on returns manifests both because the change in the social security instrument changes the aggregate income and because the replacement rate acts, conceptually, as another risk-free asset. Thus, the planner needs to consistently coordinate the two risk-free instruments at the optimum.

With this rich theoretical setting at disposal, we then turn to the quantitative implication of our model. To allow for a richer and more credible discussion of the policy implications we derive in our numerical exercise, we calibrate our model on three different economies: the United States, Italy, and the Netherlands. These countries exhibit different mixes of current policy and macroeconomic parameters, thus enriching our calibration's power to credibly inform policy. For computational feasibility, we focus on the deterministic steady state of the model for each country.

First, we compare the allocations achieved in the benchmark economy, i.e., those stemming from calibrating the model with the status-quo parameters, with the social optimum and Ramsey allocations, i.e., the first-best allocations that maximize social welfare with no use of the policy instruments, and the feasible allocations a Ramsey planner reach at the optimum. Overall, the social optimum implements the highest consumption for every household's type and a more efficient labor supply. Still, the status-quo economy and the Ramsey optimum achieve allocations mutually more similar than those of the social optimum since they can use the same set of policy instruments. In particular, the consumption of young agents increases across types for all three equilibria. We can show that marked differences in the consumption inequalities across the three countries persist and that the three face different gains in moving from the benchmark economy to the Ramsey optimum. For example, the US shows a consumption pattern for the young households, which is remarkably close to the optimal Ramsey allocation, while Italy could obtain much larger benefits from a policy shift. Moreover, the Ramsey optimal allocation of old consumption is steeper for all the countries except the US, where this quantity is suboptimally high for the wealthiest. Instead, old households consume the same independently of their type at the social optimum. Finally, labor supply is flat at the Ramsey optimum because of the separability of consumption and labor, and all the three countries exhibit too much labor effort exerted by low types. The social planner, on the contrary, maximizes efficiency by reaching a corner solution where the lowest types do not work at all. Overall, these numerical results provide valuable insights into the direction of policy-adjusting intervention in the three countries from both a redistributional and an efficiency perspective.

To further investigate the gains from a policy shift from the status-quo to a Ramsey optimal allocation, we analyze the consumption-equivalent variation by household type in each country. Intuitively, this measure quantifies the percentage variation in consumption each type would need to be indifferent between the two scenarios, keeping labor constant: the larger it is, the higher the gains from the shift for that type. This exercise informs in particular how (un)equal such compensation should be—something insightful on the distribution of the distance from the Ramsey optimality across types in the status-quo. For example, in the Netherlands, the gains from the shift are much more equal than in the US, where the less productive types demand a substantial compensation to be indifferent after the change.

We then turn to social security. We find that retirement benefits in the benchmark economies are higher compared to those the Ramsey optimum prescribes—with a much wider gap for Italy and the Netherlands than for the US. We treat this more as a qualitative rather than quantitative result since our two-period OLG structure mechanically calls for a low rate of work-to-retirement years, thus artificially depressing the optimal replacement rate. Remarkably, our simulations prescribe a positive funded social security component, i.e., a mandatory saving the government imposes on workers to finance their pension. This instrument is currently unused in all three countries we consider.

1.1 Related literature

Our contribution lies at the intersection between the literature on Ramsey optimal taxation and the one addressing retirement policies. Classic findings in former include results on income tax smoothing (Lucas and Stokey, 1983; Werning, 2007) and no capital taxes (Judd, 1985; Chamley, 1986). The optimal taxes have been characterized under further assumptions such as incomplete markets (Aiyagari, 1995; Aiyagari et al., 2002; Farhi, 2010) or heterogeneous households (Werning, 2007). Economies in these works feature a standard infinitely-lived cohort of households, where intergenerational redistribution is not a concern. Other papers have used the OLG setting to discuss optimal Ramsey taxation both theoretically (Atkinson and Sandmo, 1980; Erosa and Gervais, 2002; Garriga, 2019) and quantitatively (Conesa, Kitao and Krueger, 2009). These papers do not include aggregate uncertainty, and so they do not inform optimal policies for responding to stochastic shocks. Hipsman (2018) adds aggregate stochasticity to a rich OLG setting similar to ours, but with no households' heterogeneity—thus, not discussing within-generation redistribution—and no social security. Saitto (2020), on the contrary, builds an OLG optimal taxation model with households' heterogeneity but no aggregate uncertainty and, again, no social security. More precisely, none of these papers exploit social security instruments as Ramsey policies, while we characterize their interaction with classical income and capital taxes.

The literature on retirement policies has separately addressed some relevant aspects we discuss in this work, such as financial sustainability—while much less is available on intergenerational inequality. Still, no study adopts a comprehensive setting such as ours. While İmrohoroğlu, İmrohoroğlu and Joines (1998) and Krueger and Kubler (2006) have investigated the welfare effect of introducing a PAYGO social security system, they respectively allow for either idiosyncratic or aggregate risk only, while we pool the two together. Harenberg and Ludwig (2019) address the same question considering both risks, but under benefit schemes that are extremely simplified, thus different from those that are currently present in advanced economies. Ciurila (2017) explores the relationship between different benefit schemes (defined benefit vs. notional defined contribution) on long-run macroeconomic variables and welfare, without considering differences in the system funding. Bonenkamp et al. (2017) analyze how pension reforms may be an instrument to respond to demographic and financial aggregate risk only. Hosseini and Shourideh (2019) look for Pareto-efficient reforms in social security systems, proposing a test on earning and asset taxes.

The remainder of the paper is organized as follows. Section 2 lays out the model's primitives, the agents' problems and the social welfare function. Section 3 discusses the social planner's problem, while Section 4 presents and comments the Ramsey problem. Sections 5 and 6 illustrate our results on optimal labor and capital taxes, respectively, and their interactions with social security instruments. Section 7 discusses the relationship between the replacement rate and the risk-free bonds. In Section 8 we calibrate the model and in Section 9 we discuss our quantitative results in steady state. Section 10 concludes.

2 Model

We study a closed, neoclassical overlapping generation economy with two generations, nesting an infinite horizon model, in the spirit of Atkinson and Sandmo (1980). Agents are heterogeneous in their productivity and work when young in the period they are born, whereas they retire in the following period, once old. Firms are homogeneous, and maximize their profit statically. The economy faces aggregate shocks, and markets are incomplete.

2.1 Uncertainty and heterogeneity

We model aggregate risk as a discrete set of states $s_t \in \mathcal{S}$ and histories of those states $s^t = (s_0, s_1, \ldots, s_{t-1}, s_t)$. The state s_t is characterized by a shock to the economy γ_t, and evolves according to a Markov process described by the transition matrix $\boldsymbol{M}$. Government consumption is exogenous and takes the form of a government spending shock $G(s_t)$ that only depends on the present state, and not on the history. For each history s^t, there exists a one-period risk-free bond that pays a premium $R^b(s^t) = 1 + r^b(s^t)$ at all histories $s^{t+1} \geq s^t$.[1] Throughout the paper, we suppress the argument of all the terms depending on the stories, replacing it with according subscripts: for each symbol x, the reader can think of x_t as equivalent to $x(s^t)$, unless differently specified.

Households have heterogeneous labor productivity, characterized by the discrete type θ, exogenously taken from a time-varying probability density Θ_t on the support $\mathbb{S}_t$. We denote with $\Theta_t(\theta)$ the probability that a worker is of type θ at time t.

2.2 Demographics and timing

Young. Time is discrete. At each time t a new generation of size n^y_t is born.[2] In this period, agents observe their type θ and the realized state of the world s_t, and choose their labor supply $l_{\theta,t}$, the investment in risky capital $q_{\theta,t}$, the investment in public debt $b_{\theta,t}$, and their consumption $c^y_{\theta,t}$.

Old. A share ψ_{t+1} of the young generation at time t survives in the following period so that $\psi_{t+1} n^o_{t+1}$ old workers retire at $t+1$. Old households consume the returns from the previous-period investment choices and receive a social security benefit from the government. The aggregate capital income not enjoyed by old agents due to mortality risk becomes unintended bequests and is added to the government resources for the sake of the model's tractability.[3]

2.3 Government policies

We let the government choose an optimal policy mix of public debt, capital and labor taxes. In addition, we allow for an additional instrument to redistribute across generations by including a social security system that encompasses two benefit schemes and two financing mechanisms. Retirees obtain a mix of defined benefits and work-related contributions. The government finances these benefits through a funded and an unfunded component.

[1] Since the bond is risk-free, its return at $t+1$ only depends on history s^t.

[2] These agents are then *young* at t and *old* at $t+1$. Therefore, $n^o_t = n^y_{t-1}$.

[3] Optimal bequest taxation with heterogeneous households has been extensively studied in papers as Farhi and Werning (2010) and Piketty and Saez (2013).

Formally, the government sets capital taxes τ_t^K one period in advance so that returns at time $t+1$ are taxed at a rate that depends on history s^t (Farhi, 2010). These taxes are levied on the returns on risk-free bonds, and on capital returns net of the exogenous depreciation rate δ. Thus, the after-tax return is $R_t^{b,\tau} = 1 + (1-\tau_t^K)\, r_t^b$ for the risk-free bond with return r_t^b, and $R_{t+1}^{K,\tau} = 1 + (1-\tau_t^K)\, r_{t+1}^K$ for the risky capital, where r_{t+1}^K is the risky rate in $t+1$, net of capital depreciation. Moreover, we call $R_{t+1}^K = 1 + r_{t+1}^K$ the pre-tax return on risky capital and with $R_t^b = 1 + r_t^b$ the pre-tax return on the bond.

Moreover, the government levies the following total income taxes on a worker with productivity θ earning wage $w_{\theta,t}$:

$$T\left(w_{\theta,t} l_{\theta,t}\right) = \left[\tau_t^l + \tau_t^{\text{ss}}\right] w_{\theta,t} l_{\theta,t}$$

where $\tau^l \in (-\infty, 1)$ is the linear labor income tax rate, and $\tau^{\text{ss}} = \tau^{\text{ss,U}} + \tau^{\text{ss,F}}$ is the income tax financing the social security, consisting of the unfunded component $\tau^{\text{ss,U}}$, which bankrolls current benefits, and the funded one, $\tau^{\text{ss,F}}$, invested in the capital market to finance future benefits. Moreover, $w_\theta(s^{t-1})$ is the wage paid to a worker with productivity θ. Notice that, from the point of view of the household, such a tax structure implies that it is impossible to distinguish the unfunded component of the social security from the labor income tax. Thus, we will treat the former as an implicit component of the latter in the whole paper.

Finally, the government transfers a social security benefit $y_{\theta,t}^{\text{ss}}$ to a retired worker of type θ in each period. The benefit consists of a work-related and a fixed component (i.e., a minimum pension benefit):

$$y_{\theta,t}^{\text{ss}} = \kappa_{t-1}^{\text{ss}} w_{\theta,t-1} l_{\theta,t-1} + \bar{y}_{t-1}^{\text{ss}}$$

κ^{ss} is the replacement rate, regulating the amount of labor income converted to pension, and $\bar{y}_{t-1}^{\text{ss}}$ is the fixed component of the benefit—both decided in $t-1$.

2.4 The agents' problems

The economy is populated by three type of agents: homogeneous firms, heterogeneous households and the government.

Firms A representative firm has constant returns to scale technology $F(K, L)$ using capital and efficient labor and statically maximizes its profit given a collection of prices $Q \equiv \{w_{\theta,t}, r_t^K\}$. Thus, it solves the following problem

$$\begin{aligned} \underset{K_{t-1},\, \{l_{\theta,t}\}_{\theta\in\mathbb{S}}}{\text{maximize}} \quad & \gamma_t F\left(K_{t-1}, L_t\right) - n_t^y \sum_\theta w_{\theta,t} l_{\theta,t} \Theta_t(\theta) - r_t^K K_{t-1} \\ \text{subject to} \quad & L_t = n_t^y \sum_\theta \theta l_{\theta,t} \Theta_t(\theta) \end{aligned}$$

Optimality conditions equate capital and labor prices to their marginal product[4]:

$$\theta \gamma_t F_{L,t} = w_{\theta,t} \quad \text{and} \quad \gamma_t F_{K,t} = r_t^K$$

[4]For better readability, we omit everywhere the arguments of first and second derivatives of the production function.

To stress the dependence of prices from θs and the aggregate shock, we denote $\bar{w}(s^t) = F_{L,t}$ and $\bar{r}^K(s^t) = F_{K,t}$ so that a worker earns $w_{\theta,t} = \gamma_t \bar{w}(s^t)\theta$ per unit of labor and receives $r_t^K = \gamma_t \bar{r}^K(s^t)$ per unit of capital invested.

Government At time t, the government sets a collection of policies $\mathcal{P} \equiv \{\tau_t^l, \tau_t^{\text{ss,U}}, \tau_t^{\text{ss,F}}, \kappa_t^{\text{ss}}, \tau_{t-1}^K, r_{t-1}^b\}$, given the exogenous spending G_t, that satisfies the following budget constraint

$$
\begin{aligned}
& n_t^o R_t^{b,\tau} \sum_\theta b_{\theta,t-1}\Theta_t(\theta) + G_t + n_t^o \psi_t \sum_\theta y_{\theta,t}^{\text{ss}} \Theta_t(\theta) \\
& \leq n_t^y \sum_\theta (\tau_t^l + \tau_t^{\text{ss,U}}) w_{\theta,t} l_{\theta,t} \Theta_t(\theta) + n_t^o \tau_{t-1}^K r_t^K \sum_\theta q_{\theta,t-1}\Theta_t(\theta) \\
& + n_t^y \sum_\theta b_{\theta,t}\Theta_t(\theta) + n_t^o \tau_{t-1}^{\text{ss,F}} R_t^K \sum_\theta w_{\theta,t-1} l_{\theta,t-1}\Theta_t(\theta) + n_t^o (1-\psi_t) R_t^K \sum_\theta q_{\theta,t-1}\Theta_t(\theta)
\end{aligned}
$$

i.e., the government uses debt to balance the difference between expenditures and revenues from income, capital and social security taxes.

Households A worker of type θ from a given generation who is young at time t chooses the allocations $\mathcal{H} \equiv \{c_{\theta,t}^y, c_{\theta,t+1}^o, l_{\theta,t}, q_{\theta,t}, b_{\theta,t}\}$ to maximize life-time utility discounted at rate $\beta \in (0,1)$

$$
\begin{aligned}
\underset{\mathcal{H}}{\text{maximize}} \quad & U\left(c_{\theta,t}^y, c_{\theta,t+1}^o, l_{\theta,t}\right) = u\left(c_{\theta,t}^y, l_{\theta,t}\right) + \beta \psi_{t+1} \mathrm{E}_t\left[u\left(c_{\theta,t+1}^o, 0\right)\right] \\
\text{subject to} \quad & c_{\theta,t}^y + q_{\theta,t} + b_{\theta,t} \leq \left[1 - \tau_t^l - \tau_t^{\text{ss,F}} - \tau_t^{\text{ss,U}}\right] w_{\theta,t} l_{\theta,t}, && (2.1) \\
& c_{\theta,t+1}^o \leq R_{t+1}^{K,\tau} q_{\theta,t} + R_t^{b,\tau} b_{\theta,t} + \kappa_t^{\text{ss}} w_{\theta,t} l_{\theta,t} + \bar{y}_t^{\text{ss}} && (2.2)
\end{aligned}
$$

Optimal allocations for a household must satisfy the no-arbitrage condition[5]

$$
\mathrm{E}_t\left[R_{t+1}^{K,\tau}\right] - R_t^{b,\tau} = \frac{-\mathrm{Cov}\left(R_{t+1}^{K,\tau}, u_{c,\theta,t+1}^o\right)}{\mathrm{E}_t\left[u_{c,\theta,t+1}^o\right]} \tag{2.3}
$$

and the labor-leisure condition

$$
-\frac{u_{l,\theta,t}^y}{u_{c,\theta,t}^y} = \left[1 - \tau_t^l - \tau_t^{\text{ss}} + \frac{\kappa_t^{\text{ss}}}{R_t^{b,\tau}}\right] w_{\theta,t} \tag{2.4}
$$

where the labor wedge depends on labor income taxes, social security taxes, and is decreased by the discounted expected return on social security, which depends on the replacement rate. Notice that an increase in the replacement rate reduces the marginal disutility of labor since it gives households incentives to work today to enjoy higher consumption tomorrow. Indeed, an instrument that links present earnings to future consumption acts as a risk-free mandatory saving program. This explains why the return on the government's debt scales it in (2.4): the government must set the two hand in hand to avoid arbitrages between the two risk-free assets. We will discuss this implication extensively later on when talking about optimal taxes.

[5]For readability, we omit the arguments of the marginal utilities in all the text.

2.5 Feasibility, market clearing, and equilibrium

Three equations regulate the flow of resources in the economy. Total investment in risky assets must equate the aggregate level of capital. Hence, the market clearing condition in the capital market requires

$$K_t = n_t^y \sum_\theta q_{\theta,t}\Theta_t(\theta) + n_t^y \tau_t^{\text{ss,F}} \sum_\theta w_{\theta,t} l_{\theta,t}\Theta_t(\theta) \tag{2.5}$$

The efficient amount of labor employed by the firm must be consistent with the households' labor supply:

$$L_t = n_t^y \sum_\theta \theta l_{\theta,t}\Theta_t(\theta) \tag{2.6}$$

Finally, the aggregate resource constraint implies the following feasibility condition

$$n_t^y \sum_\theta c_{\theta,t}^y \Theta_t(\theta) + n_t^o \psi_t \sum_\theta c_{\theta,t}^o \Theta_t(\theta) + G_t + K_t \leq \gamma_t F\left(K_{t-1}, L_t\right) \tag{2.7}$$

Using the setup outlined above, we can define a competitive equilibrium as follows:

Definition 1 (Competitive equilibrium)**.** Given the initial stocks $b_\theta(s^{-1})$ and $q_\theta(s^{-1})$ for each $\theta \in \Theta_{-1}$, and the initial prices and policies $r^K(s^{-1})$, $r^b(s^{-1})$, $\tau^K(s^{-1})$ and $\kappa^{\text{ss}}(s^{-1})$, a competitive equilibrium is a set of prices $\mathcal{Q}$, policies $\mathcal{P}$ and allocations $\mathcal{H}$ such that households and firms maximize their utilities under the budget constraints, the feasibility constraint holds, and markets clear.

2.6 Social welfare function

We measure social welfare as the discounted sum of households' utilities weighted by two sets of welfare weights. Preferences for redistribution across generations are captured by the set of generational weights $\{\phi_t\}_{t\geq 0}$, while weights $\{g(\theta)\}_{\theta\in\mathbb{S}}$ measure the social desire for redistribution within a given generation. Therefore, we define the ex-ante social welfare function as

$$W \equiv \phi_0 \beta \psi_0 n_0^o \sum_\theta g(\theta)\, u\left(c_\theta^0, 0\right)\Theta_t(\theta) + \mathrm{E}_0\left[\sum_{t\geq 0} \phi_t n_t^y \sum_\theta g(\theta)\, U\left(c_{\theta,t}^y, l_{\theta,t}, c_{\theta,t+1}^o\right)\Theta_t(\theta)\right] \tag{2.8}$$

The concavity of the utility function already provides motives for redistribution even when the government is utilitarian and $g(\theta) = \bar{g}$ for any θ. Indeed, a decreasing marginal utility implies that a government prefers allocations with low dispersion in consumption. Therefore, welfare weights $g(\theta)$ capture preferences for redistribution on top of the desire for redistribution implied by the agents' preferences.

3 Social planner

We introduce a social planner optimum to characterize the first-best allocation that will serve as a benchmark to evaluate welfare in the Ramsey optimum. Under the feasibility constraint and the market clearing condition for labor, a social planner maximizes the discounted sum of the generations average utilities by directly choosing among implementable allocations. Therefore, we state the planner's problem as

Definition 2 (Planner's problem)**.** The planner's problem is a maximization of the ex-ante social welfare function (2.8) subject the following constraints:

- the feasibility constraint (2.7);
- the market clearing condition for labor (2.6);
- the initial condition on capital $K_0 = n_0^o \sum_\theta q_\theta(s_0)\Theta_t(\theta)$.

Since the planner has direct access to allocations, the problem does not account for the instruments detailed in Section 2.3. The only relevant constraints are the feasibility and efficient labor ones, plus an initial condition on capital. We refer to the solution of the social planner problem as a *social optimum*.

Proposition 1 (Social optimum)**.** The social optimum allocations satisfy the following conditions:

$$\frac{u^y_{c,\theta,t}}{u^y_{c,\theta',t}} = \frac{g(\theta')}{g(\theta)} \text{ and } \frac{u^o_{c,\theta,t}}{u^o_{c,\theta',t}} = \frac{g(\theta')}{g(\theta)}, \quad \forall\theta,\theta' \tag{3.1}$$

$$-\frac{u^y_{l,\theta,t}}{u^y_{c,\theta,t}} = \theta\gamma_t F_{L,t} \tag{3.2}$$

$$\frac{u^y_{c,\theta,t}}{\beta u^o_{c,\theta,t}} = \frac{\phi_{t-1}}{\phi_t} \tag{3.3}$$

$$\frac{\phi_t}{\phi_{t+1}} \frac{u^y_{c,\theta,t}}{\mathrm{E}_t\left[u^y_{c,\theta,t+1}\left(1+\gamma_{t+1}F_{K,t+1}-\delta\right)\right]} = 1 \tag{3.4}$$

Equations (3.1)-(3.4) are derived from the first-order conditions on the social planner's problem as illustrated in Appendix A.2.

First, equation (3.1) shows how the planner trades off the marginal utilities of different types. In the corner case with constant welfare weights, i.e., $g(\theta) = \bar{g}$ for each θ, the planner is utilitarian and equalizes the marginal utility of consumption across types for both young and older workers. Hence, the planner allocates the same consumption to all the old types within a given generation since their marginal utility does not depend on labor. This equality result holds for the young generation if and only if the preferences are separable in labor and consumption. Outside the utilitarian benchmark, given any two types within the same generation, the social planner trades off their marginal utilities of consumption depending on the ratio of their social welfare weights. In particular, the marginal utility of consumption across two generic types θ and θ' grows proportionally to the inverse ratio of their welfare weights, guaranteeing that types with lower $g(\cdot)$ get lower consumption in equilibrium.

Equation (3.2) is the standard labor-leisure condition that equates the marginal disutility of working to the marginal product of labor. It shows that the social planner requires more productive individuals to work more for efficiency reasons. More hours of work are compensated as described by equation (3.3). Because the marginal utility of consumption decreases in labor when preferences are not separable and the consumption of old agents is constant across households, the planner must promise larger consumption when young to high ability individuals to compensate them for their higher labor supply. This compensation disappears if preferences are separable and the marginal utility of consumption is independent of labor supply.

Finally, equation (3.4) pins down the optimal savings path, which in turn determines how aggregate capital is transferred across generations. Capital is set in equilibrium such that the marginal effect of transferring one unit of consumption across young agents of two different generations equates the ratio of their generational welfare weights.

The social planner problem presumes that the government observes individual types and that there are unconstrained instruments. In the next Section, we set up a Ramsey problem where the government chooses optimal policies from a given set of instruments to maximize social welfare, keeping agents' optimal responses into account.

4 Ramsey planner

A Ramsey planner is a benevolent government that maximizes social welfare choosing an optimal mix of policies $\mathcal{P}$. Hence, the planner chooses the Competitive Equilibrium that maximizes the social welfare function. We define the Ramsey Planner's problem as follows.

Definition 3 (Ramsey Problem)**.** The Ramsey problem is a maximization of the ex-ante social welfare function (2.8) subject to the following constraints:

- the young (2.1) and old (2.2) generations budget constraints;
- the feasibility constraint (2.7);
- the market clearing conditions for capital (2.5) and labor (2.6);
- the no-arbitrage (2.3) and labor-leisure (2.4) conditions.

We rewrite the problem using a variation on the so-called *primal approach* (Lucas and Stokey, 1983) since we have a rich set of policies at the government's use. Therefore, we let the government search for the optimal allocations *given* the social security instruments $\tau^{\text{ss,F}}$ and κ^{ss}, and then we derive the other Ramsey optimal supporting policies and prices. Then, we ensure that the obtained set of policies, prices, and allocations satisfy the conditions for being a competitive equilibrium. Since in the following section we will focus our analysis on the distortion caused by the policies enacted by the government, to reduce the dimensionality of the problem we assume $\bar{y}^{\text{ss}} = 0$. The following Definition and Proposition formalize this approach.

Definition 4 (Implementable allocation)**.** Given a set of cross-periods optimal policies $\{\tau_t^{\text{ss,F}}, \kappa_t^{\text{ss}}\}_{t\geq 0}$, we say that an allocation $\left\{c_{\theta,t}^{y}, c_{\theta,t}^{o}, q_{\theta,t}, b_{\theta,t}, l_{\theta,t}\right\}_{t\geq 0}$ is implementable if the following implementability conditions (IC) hold for each s^t, $t \geq 0$ and for each $\theta \in \mathbb{S}$:

IC on the young generation:

$$u^y_{l,\theta,t} l_{\theta,t} + u^y_{c,\theta,t} \left[c^y_{\theta,t} + q_{\theta,t} + b_{\theta,t} \right] = -\beta \psi_{t+1} \kappa^{\text{ss}}_t \theta \gamma_t F_{L,t} l_{\theta,t} \, \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right] \qquad \forall \theta \in \mathbb{S} \tag{4.1}$$

IC on the old generation:

$$\begin{aligned} c^o_{\theta,t+1} =& \frac{u^y_{c,\theta,t}}{\beta \psi_{t+1} \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right]} b_{\theta,t} + \kappa^{\text{ss}}_t \theta \gamma_t F_{L,t} + \\ & q_{\theta,t} \left(1 + \frac{u^y_{c,\theta,t} - \beta \psi_{t+1} \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right]}{\beta \psi_{t+1} \mathrm{E}_t \left[\hat{F}_{K,t+1} u^o_{c,\theta,t+1} \right]} \hat{F}_{K,t+1} \right) \qquad \forall \theta \in \mathbb{S} \end{aligned} \tag{4.2}$$

IC on the feasibility constraint:

$$\begin{aligned} \gamma_t F(K_{t-1}, L_t) =& n^y_t \sum_\theta c^y_{\theta,t} \Theta_t(\theta) + n^o_t \sum_\theta c^o_{\theta,t} \Theta_t(\theta) + G_t \\ & + n^y_t \sum_\theta q_{\theta,t} \Theta_t(\theta) + n^y_t \tau^{\text{ss,F}}_t \gamma_t F_{L,t} \sum_\theta \theta l_{\theta,t} \Theta_t(\theta) \\ & - (1-\delta) \left[n^y_{t-1} \sum_\theta q_{\theta,t-1} \Theta_{t-1}(\theta) + n^y_{t-1} \tau^{\text{ss,F}}_{t-1} \gamma_{t-1} F_{L,t-1} \sum_\theta \theta l_{\theta,t-1} \Theta_{t-1}(\theta) \right] \end{aligned} \tag{4.3}$$

IC on the income taxes:

$$\left[1 - \tau^l_t - \tau^{\text{ss,F}}_t - \tau^{\text{ss,U}}_t \right] = - \frac{u^y_{l,\theta,t} + \beta \psi_{t+1} \kappa^{\text{ss}}_t F_{L,t} \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right]}{\theta \gamma_t F_{L,t} u^y_{c,\theta,t}} \qquad \forall \theta \in \mathbb{S} \tag{4.4}$$

IC on the capital taxes:

$$1 - \tau^K_t = \frac{u^y_{c,\theta,t} - \beta \psi_{t+1} \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right]}{\beta \psi_{t+1} \mathrm{E}_t \left[\hat{F}_{K,t+1} u^o_{c,\theta,t+1} \right]} \qquad \forall \theta \in \mathbb{S} \tag{4.5}$$

where $\hat{F}_{K,t+1} = \gamma_{t+1} F_{K,t+1} - \delta$.

Proposition 2 (Implementability conditions)**.** If a set of cross-periods optimal policies $\tau^{\text{ss,F}}_t, \kappa^{\text{ss}}_t$ and an allocation $c^y_{\theta,t}, c^o_{\theta,t}, q_{\theta,t}, b_{\theta,t}, l_{\theta,t}$ solves the Ramsey problem, then the allocation must be implementable, i.e., the ICs (4.1), (4.2), (4.3), (4.4), (4.5) must hold.

Proof. The proof can be found in Appendix A.3. □

To solve the Ramsey problem, we attach an agent-specific—except for the feasibility constraint—Lagrange multiplier for every $t \geq 0$ to each constraint in Proposition 2. In particular, we define the multipliers $\lambda^y_{\theta,t}$ for (4.1), $\lambda^o_{\theta,t}$ for (4.2), and λ^f_t for (4.3). Moreover, since the policies are not

θ-specific, ICs on income (4.4) and capital (4.5) taxes ensure that the agents' choices are consistent across all workers' types. Thus, the differences between the rhs of the two constraints must be zero for every couple θ, θ', and we attach the multipliers $\lambda^L_{\theta,\theta',t}$ and $\lambda^K_{\theta,\theta',t}$ to these differences. This approach is similar to the one adopted in Chari, Christiano and Kehoe (1994).

5 Deadweight loss from labor taxation and social security

In this section, we study the optimal smoothing of labor taxes and their interaction with social security instruments. Since the structure of the model does not allow to derive explicit formulas for marginal tax rates (or the labor wedge), we focus on the marginal deadweight loss (DWL) from labor taxation as a measure of labor distortions. The DWL is the net loss in efficiency raising from collecting an additional dollar of revenue through the labor tax.

We link the DWL to our formulas with the following argument. At a Ramsey optimum, the DWL from labor taxes must equate the monetary social value of providing a lump-sum transfer to all individuals. The Lemma below formalizes this intuition.

Lemma 1. At the optimum, the marginal deadweight loss associated with an increase in labor taxes at the story s^t, $\mathcal{D}^l_t$, is equal to the total government value of collecting a dollar through a lump-sum transfer from young agents:

$$\mathcal{D}^l_t = -\sum_{\theta} \frac{\lambda^y_{\theta,t}}{\mathcal{W}_{\theta,t}} \tag{5.1}$$

where $\mathcal{W}_{\theta,t} = n^y_t \phi_t g(\theta)\, \Theta_t(\theta)$ is the social marginal value of increasing utils for a young worker of type θ at time t.

Proof. The proof can be found in Appendix A.4. □

The multipliers $\lambda^y_{\theta,t}$ in the numerator of (5.1) capture the value of transferring a dollar to young individuals, while the denominator represents the government value of increasing by one util the utility of young agents and converts utils into money.

Due to the richness of our model, it is not always possible to make $\mathcal{D}^l$ explicit in our formulas. Thus, we also discuss the individual contribution to the DWL, *i.e.*, the value of relaxing the implementability condition for a young agent of type θ, normalized by the marginal social utility of increasing their consumption. Formally:

Definition 5 (Individual contribution to the DWL from labor taxation)**.** The individual contribution of a type-θ individual born at s^t to the DWL caused by labor taxation, is defined as $d^y_{\theta,t} = -\frac{\lambda^y_{\theta,t}}{\mathcal{W}_{\theta,t}}$.

5.1 The deadweight loss from labor taxation in the quasi-linear case

For illustrative purposes, we start our discussion on the optimal DWL from a model with quasi-linear preferences, which delivers simple formulas. We assume that $u\left(c^y_{\theta,t}, l_{\theta,t}\right) = c^y_{\theta,t} + h(l_{\theta,t})$, where h is a decreasing concave function. The marginal utility of consumption is therefore constant across θs and equal to 1.

Proposition 3. At the Ramsey optimum, the DWL from labor taxation under quasi-linear preferences evolves as follows

$$\phi_t n_t^y \left(1 + \bar{\mathcal{D}}_t^l\right) = \frac{\phi_{t+1}}{\beta} n_{t+1}^y \, \mathrm{E}_t \left[1 + \bar{\mathcal{D}}_{t+1}^l\right] \tag{5.2}$$

where $\bar{\mathcal{D}}_t^l$ is the weighted average DWL across θs. Similarly, the individual contribution follows

$$\phi_t n_t^y \, \Theta_t(\theta) \left(1 + d_{\theta,t}^y\right) = \frac{\phi_{t+1}}{\beta} n_{t+1}^y \, \Theta_{t+1}(\theta) \, \mathrm{E}_t \left[1 + d_{\theta,t+1}^y\right] \tag{5.3}$$

Proof. Proof can be found in Appendix A.4.2. □

Equations (5.2) and (5.3) show the government's preferences for smoothing the DWL from labor taxation over time. The former establishes that the evolution of the aggregate DWL follows a martingale. Such a structure is a byproduct of market incompleteness and it echoes the tax smoothing results in incomplete markets established in Barro (1979), and in Farhi (2010) who obtains a similar martingale structure for the government needs for funds and the marginal cost of taxation.

Equation (5.3) extends this result to the smoothing of individual DWL contributions. It shows that the DWL smoothing is independent of θ. In other words, the ratio between the contribution of young agents' of a certain θ to the DWL for two contiguous generations only depends on the proportion of the generational welfare weights. As an implication, the government chooses the same distortion pattern overtime for every young agent's type so that the growth in individual distortions is only pinned down by preferences for redistribution across generations. This result a byproduct of the quasi-linearity: since the marginal utility of consumption is constant, the government has no incentive to compress consumption heterogeneously across types. Concavity in consumption utility breaks down this result as we illustrate in the next Subsection.

5.2 The deadweight loss from labor taxation in the general case

We discuss now the general case where preferences are concave and income effects exist. To make formulas more intuitive, we preserve the assumption of separability between consumption and labor, but we report the generic formula in Appendix A.4. We start with optimal DWL smoothing in the homogeneous agents case to distinguish between taxation motives based on the intertemporal balancing of income and substitution effects and those based on the preferences for redistribution across different θs.

Proposition 4. (DWL with homogeneous agents). Suppose agents preferences are separable in consumption and labor and that there is a representative agent for each generation, at the Ramsey optimum the DWL evolves following

$$\tilde{\mathcal{W}}_t \left[1 + \mathcal{D}_t^l \left(1 + MB_t^L\right)\right] = R_t^{b,\tau} \psi_{t+1} \, \mathrm{E}_t \left[\tilde{\mathcal{W}}_{t+1} \left(1 + \mathcal{D}_{t+1}^l \left(1 + MB_{t+1}^L\right)\right)\right] \tag{5.4}$$

where

$$MB_t^L = \varepsilon_{b,t}^{R_t^{b,\tau}} \left(1 - \frac{\kappa_t^{\text{ss}}}{R_t^{b,\tau}} \frac{w_t l_t}{b_t} \Pr\left(s^{t+1}|s^t\right)\right) + \frac{\mathrm{d}\tau_t^K}{\mathrm{d}b_t} \frac{q_t \, \mathrm{E}_t \left[\hat{F}_{K,t+1}\lambda_{t+1}^o\right]}{R_t^{b,\tau} \, \mathrm{E}_t \left[\lambda_{t+1}^o\right]} \tag{5.5}$$

and

$$MB_{t+1}^L = -\sigma_{t+1}^y \left[1 - \frac{q_{t+1}}{c_{t+1}^y} \left(\frac{\mathrm{E}_{t+1}\left[\lambda_{t+2}^o \hat{F}_{K,t+2}\right]}{r_{t+1}^b \, \mathrm{E}_{t+1}\left[\lambda_{t+2}^o\right]} - 1\right)\right] \tag{5.6}$$

Proof. Proof can be found in Appendix A.4.3. □

Equation (5.4) mimics the optimal smoothing in (5.2) up to the use of social marginal utilities $\tilde{\mathcal{W}}_t$ that arise from concave preferences, and the presence of the intertemporal marginal benefits of labor taxation MB^L. The formula balances (in expectation) the marginal costs of labor taxation at two consecutive histories. An increase in the labor tax has a mechanical effect of 1 on consumption, captured by the first terms on the left- and right-hand-side. Moreover, the tax increase generates mechanical distortions at both histories that are measured by the second terms on both sides. Finally, intertemporal marginal benefits of distortions that depend on income effects arise and are measured by MB^L. These terms arise because income effects are not considered pure distortions and therefore should not be incorporated in the marginal cost of taxation.[6]

The intertemporal marginal benefit at s^t (i.e., MB_t^L) has two components that partly resemble those in Hipsman (2018). First, an increase in labor taxes reduces the number of issued bonds, whose return will decrease proportionally to the elasticity of risk-free returns to issued bonds $\varepsilon_{b,t}^{R_t^{b,\tau}}$. This change in $R_t^{b,\tau}$ will make the young poorer. To offset it, they will work more because of the income effect and will thus diminish the distortionary cost of taxation. The social security benefit partially mitigates this effect since lower interest rates make the replacement rate κ^{ss} more appealing for the household reducing the incentives to supply labor in response to a change in $R_t^{b,\tau}$. Second, a change in the risk-free interest rate implies that the government must increase taxes by $\mathrm{d}\tau_t^K/\mathrm{d}b_t$ to reduce the after-tax return on capital (by a no arbitrage argument). Here, $\mathrm{d}\tau_t^K/\mathrm{d}b_t$ is the increase in capital taxes needed to keep young agent's consumption constant after a change in issued bonds. Therefore the second term in MB_t^L captures how this change in capital taxes will mechanically affect the old generation at time $t+1$ (i.e. the young at time t).

The intertemporal marginal benefit at s^{t+1} (i.e., MB_{t+1}^L) also consists of two components. First, the direct intertemporal distortionary cost generated from the consumption drop, which is proportional to the intertemporal elasticity of substitution $\sigma_{\theta,t+1}^y$. Second, because the tax hike induced by the no arbitrage condition will be phased out after s^{t+1}, the last term balances the capital tax term in MB_{t+1}^L.

5.2.1 Introducing heterogeneity and the need for redistribution

So far, we have ignored the within-generation heterogeneity, and all the motives to tax labor that arise from this model's feature. In this Section, we add this ingredient and discuss how it augments

[6]Notice that the MB terms are equal to zero in the case of quasi-linear preferences since they only depend on income effects that are absent in the quasi-linear case.

the formula in (5.4). Let us define with $MB^L_{\theta,t}$ and $MB^L_{\theta,t+1}$ the versions of equations (5.5) and (5.6) where we employ the θ-specific quantities or prices for all quantities and prices that can be θ-specific. The following Proposition provides a DWL formula for the heterogeneous agents model.

Proposition 5. (DWL with heterogeneous agents). Suppose agents preferences are separable in consumption and labor and that there are two types $\theta > \theta'$ for each generation. At the Ramsey optimum, the DWL evolves as follows

$$\tilde{\mathcal{W}}_{\theta,t}\left[1 + d^y_{\theta,t}\left(1 + \widetilde{MB}^L_{\theta,t}\right)\right] = R^{b,\tau}_t \psi_{t+1}\,\mathrm{E}_t\left[\tilde{\mathcal{W}}_{\theta,t+1}\left(1 + d^y_{\theta,t+1}\left(1 + \widetilde{MB}^L_{\theta,t+1}\right)\right)\right] \tag{5.7}$$

where

$$\widetilde{MB}^L_{\theta,t} = MB^L_{\theta,t} - \frac{1}{R^{b,\tau}_t\,\mathrm{E}_t\left[\lambda^o_{\theta,t+1}\right]}\left(\varepsilon^{R^{b,\tau}_t}_{b_{\theta,t}}\frac{\lambda^L_{\theta,\theta',t}}{b_{\theta,t}}\frac{\kappa^{ss}_t}{R^{b,\tau}_t}\psi_{t+1} + \frac{\mathrm{d}\tau^K_t}{\mathrm{d}b_{\theta,t}}\frac{\lambda^K_{\theta,\theta',t}}{\beta}\right) \tag{5.8}$$

and

$$\widetilde{MB}^L_{\theta,t+1} = MB^L_{\theta,t+1} - \frac{\sigma^y_{\theta,t+1}}{c^y_{\theta,t+1}}\frac{\lambda^K_{\theta,\theta',t+1} - \left(1 - \tau^l_{t+1}\right)\lambda^L_{\theta,\theta',t+1}}{r^b_{t+1}\beta\,\mathrm{E}_{t+1}\left[\lambda^o_{\theta,t+2}\right]} \tag{5.9}$$

Proof. Proof can be found in Appendix A.4.4. □

The smoothing of DWL contributions closely follows the structure of (5.4), where aggregate DWL is replace by individual contributions to the DWL, and the intertemporal marginal benefits of taxation are augmented by two factors presented in equations (5.8) and (5.9) that capture the motives for redistribution. These extra terms depend on the income effects discussed in (5.5) and (5.6). Indeed, while income effects do not affect efficiency, they do have an impact on the redistribution across agents within a generation, and should therefore be taken into account.

The second term in (5.8) measures how the changes in bond returns and in capital taxes discussed in the discussion to Proposition 4 affect the margins of choice of an agent θ. Under the assumption that $\theta > \theta'$, we obtain that $\lambda^L_{\theta,\theta',t}, \lambda^K_{\theta,\theta',t} > 0$. It follows that larger labor supply reductions on richer individuals (from the fact that κ^{ss} becomes more attractive relative to bonds) decrease the intertemporal marginal benefit and generate a motive to increase distortions on type θ. At the same time, the increase in the capital tax $\mathrm{d}\tau^K_t/\mathrm{d}b_{\theta,t}$ decreases the intertemporal marginal benefit if it affects richer individuals more, and therefore induces the government to set a larger DWL contribution for θ. The term would have the opposite sign if the agent was of type θ'. Similarly to what we discussed for the marginal benefit in (5.6), the second term in (5.9) balances the latter term in (5.8) and measures the benefits from reducing capital taxes between period $t+1$ and $t+2$. If the benefit is larger for richer individuals, it generates motives to reduce the DWL contributions of these individuals.

6 Capital taxes and social security

This section presents the results regarding optimal capital taxes and their interactions with social security through an incremental approach, starting from a benchmark case with homogeneous agents and no social security instruments, then building the intuition of the marginal effects of relaxing these initial assumptions.

We express τ^K as a sum of different components, each of which represents an additional motive to tax capital. Some of these terms are related to the benchmark result on optimal capital taxes of Farhi (2010), but with notable differences that stem from the relevant additions in our model: the OLG structure, the households heterogeneity, and the social security instruments.

In particular, we show that the funded component of social security taxes crowds out the capital taxation motives related to aggregate quantities in the homogeneous workers' case and adds new motives when assuming heterogeneity among households.

6.1 Capital taxes with homogeneous agents and no social security

Proposition 6 (Optimal capital taxes with homogeneous agents and no social security)**.** If agents are homogeneous and the government does not have access to social security instruments, optimal capital taxes at the story s^t with $t \geq 1$ are composed of three terms: aggregate resources, intertemporal, and hedging—each scaled by the mechanical budget component.

$$\tau_t^K = \frac{1}{B_t(\lambda_{t+1}^o)} \left[T_t^A \left(\lambda_t^f, \lambda_{t+1}^f \right) + T_t^I \left(\lambda_t^y, \lambda_{t+1}^o \right) + T_t^H \left(\lambda_{t+1}^o \right) \right] \tag{6.1}$$

Proof. Proof can be found in Appendix A.5.1. □

Through the decomposition given in Proposition 6, we provide a distinct interpretation of each motive of the optimal capital tax. In the remainder of this section, we separately discuss the contribution of each term of the decomposition.

Budget component. The budget component is given by

$$B_t(\lambda_{t+1}^o) = \frac{1}{n_t^y} \beta \, \mathrm{E}_t \left[\lambda_{t+1}^o \hat{F}_{K,t+1} \right] \tag{6.2}$$

An increase in τ_t^K mechanically strengthens the budget constraint on the old generation, proportionally to the capital tax base $\hat{F}_{K,t+1} = \gamma_{t+1} F_{K,t+1} - \delta$. Since capital taxes are fixed one period in advance, this mechanical effect is appropriately discounted by β. Multiplying by the marginal social utility converts the quantity in welfare utils, while dividing by n_t^y delivers an average for each agent subject to the tax.

Aggregate resources component. The term representing the motive for capital taxes linked to aggregate resources is

$$T_t^A \left(\lambda_t^f, \lambda_{t+1}^f \right) = \mathrm{E}_t \left[\lambda_{t+1}^f R_{t+1}^K - \lambda_t^f \right]$$

This component accounts for the mechanical effect on the economy's resources at large that follows a rise in capital taxes at story s^t. Capital taxes affect capital accumulation, influencing the amount of available resources in the whole economy. Therefore, an increase in the tax affects the feasibility constraints of two consequent periods. First, when the tax variation is decided since agents incorporate the information responding to the change in the tax when choosing optimal consumption and saving. Then, it affects the feasibility constraint of the subsequent period to an extent proportional to the return on risky capital. The difference between the two terms controls part of the optimal smoothing behavior of capital taxes. Indeed, a negative shock at the story s^t increases the multiplier at that time (the social welfare effect of a marginal relaxation on the feasibility constraint), reducing the motives of taxation. Intuitively, the government expects the shock to be transitory and the economy to grow in the next period. Therefore it decides to shift the tax burden accordingly, moving resources from one period to another. This two-period dynamic also informs the optimal capital taxes behavior in the secular trend. A government expecting positive growth in the future prefers to shift the tax burden away from the present and progressively increase taxes over time. By the same argument, when expecting a slowdown in the economy, such an increasing path becomes more concave, increasing the smoothness of the tax burden.

Intertemporal component. The component of the optimal capital tax that accounts for intertemporal taxation motives is given by

$$T_t^I\left(\lambda_t^y, \lambda_{t+1}^o\right) = \frac{\mathrm{E}_t\left[\beta \lambda_{t+1}^o R_{t+1}^K\right] - \lambda_t^y u_{c,t}^y}{n_t^y} \tag{6.3}$$

This term highlights motives for capital taxation that stem from the agents' intertemporal consumption tradeoff. It captures the incidence of an increase capital taxes on the taxed generation and its distortions on the life-time consumption-saving path. Its structure recalls the one of the aggregate resource component, although it focuses on a single generation. Adding an OLG structure to the model, requires the government to take into account the effect of a capital tax not only on aggregate resources, but also on the agents that directly pay for it. Indeed, increasing capital taxes shifts consumption from old to young agents of the taxed generation. Waiving a consumption unit today to accumulate capital relaxes the young budget constraint proportionally to the marginal utility of consumption while mechanically impacting the old budget constraint proportionally to the risky capital return. Such a relationship pinpoints the social value of this intertemporal shifting as an accounting association between the Lagrange multipliers on the two households' constraints.

Hedging component. The hedging term reads as

$$T_t^H\left(\lambda_{t+1}^o\right) = -\beta\, \mathrm{E}_t\left[(1-\tau_t^K) q_t \gamma_{t+1} F_{KK,t+1} u_{c,t+1}^o\right] C^H\left(\lambda_{t+1}^o\right)$$

where

$$C^H\left(\lambda_{t+1}^o\right) = \frac{\mathrm{Cov}_t\left[\lambda_{t+1}^o, q_t \hat{F}_{K,t+1}\right]}{\mathrm{E}_t\left[q_t \hat{F}_{K,t+1} u_{c,t+1}^o\right]} - \frac{\mathrm{Cov}_t\left[\lambda_{t+1}^o, q_t \gamma_{t+1} F_{KK,t+1}\right]}{\mathrm{E}_t\left[q_t \hat{F}_{K,t+1} u_{c,t+1}^o\right]}$$

This component echoes the one derived by Farhi (2010). The term in parenthesis is proportional to the inverse of the elasticity of capital to capital taxes. The higher is this elasticity, the lower will be the tax rate's absolute value.

The term C^H characterizes a motive for capital taxation that attains with smoothing the effects of a tax increase on the agents' marginal consumption across states (hence, hedging). This term balances the opposite direct and indirect impact of the tax increase, consisting of the difference between two covariances. The first measures the relationship between the direct effect of increased capital taxes on investments and the multiplier on the old budget constraint. A larger correlation implies a lower optimal τ_t^K because of the depressive effects on retirees' consumption. The other covariance accounts for an indirect effect of an increase in the tax, which distorts labor and investment allocations to a magnitude pinned down by the covariance between the size of the tax base adjustment (i.e., the derivative of the marginal product of the tax base F_{KK}) and the multiplier on the old budget constraint.

6.2 Capital taxes with heterogeneous agents and no social security

We now introduce agents' heterogeneity while keeping the government from using social security instruments. We characterize the composition of the optimal capital taxes at the Ramsey equilibrium in the following Proposition.

Proposition 7 (Optimal capital taxes with heterogeneous agents and no social security)**.** Suppose, without loss of generality, there exists two agents' types $\theta > \theta'$ for each generation and that the government does not have access to social security instruments. Then, optimal capital taxes at the story s^t, with $t \geq 1$, read as

$$\tau_t^K = \sum_\theta \frac{1}{B_{\theta,t}(\lambda^o_{\theta,t+1})} \left[T^A_{\theta,t}\left(\lambda^f_t, \lambda^f_{t+1}\right) + T^I_{\theta,t}\left(\lambda^y_{\theta,t}, \lambda^o_{\theta,t+1}\right) + T^H_{\theta,t}\left(\lambda^o_{\theta,t+1}\right) + T^R_{\theta,t}\left(\lambda^K_{\theta,\theta',t}\right)\right] \Theta_t(\theta) \tag{6.4}$$

where $B_{\theta,t}(\cdot)$, $T^A_{\theta,t}(\cdot)$, $T^I_{\theta,t}(\cdot)$ and $T^H_{\theta,t}(\cdot)$ are the same terms of the Ramsey optimal tax in the homogeneous case (6.1), where we use the θ-specific quantities or prices for all quantities and prices that can be θ-specific; and $T^R_{\theta,t}(\lambda^K_{\theta,\theta',t})$ is a redistributional component.

Proof. The proof can be found in Appendix A.5.2. □

Redistributional component The redistributional component in (6.1) is given by

$$T^R_{\theta,t}(\lambda^K_{\theta,\theta',t}) = \frac{(1-\tau_t^K)}{r_t^b} \lambda^K_{\theta,\theta',t} \Theta_t(\theta)\, C(\theta,\theta') \tag{6.5}$$

where

$$C(\theta,\theta') = \frac{\text{Cov}_t\left[\gamma_{t+1} F_{KK,t+1}, u^o_{c,\theta,t+1}\right]}{u^o_{c,\theta,t+1}} - \frac{\text{Cov}_t\left[\gamma_{t+1} F_{KK,t+1}, u^o_{c,\theta',t+1}\right]}{u^o_{c,\theta',t+1}}$$

As with labor taxes, by allowing for agents' heterogeneity, we introduce an additional motive of capital taxation, i.e., one for redistribution. The first term in equation 6.5 illustrates how the government increases distortions on the θ-type agent. If $\theta > \theta'$, the concavity in preferences implies $\lambda^K_{\theta,\theta',t} > 0$. This, in turn, generates taxation motives to impose distortions on the more productive type. The magnitude of such a distortion is pinned down by a covariance term $C(\theta,\theta')$ which measures how differently the two types respond to the indirect adjustments of the aggregate tax base that follows an increase in capital taxes. In other words, each covariance term sizes how the θ-specific marginal utility of consumption is sensitive to changes in the return on risky capital caused by changes in the tax rate. Therefore, a positive covariance means that the change in the yield is positive in the states for which it generates positive marginal utility. How differently the two types respond to such a change pinpoints the extent of the redistributional intervention through capital taxes. If the covariance for the lower type θ' is higher than the one for θ, motives for reducing capital taxes stem since their marginal increase would harm the former more than the latter. Notice that the covariance term would be zero under agents' homogeneity, thus shutting down the entire component in that case.

6.3 Capital taxes with homogeneous agents and social security

In this subsection we revert to the homogeneous case while letting the government use the social security instruments. The following Proposition shows how some of the motives for capital taxation are crowded out by social security instruments.

Proposition 8. If agents are homogeneous and the government has access to the social security instruments, capital taxes at the Ramsey optimum are given by:

$$\tau_t^K = \frac{T_t^I\left(\lambda_t^y, \lambda_{t+1}^o\right)}{B_t(\lambda_{t+1}^o)} \tag{6.6}$$

where $T_t^I(\cdot)$ is given by (6.3) and $B_t(\cdot)$ by (6.2).

Proof. The proof can be found in Appendix A.5.3. □

Here, we show that the motives to tax capital boil down to the intertemporal component scaled by the welfare one when the government can levy taxes to fund the social security system. This is because $\tau^{\text{ss,F}}$ takes now care of the terms in (6.1) that do not appear in (6.6), i.e., the aggregate resources and the hedging components. Notice that these two components reflect aggregate motives, which the Ramsey planner can thus shift on social security taxes.

6.4 Capital taxes with heterogeneous agents and social security

Finally, we discuss here the case in which the government can impose social security taxes in a setting with heterogeneous agents. Under these assumptions, a new motive for capital taxes arises, substituting the hedging component of (6.4), as formalized in the following Proposition.

Proposition 9. Suppose, without loss of generality, there exists two agents' types $\theta > \theta'$ for each generation and that the government has access to social security instruments. Then, optimal capital taxes at the story s^t, with $t \geq 1$, read as

$$\tau_t^K = \sum_\theta \frac{1}{B_{\theta,t}(\lambda_{\theta,t+1}^o)} \left[T_{\theta,t}^A\left(\lambda_t^f, \lambda_{t+1}^f\right) + T_{\theta,t}^I\left(\lambda_{\theta,t}^y, \lambda_{\theta,t+1}^o\right) + T_{\theta,t}^R\left(\lambda_{\theta,\theta',t}^K\right) + T_{\theta,t}^{\text{ss}}\left(\lambda_{t+1}^f, \lambda_{t+2}^f\right)\right] \Theta_t(\theta)$$

where $B_{\theta,t}$, $T_{\theta,t}^A(\cdot)$, $T_{\theta,t}^I(\cdot)$, $T_{\theta,t}^R(\cdot)$ are the same terms of (6.4), and $T_{\theta,t}^{\text{ss}}(\cdot)$ is a funded social security component.

Proof. The proof can be found in Appendix A.5.4. □

Funded social security component The funded social security component is

$$T^{\text{ss}}_{\theta,t}\left(\lambda^f_{t+1},\lambda^f_{t+2}\right)=\mathrm{E}_t\left[\left((1-\delta)\lambda^f_{t+2}-\lambda^f_{t+1}\right)n^y_{t+1}\tau^{\text{ss,F}}_{t+1}\gamma_{t+1}F_{LK,t+1}\left(\sum_\theta \theta l_{\theta,t+1}\Theta_{t+1}(\theta)\right)\right]$$

Since the funded component of social security acts as a mandatory saving in risky capital for the households, it distorts their labor choices with effects on the amount of produced resources in the economy. This term shows that even considering this new channel, the government looks for tax smoothing when it expects adverse shocks. In fact, an increase in the capital tax rate at the story s^t has a mechanical effect on the aggregate income tax base of young agents at s^{t+1}. The magnitude of such an effect is pinned down by the overall wage change and the aggregate labor supply's measure. This shrink in the social security tax base affects the feasibility constraint at s^{t+1} since it reduces the funded component of social security in the economy and the feasibility constraint at s^{t+2} by an amount proportional to the net-of-depreciation rate. Intuitively, the Ramsey planner looks to smooth taxes across two periods when it expects a negative transitory shock by an argument similar to the one discussed for the aggregate resources component. With respect to that case, here, the timing is translated by one period due to the channel through which taxes are linked to the resource constraint, i.e., labor income. The expectation of a negative transitory shock at the story s^{t+1} reduces the motives for taxation at s^t to avoid a harmful distortion of the aggregate income tax base at the time of the shock. At the same time, the smaller the depreciation rate of capital, the higher the optimal capital taxes since the erosion effect on aggregate capital—which would call for higher tax smoothing—decreases.

7 Social security replacement rate

We now briefly turn the discussion to the optimal social security replacement rate and, in particular, to its relationship with the government's debt.

The first order condition of the Ramsey problem on κ^{ss}_t reads as

$$\sum_\theta w_{\theta,t}l_{\theta,t}\lambda^y_{\theta,t}u^y_{c,\theta,t}=\beta R^{b,\tau}_t\sum_\theta w_{\theta,t}l_{\theta,t}\,\mathrm{E}_t\left[\lambda^o_{\theta,t+1}\right] \tag{7.1}$$

The optimal replacement rate is chosen to balance, in aggregate, labor distortion today with risk-free capitalization tomorrow. Indeed, an increase in κ^{ss}_t improves the long-term returns to labor and creates incentives for all θs to increase their labor supply. The size of this incentive is proportional to the current income $w_{\theta,t}l_{\theta,t}$ earned by each type and is rescaled by $\lambda^y_{\theta,t}$ to capture the amount of the distortion. The right-hand-side instead measures the future returns of increasing κ^{ss}, which depend on the current income and the value of transferring money to older workers as measured by $\lambda^o_{\theta,t+1}$. Importantly, returns are determined by the risk-free premium on bonds since the replacement rate acts in practice as a risk-free asset. This has to do with the timing structure of the social security system. In particular, the replacement rate is set one period in advance: for any amount of current income, agents know with certainty how much they will get in social security benefits when they are old. An immediate consequence of this observation is that the replacement rate and debt management must show some consistency at the optimum.

We show this equilibrium coordination between replacement rate and debt management by comparing the optimality conditions of the Ramsey problem for the two instruments. In particular, we find that equation (7.1) is implied by the first-order condition on the individual investment in bonds, which reads

$$\lambda^{y}_{\theta,t} = \frac{\mathrm{E}_t\left[\lambda^{o}_{\theta,t+1}\right]}{\psi_{t+1}\,\mathrm{E}_t\left[u^{o}_{c,\theta,t+1}\right]}.$$

This suggests that at the optimum the government sets the two rates following a non-arbitrage logic between the two risk-free assets to which agents have access.

8 Calibration

This section discusses how we calibrate the model outlined in Section 2, choosing the parameters' values and the parametric specification of the households' distribution. We let the model match targets from three relevant western economies: the United States, Italy, and the Netherlands. We chose these countries because we argue they represent an insightful example of different possible policy mixes. In particular, the USA has shown a recent trend of increasing debt-to-GDP ratio, with constant relatively low average income and capital taxes, and low social security replacement rate. On the contrary, Italy has one of the highest public debts in the world, supported by considerable tax rates. Moreover, it shows a high replacement rate. The Netherlands is characterized instead by a small stock of public debt, significant income taxes, average capital tax rates, and a large replacement rate. Moreover, the USA exhibits a sharply lower survival probability rate for the male population with respect to the other two countries.

In what follows, we illustrate how we calibrate every parameter in the model, breaking down our description by group. Table 1 presents the summary of the calibration exercise on the three benchmark economies.

Period length and discount rate We calibrate the model so that a period corresponds to 40 calendar years, and accordingly we set the discount factor β to 0.50, matching an yearly discount of 0.98 common in the literature.

Households' heterogeneity To calibrate the distribution of the discrete household types, we non-parametrically fit different quantiles of the per-hour labor income that we compute from the Luxembourg Income Study (LIS) database. In particular, we focus on the 2004 wave as the first pre-crisis year common to the three countries. Notice that, in our model, individual productivity is the only driver of wages' heterogeneity, thus this information is sufficient for our calibration purposes.

Demographics We assume that population growth is constantly equal to zero since we focus on the steady-state. We also assume a constant survival rate calibrated on the males' survival probabilities to 65 years old from the World Development Indicators of the World Bank. We obtain two very close values for Italy and the Netherlands—0.90 and 0.89, respectively—while the fraction of American males that reach 65 years old is, on average, only 0.80. Given the time structure of our model, this number well informs the probability for a given generation to reach the retirement period.

Government policies We calibrate average capital taxes paid by households on their saving using a dedicated institutional source for each country.[7] In particular, we set a 15% capital tax rate for the US given the information offered by the Internal Revenue Service (IRS) of the US Government; a 26% rate for Italy as provided by the parliamentary documentation on the financial revenues taxation[8]; and a 31% in the Netherlands as discussed in Klemm, Hebous and Waerzeggers (2021). All the three tax rates are taken from 2021 data.

For income taxes, we start from the tax wedge decomposition provided in the OECD's *Taxing wages* database (OECD, 2021*c*).[9] This is because, in our model, the total labor cost a firm incurs corresponds to the labor income of a household. Firms indeed do not make profits, nor do they pay taxes. Conceptually, they transfer all of their labor costs to workers' wages, who then pay income taxes. From the OECD data, we are able to observe the breakdown of the labor wedge in four relevant components: employer social security contributions, employee social security contributions, income taxes, and cash benefits. Given the structure of our model, the sum of all these four components constitutes the income tax rate we calibrate. Our income taxes calibration closely meets the results from Erosa, Fuster and Kambourov (2012) for all three countries.

To calibrate the replacement rate, we use data from OECD's *Pensions at a Glance* 2021 (OECD, 2021*b*), which reports the gross pension replacement rate by country as the ratio between the gross pension entitlement and the gross pre-retirement earnings. This number needs to be re-scaled as the percentage of our income tax base net of social security contributions and income taxes. Therefore, we compute our effective country-specific replacement rate as

$$\kappa^{\mathrm{ss}} = \kappa_{\mathrm{G}}^{\mathrm{ss}} \underbrace{(1 - EC)(1 - \tau_{\mathrm{E}}^{l})}_{\text{Net income}}$$

where $\kappa_{\mathrm{G}}^{\mathrm{ss}}$ is the gross replacement rate as reported in the data, EC is the employer's social security contribution as a fraction of the total labor cost, and τ_{E}^{l} is the effective labor tax levied on the average worker's income net of social security contributions. This way, we assign to each country an average replacement rate value that is fully coherent with our model and reads as 0.32 for the USA, 0.46 for Italy, and 0.51 for the Netherlands.

Finally, we set the funded component of social security to zero for all three countries. Indeed,

[7]Despite some works proposed estimated measures of savings taxes (see, e.g., the review in Sørensen and Sørensen 2004) there exists a strong model dependence in those estimates, as discussed by Hosseini and Shourideh (2019). Therefore, we opted for a parametrization that relies more on institutional documents than other academic works.

[8]In particular, we refer to the March 31st, 2021 "focus" accessible here.

[9]Notice that *tax wedge* here refers to "the ratio between the amount of taxes paid by an average single worker (a single person at 100% of average earnings) without children and the corresponding total labour cost for the employer. The average tax wedge measures the extent to which tax on labour income discourages employment," as the OECD documentation reports.

Parameter	Description	Value USA	ITA	NED	Source(s)
Demographics					
ψ	Survival probability (65 yo)	0.80	0.90	0.89	WB World Development Indicator (2019)
β	Discount factor	0.50	0.50	0.50	Common in literature
Government policies					
τ^K	Capital taxes	0.15	0.26	0.31	IRS Topic No. 409 (US); Parliamentary docs (ITA); Klemm, Hebous and Waerzeggers (2021) (NED)
τ^l	Labor taxes	0.28	0.46	0.36	OECD (2021*c*)
k^{ss}	Replacement rate	0.32	0.46	0.51	See text
$\tau^{ss,F}$	Funded component of social security	0	0	0	See text
Macroeconomic parameters					
B/Y	Debt to GDP ratio	1.61	1.83	0.66	OECD (2021*a*)
G/Y	Expenditure to GDP ratio	0.48	0.57	0.48	OECD (2021*a*)
Technology					
α	Capital share	0.36	0.36	0.36	Common in literature
δ	Capital depreciation rate	0.88	0.88	0.88	Hosseini and Shourideh (2019)

Table 1: Model's parameters

none of the countries we examine has a mandatory saving program to finance social security. In particular, there is no mandatory program for a funded pension plan in the USA and Italy. The Netherlands commits its employees and employers to pay contributions into pension funds, but these contributions are agreed upon in collective employment agreements—thus, not chosen by the government. This is why we opt to set $\tau^{ss,F} = 0$ even in this case.

Macroeconomic parameters We target two macroeconomic parameters: debt-to-GDP ratio and expenditure-to-GDP ratio. For both, we use data from the last available year in OECD's *Government at a Glance* 2021 (OECD, 2021*a*).

Preferences and technology For the numerical simulations, we use the same functional form as the household utility function adopted in Conesa, Kitao and Krueger (2009):

$$u(c,l) = \frac{\left(c^{\eta}(1-l)^{1-\eta}\right)^{1-\sigma}}{1-\sigma} \tag{8.1}$$

where η is a share parameter that tunes the relative importance of consumption to labor, and σ determines the household's risk aversion. We take $\eta = 0.18$ and $\sigma = 1.5$. Notice that, for $\sigma = 1$, this parametric form collapses on the standard log-log specification used in Chari, Christiano and Kehoe (1994), and later in Farhi (2010) and Hipsman (2018). We also assume a Cobb-Douglas production function with constant capital share of 0.36 across countries, which is standard in the literature. Moreover, we set the capital depreciation rate as equal to 0.88 for our period, to meet an annual rate of approximately 0.05 as in Hosseini and Shourideh (2019).

9 Quantitative results for the steady state

In this section, we present the results of our numerical simulation. We focus on an economy in a steady state without aggregate uncertainty. In this environment, the scope of debt is smaller since it cannot satisfy the households' need for safe assets, and it can be replicated by a combination of income and funded social security taxes. Since these simplifying assumptions reduce the policy space, we comment on our results focusing on allocations and welfare metrics rather than policy parameters. Moreover, we assume that the planner is utilitarian so that welfare weights are constant across all θs. We start by defining the main welfare metrics that we employ to quantify the improvements of the Ramsey optimum on the benchmark economies.

9.1 Welfare metrics

We quantify the welfare gains and losses of moving across two different policy scenarios with a measure of equivalent variation that keeps labor constant at the reference one. We call this measure a *consumption-equivalent variation*, and we quantify it as the percentage increase in consumption that each type would need to experience to be indifferent between the benchmark economy and the Ramsey optimum, given the constant labor. In a steady state, the levels of consumption and labor along the life cycle are constant across generations, so the definition of this measure simplifies to the following:

Definition 6 (Consumption-equivalent variation)**.** Denote with $(c^{y,B}_{\theta,t}, c^{o,B}_{\theta,t+1}, l^{y,B}_{\theta,t})$ the optimal allocation in a benchmark economy for the θ-type agents of a generation born at s^t, and $(c^{y,R}_{\theta,t}, c^{o,R}_{\theta,t+1}, l^{y,R}_{\theta,t})$ the optimal allocations for the same agents at the Ramsey optimum. Moreover, define $\tilde{c}^{y,B}_{\theta,t} = c^{y,B}_{\theta,t}(1+\Delta_\theta)$ and $\tilde{c}^{o,B}_{\theta,t+1} = c^{o,b}_{\theta,t+1}(1+\Delta_\theta)$ the Δ-augmented consumption for a given type in the benchmark scenario. Then, the consumption-equivalent variation for a type θ is a value of Δ_θ that satisfies

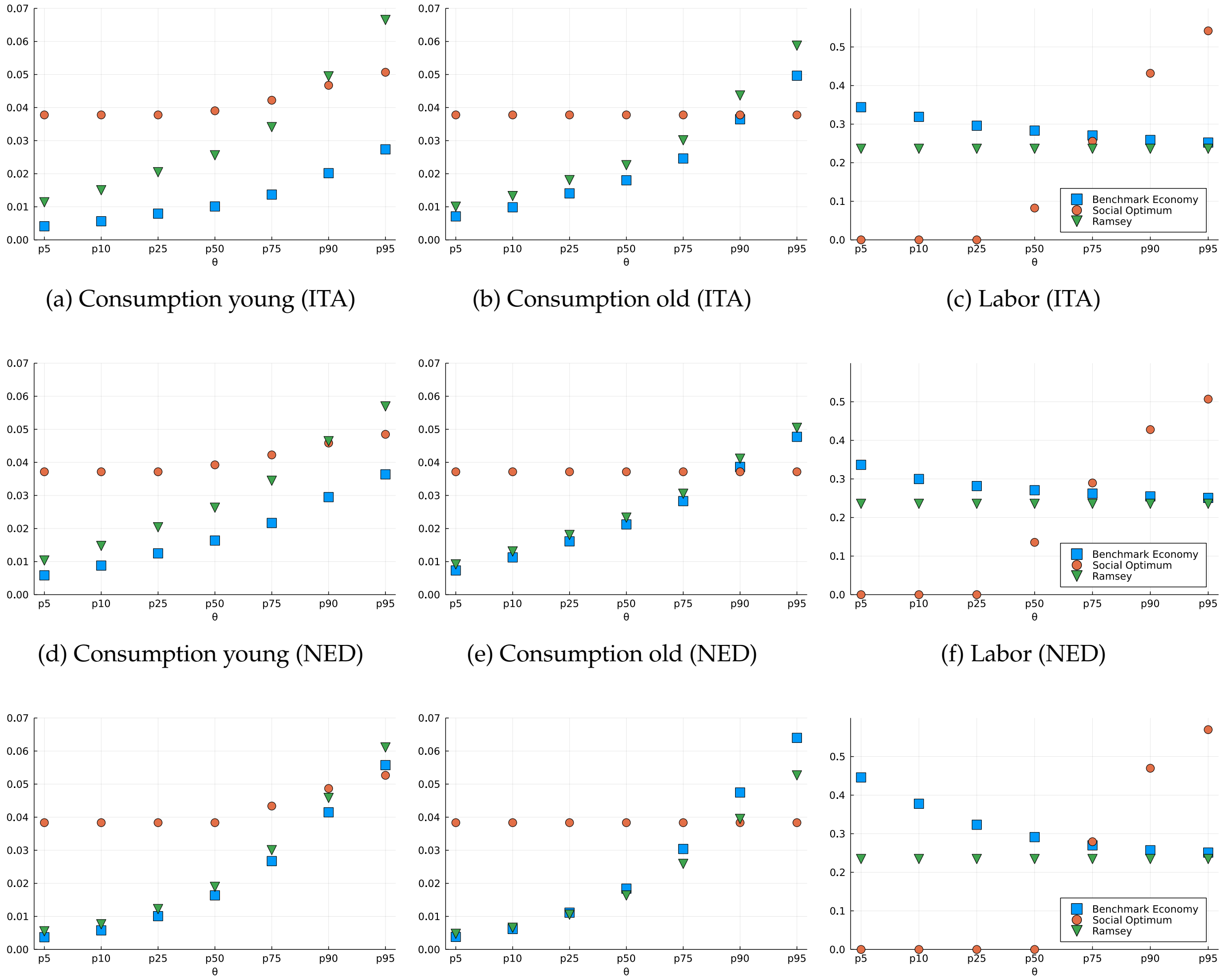
$$u\left(\tilde{c}^{y,B}_{\theta,t}, l^{y,B}_{\theta,t}\right) + \beta\psi_{t+1} u\left(\tilde{c}^{o,B}_{\theta,t+1}, 0\right) = u\left(c^{y,R}_{\theta,t}, l^{y,R}_{\theta,t}\right) + \beta\psi_{t+1} u\left(c^{o,R}_{\theta,t+1}, 0\right) \tag{9.1}$$

The consumption-equivalent variation measures the willingness to pay for each type to avoid moving from the benchmark economy to the Ramsey optimum. Thus, $\Delta_\theta > 0$ if the θ-type agent would be better off at the Ramsey optimum relative to the benchmark. Therefore, we can compare Δs across the types to understand the redistributive effects of moving between the two economies.

9.2 Numerical results: allocations and welfare

We start the discussion on our numerical simulations by comparing allocations in three regimes: *i)* the benchmark economy, *ii)* the Ramsey optimum, and *iii)* the social optimum. In particular, we study the components that determine agents' utility at the optimum, i.e., young and old consumption and labor. Moreover, our three different calibrations allow us to compare the allocations across countries to discuss differences and similarities which provide more insights about the model.

Figure 1 reports the allocations for the three countries in the three regimes. Overall, the benchmark economy and the Ramsey optimum tend to behave more similarly since they rely on the same



(a) Consumption young (ITA) (b) Consumption old (ITA) (c) Labor (ITA)

(d) Consumption young (NED) (e) Consumption old (NED) (f) Labor (NED)

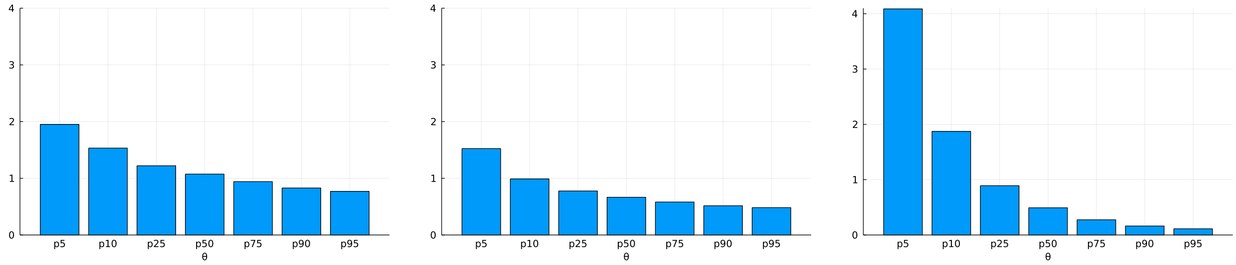
(g) Consumption young (USA) (h) Consumption old (USA) (i) Labor (USA)

Figure 1: Consumption allocations and labor: benchmark economy, Ramsey optimum, social optimum

Note: This panel shows the allocations of consumption of young agents (first column), old agents (second column) and labor supply (third column). Each row refers to a different country: Italy (ITA), the Netherlands (NED) and the United States (USA), from the top to the bottom. Each figure plots the benchmark economy allocation (blue square), the Ramsey optimum one (green triangle) and the social optimum one (orange circle).

set of instruments. On the other hand, the social optimum dominates the other two through two channels. First, it lets households enjoy more significant consumption when young and old for almost every type. Second, it manages to improve efficiency through an increasing pattern of labor supply over the type space so that more productive households work significantly more than less productive ones. In this calibration, the social optimum achieves a corner solution where no labor supply is required from the least productive types, making the labor-type profile very steep. On the contrary, the Ramsey optimum and the benchmark economies have flatter or even decreasing labor patterns due to limited instruments.

Young consumption increases across types in all countries in the benchmark economy so that wealthier households enjoy more significant consumption levels. This pattern is particularly accentuated due to low labor taxes in the US economy. At the same time, Italy is the country with the lowest inequality in consumption in the benchmark economy since it is the one where labor and capital taxes are larger. Moreover, at the Ramsey optimum young-age consumption is higher



(a) ITA

(b) NED

(c) USA

Figure 2: Consumption-equivalent variation between benchmark economy and Ramsey optimum

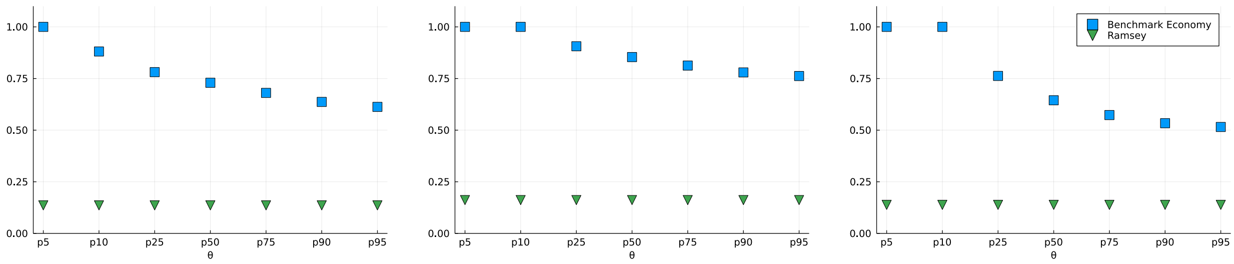
Note: This panel shows the consumption-equivalent variation, as stated in Definition 6, for three countries: Italy, the Netherlands and the US. Each figure shows the share of consumption a type should receive to be indifferent between moving to the Ramsey optimum of staying in the benchmark economy while keeping labor constant.

for all types in all countries, and the profile on the types' dimension is steeper. Italy has more considerable consumption gains in moving from the benchmark economy to the Ramsey optimum, especially for high-productivity households who enjoy sub-optimally low consumption levels in the status-quo. The consumption pattern at young age in the United States is instead remarkably close to the Ramsey optimum. In contrast with the limited-instrument scenarios, the social optimum displays a non-linear and convex consumption pattern due to the need to compensate high productivity types for their labor effort as prescribed by Equation (3.3).

Consumption at old age also increases in the agents' types and tends to be higher than young-age consumption for most of θs. The Ramsey optimum increases consumption for everyone making the pattern steeper except for the US, where old-age consumption is suboptimally high in the benchmark economy for the wealthiest types. This is due to the low capital tax rate that US households enjoy compared to the other two countries. At the social optimum, old consumption is flat, as suggested by equation (3.1).

Labor effort is flat across types at the Ramsey optimum. This is a well-known result in the case of log-log separable preferences in consumption and labor—a case we are not far from given our utility function's parametrization in (8.1). At the same time, the decreasing pattern in labor supply observed in the benchmark economies is explained by significant replacement rates that incentivize labor efforts from low productivity agents who strive to increase their consumption in retirement. As already discussed, the social optimum would require a steep pattern of labor supply to maximize efficiency.

We conclude our discussion by investigating the gains of a Ramsey optimal policy relative to the benchmark economy. Figure 2 reports the consumption-equivalent variations by type in the three countries, as described in the Definition 6. Overall, every household type benefits from the Ramsey optimum, as the variations' positivity suggests. In all countries, we observe a decreasing pattern in the willingness to pay, which implies that moving from the status quo to the Ramsey optimum would benefit poorer households more than richer ones. This result implicitly suggests that the current policy mix in the three countries is less fair to low-productivity households than a Ramsey optimum would prescribe. Notably, the redistributive gain for poorer households would be even larger if we solved the Ramsey optimum with decreasing welfare weights instead of focusing on the utilitarian case. The Netherlands is the country where gains seem to be more



(a) ITA (b) NED (c) USA

Figure 3: Social security benefit as a share of old consumption: benchmark economy and Ramsey optimum

Note: This panel shows the amount of social security benefit each household type receives, as a share of old-age consumption, under the benchmark economy (blue square) and at the Ramsey optimum (green triangle). The panel reports the exercise for three countries: Italy, the Netherlands, and the US.

equal across productivity types, which implies that the benchmark policy in the country is closer to the optimal level of redistribution that can be achieved in a utilitarian Ramsey equilibrium. The United States, on the other hand, penalizes the low-end of the productivity distribution in the benchmark calibration and emerge as the most unequal tax system. Italy seems to be placed between the other two countries.

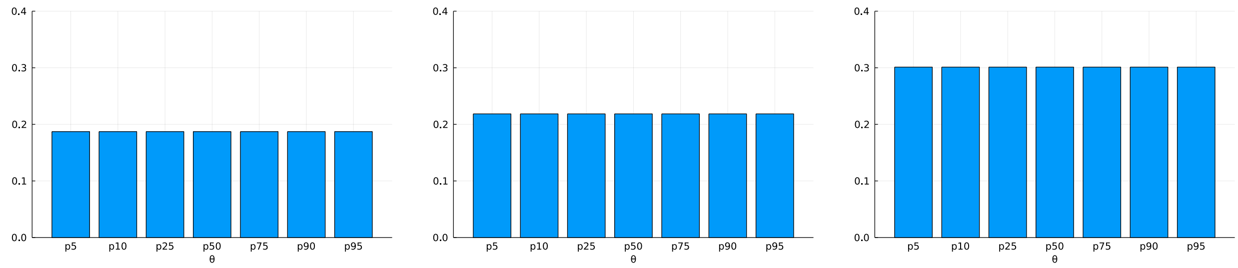
9.3 Numerical results: social security

We now turn the discussion to the results on social security, focusing on two aspects. First, we assess the generosity of the benefits system by quantifying the share of old-age consumption that relies on social security transfers. Second, we quantify the importance of the funded component by looking at what share of consumption young households save on funded social security.

Figure 3 shows that benefits in the benchmark economies exceed those at the Ramsey optimum in all three countries, suggesting that the status-quo social security systems tend to be too generous. The gap between the two scenarios is much larger in Netherlands and Italy than in the United States, which has a lower replacement rate in the benchmark economy. At the optimum, the Ramsey planner would indeed rely more on private savings rather than on social security. Moreover, the old-age consumption share funded by social security benefits is flat across household types in the Ramsey optimum, while it decreases in the benchmark economies. This is the byproduct of the labor supply patterns discussed in the previous section.

While the qualitative result that the systems seem to be more generous than what is prescribed by the Ramsey optimum is robust to alternative calibrations, the exact extent to which it is generous should be taken with some grain of salt. Indeed, our calibration is limited by the two-period structure of the OLG model, which implements an artificially low rate of work-to-retirement years that depresses the optimal replacement rate in the optimum. Our simulation allows agents to enjoy retirement for the same number of years they work, while a more realistic calibration with multiple periods per generation and a lower share of the life-cycle in retirement would certainly deliver a greater replacement rate in the Ramsey optimum.

Figure 4 displays our simulations for the funded component of social security, shown as the share of consumption during the working period. As discussed early, this pillar of the social



(a) ITA (b) NL (c) US

Figure 4: Social security funded contribution as a share of young consumption at the Ramsey optimum

Note: This panel shows the amount of funded social security distribution, as a share of young-age consumption, implied by the Ramsey optimum for each household type. The panel reports the exercise for three countries: Italy, the Netherlands, and the US.

security system is absent in all three countries in our benchmark economies calibration. Instead, the Ramsey optimum prescribes a positive funded social security tax and induces the agents to save a positive and reasonably sized share of their consumption when young. This share is wider in the US compared to the Netherlands and Italy because the US exhibits a lower survival rate to retirement age. Thus, young agents face incentives to invest less, which in turn causes a suboptimal capital level in aggregate. Since the funded social security contribution acts as a mandatory saving plan, the Ramsey planner implements positive levels of this instrument when disincentives to investment are more significant. At the same time, the planner aims to subsidize capital through negative capital taxes, as in Saitto (2020). The share of funded contribution is also flat across household types since both consumption and income are proportional to the agents' productivity. Taken together, these results suggest that a funded component in the social security system is desirable for all three countries.

10 Conclusions

This paper has provided a theoretical and quantitative analysis of optimal Ramsey taxation when the households are heterogeneously productive, the economy faces aggregate shocks, and the government has access to social security instruments. To model these instruments insightfully, we allow them to account for both a defined benefit and a defined contribution scheme and to be financed through a mix of funded and unfunded systems.

Our theoretical results show that introducing such a pension scheme changes optimal labor taxation smoothing across periods along two directions. First, it erodes the income effects of a tax rate increase since the replacement rate reduces the incentives for a labor supply adjustment. Second, it adds new taxation motives that come from redistributional objectives. At the same time, social security impacts capital taxes crowding out the taxation motives related to aggregate distortions and bringing in a new motive linked to the change in labor supply caused by a change in the capital tax rate. Moreover, we argue that the structure of our risk-free capital assets calls for equilibrium coordination between the social security replacement rate and the public debt.

To keep our numerical analysis tractable, we have focused on the deterministic steady state of three economies: the benchmark one, the Ramsey optimum, and the social optimum. Calibrating our model on three countries (the US, Netherlands, and Italy) shows space for redistributive and efficiency gains by moving from the status-quo allocations to the Ramsey optimal ones. In particular, we show that the Ramsey-optimal social security benefits are lower than the actual policies in all the three economies. At the same time, optimal taxes include non-zero social security funded component that the government uses to increase aggregate capital.

Two limitations of our work are the low number of generations in our model and the linear structure of the policies we consider. Our OLG model features two periods, thus unnaturally weighting the working period as much as the retirement one—a modeling choice that artificially depresses the optimal replacement rates. Moreover, linear income taxes allow for higher tractability at the expense of lower efficacy in dealing with redistributive motives. Thus, we plan to include non-linear income taxes and add multiple periods for each generation to refine the policy prescriptions on social security in a future version of this work.

References

Aiyagari, S. Rao. 1995. "Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting." *Journal of political Economy*, 103(6): 1158–1175.

Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppälä. 2002. "Optimal taxation without state-contingent debt." *Journal of Political Economy*, 110(6): 1220–1254.

Atkinson, Anthony B., and Agnar Sandmo. 1980. "Welfare implications of the taxation of savings." *The Economic Journal*, 90(359): 529–549.

Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy*, 87(5, Part 1): 940–971.

Bonenkamp, Jan, Lex Meijdam, Eduard Ponds, and Ed Westerhout. 2017. "Ageing-driven pension reforms." *Journal of Population Economics*, 30(3): 953–976.

Chamley, Christophe. 1986. "Optimal taxation of capital income in general equilibrium with infinite lives." *Econometrica: Journal of the Econometric Society*, 607–622.

Chancel, Lucas, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman. 2021. "World inequality report 2022."

Chari, V. V., Lawrence J. Christiano, and Patrick J. Kehoe. 1994. "Optimal Fiscal Policy in a Business Cycle Model." *Journal of Political Economy*, 102(4): 617–652.

Ciurila, Nicoleta. 2017. "How should pension benefits be linked to pension contribution." *Mimeo*.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing capital? Not a bad idea after all!" *American Economic Review*, 99(1): 25–48.

Dieppe, Alistair. 2021. *Global productivity: Trends, drivers, and policies.* World Bank Publications.

Erosa, Andrés, and Martin Gervais. 2002. "Optimal taxation in life-cycle economies." *Journal of Economic Theory*, 105(2): 338–369.

Erosa, Andrés, Luisa Fuster, and Gueorgui Kambourov. 2012. "Labor supply and government programs: A cross-country analysis." *Journal of Monetary Economics*, 59(1): 84–107.

Farhi, Emmanuel. 2010. "Capital taxation and ownership when markets are incomplete." *Journal of Political Economy*, 118(5): 908–948.

Farhi, Emmanuel, and Iván Werning. 2010. "Progressive estate taxation." *The Quarterly Journal of Economics*, 125(2): 635–673.

Garriga, Carlos. 2019. "Optimal fiscal policy in overlapping generations models." *Public Finance Review*, 47(1): 3–31.

Harenberg, Daniel, and Alexander Ludwig. 2019. "Idiosyncratic risk, aggregate risk, and the welfare effects of social security." *International Economic Review*, 60(2): 661–692.

Hipsman, Nathaniel E. 2018. "Optimal Taxation in Overlapping Generations Economies with Aggregate Risk." *Mimeo.*

Hosseini, Roozbeh, and Ali Shourideh. 2019. "Retirement financing: An optimal reform approach." *Econometrica*, 87(4): 1205–1265.

İmrohoroğlu, Ayşe, Selahattin İmrohoroğlu, and Douglas H. Joines. 1998. "The effect of tax-favored retirement accounts on capital accumulation." *American Economic Review*, 749–768.

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

Klemm, Alexander D, Shafik Hebous, and Christophe J Waerzeggers. 2021. "Capital Income Taxation in the Netherlands." *IMF Working Papers*, 2021(145).

Kotlikoff, Laurence J. 2009. "The Intergenerational Conflict Après le Deluge." In *Growth, Demography, Finance: From Economic Breakdowns to New Balances, Actes des Rencontres Economiques dAix-en-Provence*. Le Cercle des Economistes.

Krueger, Dirk, and Felix Kubler. 2006. "Pareto-improving social security reform when financial markets are incomplete!?" *American Economic Review*, 96(3): 737–755.

Lucas, Robert E., Jr., and Nancy L. Stokey. 1983. "Optimal fiscal and monetary policy in an economy without capital." *Journal of Monetary Economics*, 12(1): 55–93.

OECD. 2019. *Pensions at a Glance 2019.* OECD Publishing, Paris.

OECD. 2021*a*. *Government at a Glance 2021.* OECD Publishing, Paris.

OECD. 2021*b*. *Pensions at a Glance 2021.* OECD Publishing, Paris.

OECD. 2021*c*. *Taxing Wages 2021.* OECD Publishing, Paris.

Piketty, Thomas, and Emmanuel Saez. 2013. "A theory of optimal inheritance taxation." *Econometrica*, 81(5): 1851–1886.

Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *The Economic Journal*, 37(145): 47–61.

Saitto, Roberto. 2020. "Optimal taxation in an overlapping generations economy." *Mimeo*.

Sørensen, Peter Birch, and Peter Norman Sørensen. 2004. *Measuring the tax burden on capital and labor.* Vol. 6, MIT Press.

Werning, Ivan. 2007. "Optimal fiscal policy with redistribution." *The Quarterly Journal of Economics*, 122(3): 925–967.

Appendix A: Proofs and derivations

A.1 Households

The first order conditions for a θ-type household problem with respect to savings, borrowings and labor at time t, respectively, read as

$$\beta\psi_{t+1}\,\mathrm{E}_t\left[R^{K,\tau}_{t+1}u^{o}_{c,\theta,t+1}\right] = u^{y}_{c,\theta,t} \tag{A.1a}$$

$$\beta\psi_{t+1}R^{b,\tau}_{t}\,\mathrm{E}_t\left[u^{o}_{c,\theta,t+1}\right] = u^{y}_{c,\theta,t} \tag{A.1b}$$

$$\begin{aligned} u^{y}_{c,\theta,t}\left[1-\tau^{l}_{t}-\tau^{\mathrm{ss,F}}_{t}-\tau^{\mathrm{ss,U}}_{t}\right]w_{\theta,t} \\ +u^{y}_{l,\theta,t}+\beta\psi_{t+1}\kappa^{\mathrm{ss}}_{t}w_{\theta,t}\,\mathrm{E}_t\left[u^{o}_{c,\theta,t+1}\right] = 0 \end{aligned} \tag{A.1c}$$

First we notice that

$$R^{b,\tau}_{t} = \frac{u^{y}_{c,\theta,t}}{\beta\psi_{t+1}\,\mathrm{E}_t\left[u^{o}_{c,\theta,t+1}\right]} \qquad \text{and} \qquad \frac{u^{y}_{c,\theta,t}}{\beta\psi_{t+1}\,\mathrm{E}_t\left[R^{K,\tau}_{t+1}u^{o}_{c,\theta,t+1}\right]} = 1$$

Substituting the expression for $R^{K,\tau}_{t+1}$ at the optimum and rearranging the terms, we can express capital taxes as

$$\beta\psi_{t+1}\left(1-\tau^{K}(s^{t})\right) = \frac{u^{y}_{c,\theta,t}-\beta\psi_{t+1}u^{o}_{c,\theta,t+1}}{\mathrm{E}_t\left[\left(\gamma_{t+1}F_{K,t+1}-\delta\right)u^{o}_{c,\theta,t+1}\right]}$$

Moreover, the first order condition with respect to labor gives an expression for income taxes:

$$\left[1-\tau^{l}_{t}-\tau^{\mathrm{ss,F}}_{t}-\tau^{\mathrm{ss,U}}_{t}\right]w_{\theta,t} = -\frac{u^{y}_{l,\theta,t}+\beta\psi_{t+1}\kappa^{\mathrm{ss}}_{t}w_{\theta,t}\,\mathrm{E}_t\left[u^{o}_{c,\theta,t+1}\right]}{u^{y}_{c,\theta,t}}$$

A.2 Social planner

The first order conditions of the social planner problem with respect to young's consumption, old's consumption, capital and labor, respectively, read as:

$$g(\theta)\,u^{y}_{c,\theta,t} = \frac{\eta_t}{\phi_t}$$

$$\beta g(\theta)\,u^{o}_{c,\theta,t} = \frac{\eta_t}{\phi_{t-1}}$$

$$\mathrm{E}_t\left[\eta_{t+1}\left(\gamma_{t+1}F_{K,t+1}+1-\delta\right)\right] = \eta_t$$

$$\frac{g(\theta)\,u^{y}_{l,\theta,t}}{\theta\gamma_t F_{L,t}} = -\frac{\eta_t}{\phi_t}$$

where η_t is the Lagrange multiplier of the feasibility constraint at time t.

Combining the first two, we obtain

$$\frac{u^{y}_{c,\theta,t}}{\beta u^{o}_{c,\theta,t}} = \frac{\phi_{t-1}}{\phi_t}$$

From the first and the third, we have

$$\frac{\phi_t}{\phi_{t+1}} \frac{u^y_{c,\theta,t}}{\mathrm{E}_t \left[u^y_{c,\theta,t+1} \left(1 + \gamma_{t+1} F_{K,t+1} - \delta\right) \right]} = 1$$

While the first and the fourth give the standard labor-leisure condition

$$-\frac{u^y_{l,\theta,t}}{u^y_{c,\theta,t}} = \theta \gamma_t F_{L,t}$$

A.3 Ramsey planner

A.3.1 Proof of Proposition 2

Proof. We start by using the first order conditions for the firm to write wages and capital returns as

$$\theta \gamma_t F_{L,t} = w_{\theta,t} \quad \text{and} \quad \gamma_t F_{K,t} = r^K_t.$$

Then we note that equations (4.5) and (4.4) are equivalent to the first order conditions (2.3) and (2.4) for the household problem. Substituting them into the budget conditions (2.1) and (2.2) we can easily obtain (4.1) and (4.2). Finally (4.3) is equivalent to (2.7) after we substitute (2.5). Proceeding as in Section 4, we write the Lagrangian of the Ramsey problem as:

$$\begin{aligned}
\mathcal{L} = {} & \phi_0 \beta \psi_0 n^o_0 \sum_\theta g(\theta)\, u\left(c^0_\theta, 0\right) \Theta_t(\theta) + \mathrm{E}_0 \left[\sum_{t \geq 0} \phi_t n^y_t \sum_\theta g(\theta)\, U\left(c^y_{\theta,t}, l_{\theta,t}, c^o_{\theta,t+1}\right) \Theta_t(\theta) \right] \\
& - \mathrm{E}_0 \left[\sum_{t \geq 0} \sum_\theta \lambda^y_{\theta,t} \left[u^y_{l,\theta,t} l_{\theta,t} + u^y_{c,\theta,t} \left[c^y_{\theta,t} + q_{\theta,t} + b_{\theta,t} \right] + \beta \psi_{t+1} \kappa^{\mathrm{ss}}_t \theta \gamma_t F_{L,t} l_{\theta,t}\, \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right] \right] \right] \\
& - \mathrm{E}_0 \left[\sum_{t \geq 0} \sum_\theta \beta \lambda^o_{\theta,t+1} \left[c^o_{\theta,t+1} - \frac{u^y_{c,\theta,t}}{\beta \psi_{t+1} \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right]} b_{\theta,t} - \kappa^{\mathrm{ss}}_t \theta \gamma_t F_{L,t} - q_{\theta,t} \left(1 + \frac{u^y_{c,\theta,t} - \beta \psi_{t+1} \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right]}{\beta \psi_{t+1} \mathrm{E}_t \left[\hat{F}_{K,t+1} u^o_{c,\theta,t+1} \right]} \hat{F}_{K,t+1} \right) \right] \right] \\
& - \mathrm{E}_0 \left[\sum_{t \geq 0} \lambda^f_t \left[n^y_t \sum_\theta c^y_{\theta,t} \Theta_t(\theta) + n^o_t \sum_\theta c^o_{\theta,t} \Theta_t(\theta) + G_t - \gamma_t F\left(K_{t-1}, L_t\right) + n^y_t \sum_\theta q_{\theta,t} \Theta_t(\theta) \right. \right. \\
& \left. \left. + n^y_t \tau^{\mathrm{ss,F}}_t \gamma_t F_{L,t} \sum_\theta \theta l_{\theta,t} \Theta_t(\theta) - (1-\delta) \left[n^y_{t-1} \sum_\theta q_{\theta,t-1} \Theta_{t-1}(\theta) + n^y_{t-1} \tau^{\mathrm{ss,F}}_{t-1} \gamma_{t-1} F_{L,t-1} \sum_\theta \theta l_{\theta,t-1} \Theta_{t-1}(\theta) \right] \right] \right] \\
& - \mathrm{E}_0 \left[\sum_{t \geq 0} \sum_\theta \sum_{\theta' \neq \theta} \lambda^K_{\theta,\theta',t} \left[\frac{u^y_{c,\theta,t} - \beta \psi_{t+1} \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right]}{\beta \psi_{t+1} \mathrm{E}_t \left[\hat{F}_{K,t+1} u^o_{c,\theta,t+1} \right]} - \frac{u^y_{c,\theta',t} - \beta \psi_{t+1} \mathrm{E}_t \left[u^o_{c,\theta',t+1} \right]}{\beta \psi_{t+1} \mathrm{E}_t \left[\hat{F}_{K,t+1} u^o_{c,\theta',t+1} \right]} \right] \right] \\
& - \mathrm{E}_0 \left[\sum_{t \geq 0} \sum_\theta \sum_{\theta' \neq \theta} \lambda^L_{\theta,\theta',t} \left[\frac{u^y_{l,\theta,t} + \beta \psi_{t+1} \kappa^{\mathrm{ss}}_t F_{L,t} \mathrm{E}_t \left[u^o_{c,\theta,t+1} \right]}{\theta \gamma_t F_{L,t} u^y_{c,\theta,t}} - \frac{u^y_{l,\theta',t} + \beta \psi_{t+1} \kappa^{\mathrm{ss}}_t F_{L,t} \mathrm{E}_t \left[u^o_{c,\theta',t+1} \right]}{\theta \gamma_t F_{L,t} u^y_{c,\theta',t}} \right] \right]
\end{aligned} \tag{A.2}$$

□

A.4 The deadweight loss from labor taxation

A.4.1 Proof of Lemma 1

Proof. An individual θ solves

$$\begin{aligned} \underset{c^y_{\theta,t},\, c^o_{\theta,t+1}}{\text{maximize}} \quad & u\left(c^y_{\theta,t}, l_{\theta,t}\right) + \beta\, \mathrm{E}_t\left[u\left(c^o_{\theta,t+1}, 0\right)\right] \\ \text{subject to} \quad & c^y_{\theta,t} + q_{\theta,t} + b_{\theta,t} \le \left[1 - \tau^l_t - \tau^{\text{ss,F}}_t - \tau^{\text{ss,U}}_t\right] w_{\theta,t} l_{\theta,t} + Tr_t, \\ & c^o_{\theta,t+1} \le R^{K,\tau}_{t+1} q_{\theta,t} + R^{b,\tau}_t b_{\theta,t} + \kappa^{\text{ss}}_t w_{\theta,t} l_{\theta,t} \end{aligned}$$

where Tr_t are lump-sum transfers used by the government to compensate young agents' utility. Now, define $V\left(c^y_{\theta,t}, l_{\theta,t}, c^o_{\theta,t}\right)$ the indirect utility of the individual θ. From the envelope theorem, $V_{Tr} \equiv \partial V/\partial Tr = \mu^y_\theta$, where $\mu^y_\theta = u^y_{c,\theta,t}$ is the Lagrange multiplier on the young budget constraint. It follows that $\partial V/\partial \tau^l_t = -w_{\theta,t} l_{\theta,t} \mu^y_\theta = -w_{\theta,t} l_{\theta,t} \partial V/\partial Tr_t$.

Consider the planner's problem

$$\begin{aligned} \text{maximize} \quad & \phi_0 + \beta\psi_0 n^o_0 \sum_\theta g(\theta) u\left(c^o_{\theta,0}, 0\right) \Theta_t(\theta) \\ & + \mathrm{E}_0\left[\sum_{t\ge 0} \phi_t n^y_t \sum_\theta g(\theta) V\left(c^y_{\theta,t}, l_{\theta,t}, c^o_{\theta,t}\right) \Theta_t(\theta)\right] \\ \text{subject to} \quad & \\ n^o_t R^{b,\tau}_{t-1} \sum_\theta b_{\theta,t-1}\Theta_t(\theta) + G_t \le\ & n^y_t \sum_\theta \left[1 - \tau^l_t - \tau^{\text{ss,F}}_t - \tau^{\text{ss,U}}_t\right] w_{\theta,t} l_{\theta,t} \Theta_t(\theta) \\ & - n^y_t \sum_\theta b_{\theta,t}\Theta_t(\theta) - n^o_t \tau^K_{t-1} r^K_t \sum_\theta q_{\theta,t-1}\Theta_t(\theta) \\ & - n^o_t \tau^{\text{ss,F}}_{t-1} R^{K,\tau}_t) \sum_\theta w_{\theta,t-1} l_{\theta,t-1} \Theta_t(\theta) \end{aligned}$$

with the associated Lagrangian

$$\begin{aligned} \mathcal{L} = &\ \phi_0 + \beta\psi_0 n^o_0 \sum_\theta g(\theta) u(c^o_{\theta,0})\Theta_t(\theta) + \mathrm{E}_t\left[\sum_{t\ge 0} \phi_t n^y_t \sum_\theta g(\theta) V\left(c^y_{\theta,t}, l_{\theta,t}, c^o_{\theta,t}\right) \Theta_t(\theta)\right] \\ & - E_0\Bigg[\sum_{t\ge 0} \lambda_t \Bigg[n^o_t R^{b,\tau}_{t-1} \sum_\theta b_{\theta,t-1}\Theta_t(\theta) + G_t \\ & - n^y_t \sum_\theta \left[1 - \tau^l_t - \tau^{\text{ss,F}}_t - \tau^{\text{ss,U}}_t\right] w_{\theta,t} l_{\theta,t} \Theta_t(\theta) - n^y_t \sum_\theta b_{\theta,t-1}\Theta_t(\theta) \\ & n^o_t \tau^K_{t-1} r^K_t \sum_\theta q_{\theta,t-1}\Theta_t(\theta) - n^o_t \tau^{\text{ss,F}}_{t-1} R^{K,\tau}_t) \sum_\theta w_{\theta,t-1} l_{\theta,t-1} \Theta_t(\theta)\Bigg]\Bigg] \end{aligned}$$

Naming the different tax basis as $Y_{\theta_t} = w_{\theta,t} l_{\theta,t}$, $Y^{ss}_{\theta_t} = R^{K,\tau}_{t+1} w_{\theta,t} l_{\theta,t}$, $B_{\theta,t} = R^{b,\tau}_t b_{\theta,t}$, $Q_{\theta,t} = r^K_t q_{\theta,t-1}$, we can write the FOC for τ_t^l as

$$\phi_t n_t^y \sum_\theta g(\theta) \frac{\partial V\left(c_{\theta,t}^y, l_{\theta,t}, c_{\theta,t}^o\right)}{\partial \tau_t^l} \Theta_t(\theta) - E_t\Bigg[\overbrace{-\lambda n_t^y \sum_\theta Y_{\theta_t} \Theta_t(\theta)}^{\text{Mechanical revenue}}$$
$$+ \sum_{s\geq t} \lambda_s \sum_\theta \left(n_s^o \frac{\partial B_{\theta,s-1}}{\partial \tau_t^l} - n_s^y \left[1 - \tau_t^l - \tau_t^{\text{ss,F}} - \tau_t^{\text{ss,U}}\right] \frac{\partial Y_{\theta_s}}{\partial \tau_t^l} + n_s^y \frac{\partial b_{\theta,t}}{\partial \tau_t^l}\right.$$
$$\left. - n_s^o \tau_{t-1}^K \frac{\partial Q_{\theta,s-1}}{\partial \tau_t^l} + n_s^o \tau_{t-1}^{\text{ss,F}} \frac{\partial Y_{\theta,s-1}^{ss}}{\partial \tau_t^l} \right) \Theta_t(\theta) = 0$$

We can use the Slutsky equation and the envelope theorem on the individual's problem to obtain the following relationship

$$\underbrace{n_t^y \sum_\theta Y_{\theta_t} \left[-\frac{\phi_t g(\theta)}{\lambda_t} \frac{\partial V\left(c_{\theta,t}^y, l_{\theta,t}, c_{\theta,t}^o\right)}{\partial Tr_t} + 1 - \frac{\partial FE_{s\geq t}}{\partial Tr_t}\right] \Theta_t(\theta)}_{\text{Govt. value for decreasing an extra transfer to all } \theta\text{s}} = \underbrace{\frac{\partial FE_{s\geq t}^c}{\partial \tau_t^l}}_{\text{Revenue loss from compensated tax base change}} \tag{A.3}$$

where the compensated fiscal externalities are given by

$$\frac{\partial FE_{s\geq t}}{\partial Tr_t} = E_t\left[\sum_{s\geq t} \frac{\lambda_s}{\lambda_t} \left(n_s^o \sum_\theta \frac{\partial B_{\theta,s-1}}{\partial Tr_t} - n_t^y \sum_\theta \left[1 - \tau_t^l - \tau_t^{\text{ss,F}} - \tau_t^{\text{ss,U}}\right] \frac{\partial Y_{\theta_s}}{\partial Tr_t} \right.\right.$$
$$\left.\left. - n_s^y \sum_\theta \frac{\partial b_{\theta,t}}{\partial Tr_t} - n_s^o \tau_{t-1}^K \sum_\theta \frac{\partial Q_{\theta,s-1}}{\partial Tr_t} - n_s^o \tau_{t-1}^{\text{ss,F}} \sum_\theta \frac{\partial Y_{\theta,s-1}^{ss}}{\partial Tr_t} \right) \Theta_t(\theta) \right]$$

and

$$\frac{\partial FE_{s\geq t}^c}{\partial \tau_t^l} = E_t\left[\sum_{s\geq t} \frac{\lambda_s}{\lambda_t} \left(n_s^o \sum_\theta \frac{\partial B_{\theta,s-1}}{\partial \tau_t^l} - n_t^y \sum_\theta \left[1 - \tau_t^l - \tau_t^{\text{ss,F}} - \tau_t^{\text{ss,U}}\right] \frac{\partial Y_{\theta_s}}{\partial \tau_t^l} \right.\right.$$
$$\left.\left. - n_s^y \sum_\theta \frac{\partial b_{\theta,t}}{\partial \tau_t^l} - n_s^o \tau_{t-1}^K \sum_\theta \frac{\partial Q_{\theta,s-1}}{\partial \tau_t^l} - n_s^o \tau_{t-1}^{\text{ss,F}} \sum_\theta \frac{\partial Y_{\theta,s-1}^{ss}}{\partial \tau_t^l} \right) \Theta_t(\theta) \right]$$

Equation (A.3) equates the value of reducing the transfer of an extra unit of consumption to every young individual in t to the marginal excess burden of the tax. The latter is equal to the fiscal externality computed using the compensated responses of different tax bases, including general equilibrium price changes caused by the increase in τ^l. Since everything is normalized by λ_t, quantities are in terms of government revenues.

In our main Ramsey setup, the value of decreasing transfers to a young individual of type θ in t is captured by the multiplier $\lambda_{\theta,t}^y$, i.e., the value of relaxing the individual's implementability condition. To convert this value into government revenues, we normalize it by the marginal social value of increasing utility for a young individual of type θ in t:

$$\tilde{\mathcal{W}}_{\theta,t} = n_t^y \phi_t g(\theta) \Theta_t(\theta)$$

Finally, summing across the individuals, we obtain the total value of a reduction of one unit of transfer to all young agents:

$$\lambda^L(s^t) = -\sum_\theta \frac{\lambda^y_{\theta,t}}{\mathcal{W}_{\theta,t}}$$

which, in equilibrium, equates to the deadweight loss and is thus equivalent to the labor wedge. □

A.4.2 Proof of Proposition 3

Proof. This proposition follows from Equation (A.10) noting that in the quasi-linear case $u^y_{c,\theta,t} = u^o_{c,\theta,t+1} = 1$ and $u^y_{cc,\theta,t} = u^o_{cc,\theta,t+1} = u^y_{lc,\theta,t} = 0$. Hence in this case $\varepsilon^{R^{\tau,b}}_{b_{\theta,t}} = 0$ and $\frac{d\tau^K_t}{db_{\theta,t}} = 0$. □

A.4.3 Proof of Proposition 4

Proof. This proposition follows from Equation (A.10) noting that the terms in $\lambda^K_{\theta,\theta',t}$ and $\lambda^L_{\theta,\theta',t}$ will not appear since with homogeneous agents their terms vanish from the Lagrangian (A.2). □

A.4.4 Proof of Proposition 5

Proof. We prove the most general case (for two types θ and θ') in which the utility is not necessarily separable and then all the other cases will follow from this. To start, we compute the first order conditions of the Ramsey planner with respect to $c^y_{\theta,t}$ and $c^o_{\theta,t+1}$ by taking the appropriate derivatives of the Lagrangian (A.2). The first order condition with respect to $c^y_{\theta,t}$ reads:

$$\begin{aligned}
\lambda^f_t n^y_t \Theta_t(\theta) = {} & \phi_t n^y_t g(\theta)\Theta'_t(\theta) u^y_{c,\theta,t} - \lambda^y_{\theta,t}\left[\left(c^y_{\theta,t} + q^y_{\theta,t} + b^y_{\theta,t}\right) u^y_{cc,\theta,t} + u^y_{c,\theta,t} + u^y_{lc,\theta,t} l_{\theta,t}\right] \\
& + E_t\left[\lambda^o_{\theta,t+1} u^y_{cc,\theta,t}\left(\frac{(\gamma_{t+1}F_{K,t+1} - \delta)\, q^y_{\theta,t}}{\psi_{t+1}\,\mathrm{E}_t\left[(\gamma_{t+1}F_{K,t+1} - \delta)\, u^o_{c,\theta,t+1}\right]} + \frac{b^y_{\theta,t}}{\psi_{t+1}\,\mathrm{E}_t\left[u^o_{c,\theta,t+1}\right]}\right)\right] \\
& - \lambda^K_{\theta,\theta',t} \frac{u^y_{cc,\theta,t}}{\beta\psi_{t+1}\,\mathrm{E}_t\left[(\gamma_{t+1}F_{K,t+1} - \delta)\, u^o_{c,\theta,t+1}\right]} \\
& - \lambda^L_{\theta,\theta',t}\left[\frac{u^y_{lc,\theta,t}}{\gamma_t \theta F_{\tilde{L},t} u^y_{c,\theta,t}} - u^y_{cc,\theta,t} \frac{u^y_{l,\theta,t} + \beta\psi_{t+1}\kappa^{ss}_t \gamma_t \theta F_{\tilde{L},t} E_t\left[u^o_{c,\theta,t+1}\right]}{\gamma_t \theta F_{\tilde{L},t}\left(u^y_{c,\theta,t}\right)^2}\right]
\end{aligned} \tag{A.4}$$

while the one with respect to $c^o_{\theta,t}$ is:

$$\lambda^f_t n^o_t \psi_t \Theta_t(\theta) = \phi_{t-1} n^y_{t-1} g(\theta)\Theta'_{t-1}(\theta)\beta\psi_t u^o_{c,\theta,t} - \lambda^y_{\theta,t-1}\left(\beta\psi_t\kappa^{ss}_{t-1}\theta\gamma_{t-1}F_{\tilde{L},t-1}l_{\theta,t-1}p_{t|t-1}u^o_{cc,\theta,t}\right)$$
$$- \beta\lambda^o_{\theta,t} + E_{t-1}\left[\lambda^o_{\theta,t}\right]\frac{u^y_{c,\theta,t-1}u^o_{cc,\theta,t}}{\psi_t E_{t-1}\left[u^o_{c,\theta,t}\right]^2}b^y_{\theta,t-1}$$
$$- E_{t-1}\left[\left(\gamma_t F_{K,t}-\delta\right)\lambda^o_{\theta,t}\right]\left(\frac{\left(u^y_{c,\theta,t-1}-\beta\psi_t E_{t-1}\left[u^o_{c,\theta,t}\right]\right)u^o_{cc,\theta,t}}{\psi_t E_{t-1}\left[\left(\gamma_t F_{K,t}-\delta\right)u^o_{c,\theta,t}\right]^2}\right)\left(\gamma_t F_{K,t}-\delta\right)q^y_{\theta,t-1}$$
$$- E_{t-1}\left[\left(\gamma_t F_{K,t}-\delta\right)\lambda^o_{\theta,t}\right]\beta\frac{u^o_{cc,\theta,t}}{E_{t-1}\left[\left(\gamma_t F_{K,t}-\delta\right)u^o_{c,\theta,t}\right]}q^y_{\theta,t-1}$$
$$- \lambda^K_{\theta,\theta',t-1}E_{t-1}\left[\left[-\frac{u^o_{cc,\theta,t}}{E_{t-1}\left[\gamma_t\left(F_{K,t}-\delta\right)u^o_{c,\theta,t}\right]}\right.\right.$$
$$\left.\left.- \frac{u^y_{c,\theta,t-1}-\beta\psi_t E_{t-1}\left[u^o_{c,\theta,t}\right]}{\beta\psi_t\left(E_{t-1}\left[\left(\gamma_t F_{K,t}-\delta\right)u^o_{c,\theta,t}\right]\right)^2}\left(\gamma_t F_{K,t}-\delta\right)u^o_{cc,\theta,t}\right]\right]$$
$$- \lambda^L_{\theta,\theta',t-1}E_{t-1}\left[\frac{\beta\psi_t\kappa^{ss}_{t-1}\gamma_{t-1}\theta F_{\tilde{L},t-1}u^o_{cc,\theta,t}}{\gamma_{t-1}\theta F_{\tilde{L},t-1}u^y_{c,\theta,t-1}}\right] \tag{A.5}$$

In the following it will be useful to use the first order condition of the Ramsey planner with respect to $b_{\theta,t}$,

$$\lambda^y_{\theta,t} = \frac{E_t\left[\lambda^o_{\theta,t+1}\right]}{\psi_{t+1}E_t[u^o_{c,\theta,t+1}]} \tag{A.6}$$

and a similar identity that follows from it by using (A.1b):

$$\frac{\lambda^y_{\theta,t}u^y_{c,\theta,t}}{\beta R^{\tau,b}_t E_t\left[\lambda^o_{\theta,t+1}\right]} = 1 \tag{A.7}$$

Furthermore we recall the definition of the marginal social utility of an increase in consumption for a young worker of type θ:

$$\tilde{\mathcal{W}}_{\theta,t} = n^y_t\phi_t g(\theta)\Theta_t(\theta)u^y_{c,\theta,t} \tag{A.8}$$

From Equation (A.4) using $\tilde{\lambda}^y_{\theta,t} = \frac{\lambda^y_{\theta,t} u^y_{c,\theta,t}}{\tilde{\mathcal{W}}_{\theta,t}}$ we get:

$$\begin{aligned}
\lambda^f_t n^y_t \Theta_t(\theta) = {} & \tilde{\mathcal{W}}_{\theta,t}\left[1 - \tilde{\lambda}^y_{\theta,t}\left[-\frac{q^y_{\theta,t} + b^y_{\theta,t}}{c^y_{\theta,t}}\sigma^y_{\theta,t} + \eta^y_{\theta,t}\right]\right] \\
& + \tilde{\mathcal{W}}_{\theta,t} E_t\left[\frac{\lambda^o_{\theta,t+1}}{\mathcal{W}_{\theta,t}}\frac{u^y_{cc,\theta,t}}{u^y_{c,\theta,t}}\left(\frac{(\gamma_{t+1}F_{K,t+1} - \delta)\, q^y_{\theta,t}}{\psi_{t+1}\,\mathrm{E}_t\left[(\gamma_{t+1}F_{K,t+1} - \delta)\, u^o_{c,\theta,t+1}\right]} + \frac{b^y_{\theta,t}}{\psi_{t+1}\,\mathrm{E}_t\left[u^o_{c,\theta,t+1}\right]}\right)\right] \\
& - \tilde{\mathcal{W}}_{\theta,t}\tilde{\lambda}^K_t \frac{u^y_{cc,\theta,t}}{u^y_{c,\theta,t}\beta\psi_{t+1}\,\mathrm{E}_t\left[(\gamma_{t+1}F_{K,t+1} - \delta)\, u^o_{c,\theta,t+1}\right]} \\
& - \tilde{\mathcal{W}}_{\theta,t}\frac{\tilde{\lambda}^L_t}{u^y_{c,\theta,t}}\left[\frac{u^y_{lc,\theta,t}}{\gamma_t \theta F_{\tilde{L},t} u^y_{c,\theta,t}} - \sigma_{\theta,t}\frac{\left(1 - \tau^L_t\right)}{c^y_{\theta,t}}\right]
\end{aligned}$$

Where $\sigma^y_{\theta,t} = -\frac{u^y_{cc,\theta,t}}{u^y_{c,\theta,t}} c^y_{\theta,t}$ represents the inverse of the inter-temporal elasticity of substitution and we define $\eta^y_{\theta,t} = 1 + \frac{u_{lc,\theta,t} l_{\theta,t}}{u_{c,\theta,t}} - \sigma^y_{\theta,t}$. Using then Equations (A.6) and the households first order conditions (A.1a), and (A.1b) we get:

$$\begin{aligned}
\lambda^f_t n^y_t \Theta_t(\theta) = {} & \tilde{\mathcal{W}}_{\theta,t}\Bigg[1 - \tilde{\lambda}^y_{\theta,t}\left[-\frac{q^y_{\theta,t} + b^y_{\theta,t}}{c^y_{\theta,t}}\sigma^y_{\theta,t} + \eta^y_{\theta,t}\right] - \sigma^y_{\theta,t}\frac{b^y_{\theta,t}}{c^y_{\theta,t}}\frac{\beta R^{\tau,b}_t \psi_{t+1} E_t\left[u^o_{c,\theta,t+1}\right]}{u^y_{c,\theta,t}}\frac{\lambda^y_{\theta,t}}{\mathcal{W}_{\theta,t}} \\
& - \sigma^y_{\theta,t}\frac{q^y_{\theta,t}}{c^y_{\theta,t}}\frac{1}{\psi_{t+1} r^b_t E_t\left[u^o_{c,\theta,t+1}\right]} E_t\left[\frac{\lambda^o_{\theta,t+1}}{\mathcal{W}_{\theta,t}}(\gamma_{t+1}F_{K,t+1} - \delta)\right] \\
& + \tilde{\lambda}^K_t \frac{\sigma^y_{\theta,t}}{c^y_{\theta,t}}\frac{1}{\beta\psi_{t+1} r^b_t E_t\left[u^o_{c,\theta,t+1}\right]} - \frac{\tilde{\lambda}^L_t}{u^y_{c,\theta,t}}\left[\frac{u^y_{lc,\theta,t}}{\gamma_t \theta F_{\tilde{L},t} u^y_{c,\theta,t}} - \sigma_{\theta,t}\frac{\left(1 - \tau^L_t\right)}{c^y_{\theta,t}}\right]\Bigg]
\end{aligned}$$

where $\tilde{\lambda}^K_t = \frac{\lambda^K_{\theta,\theta',t} u^y_{c,\theta,t}}{\tilde{\mathcal{W}}_{\theta,t}}$, and $\tilde{\lambda}^L_t = \frac{\lambda^L_t u^y_{c,\theta,t}}{\tilde{\mathcal{W}}_{\theta,t}}$. After using (A.1a), and (A.6) we proceed collect $\tilde{\lambda}^y_{\theta,t}$ and we finally get:

$$\begin{aligned}
\lambda^f_t n^y_t \Theta_t(\theta) = \tilde{\mathcal{W}}_{\theta,t}\Bigg[1 - \tilde{\lambda}^y_{\theta,t}\Bigg[\eta^y_{\theta,t} + \sigma^y_{\theta,t}\frac{q^y_{\theta,t}}{c^y_{\theta,t}} E_t\left[\frac{\lambda^o_{\theta,t+1}(\gamma_{t+1}F_{K,t+1} - \delta)}{r^b_t E_t\left[\lambda^o_{\theta,t+1}\right]} - 1\right] \\
-\sigma^y_{\theta,t}\frac{\lambda^K_{\theta,\theta',t}}{c^y_{\theta,t}}\frac{1}{\beta r^b_t E_t\left[\lambda^o_{\theta,t+1}\right]} - \frac{\lambda^L_t}{R^{b,\tau}_t \beta E_t\left[\lambda^o_{\theta,t+1}\right]}\left[\frac{u^y_{lc,\theta,t}}{\gamma_t \theta F_{\tilde{L},t} u^y_{c,\theta,t}} - \sigma_{\theta,t}\frac{\left(1 - \tau^L_t\right)}{c^y_{\theta,t}}\right]\Bigg]\Bigg]
\end{aligned}$$

We now turn our attention to Equation (A.5). In order to facilitate the interpretation of this term we need the expressions for the elasticity of risk-free returns to issued bonds and the sensitivity of capital taxes to issued bonds:

$$\varepsilon^{R^{\tau,b}}_{b_\theta,t} = -R^{\tau,b}_t b_{\theta,t} \frac{E_t\left[u^o_{cc,\theta,t+1}\right]}{E_t\left[u^o_{c,\theta,t+1}\right]}, \quad \frac{d\tau^K_t}{db_t} = R^{\tau,b}_t \frac{E_t\left[u^o_{cc,\theta,t+1}\right] + \left(1-\tau^K_t\right) E_t\left[\left(\gamma_{t+1}F_{K,t+1}-\delta\right) u^o_{cc,\theta,t+1}\right]}{E_t\left[\left(\gamma_{t+1}F_{K,t+1}-\delta\right) u^o_{c,\theta,t+1}\right]} \tag{A.9}$$

The formulas above are obtained by computing the total derivative of (A.1a) and (A.1b) by keeping the young agents consumption constant and assuming that the change in the old agents consumption is given just by a change in $R^{\tau,b}_t$. After taking the expectation at time $t-1$, and rearranging some terms using the households first order conditions, Equation (A.5) becomes:

$$\begin{aligned}
E_{t-1}\left[\lambda^f_t\right] n^o_t \psi_t \Theta_t(\theta) = {} & \phi_{t-1} n^y_{t-1} g(\theta)\,\Theta_{t-1}(\theta)\,\beta\psi_t E_{t-1}\left[u^o_{c,\theta,t}\right] \\
& - \lambda^y_{\theta,t-1}\left(\beta\psi_t\kappa^{ss}_{t-1}\theta\gamma_{t-1}F_{\tilde{L},t-1} l_{\theta,t-1} p_{t|t-1} E_{t-1}\left[u^o_{cc,\theta,t}\right]\right) \\
& - \beta E_{t-1}\left[\lambda^o_{\theta,t}\right] - E_{t-1}\left[\lambda^o_{\theta,t}\right]\beta\varepsilon^{R^{\tau,b}}_{b_\theta,t-1} - E_{t-1}\left[\left(\gamma_t F_{K,t}-\delta\right)\lambda^o_{\theta,t}\right]\beta\frac{d\tau^K_{t-1}}{db_{\theta,t-1}}\frac{q^y_{\theta,t-1}}{R^{\tau,b}_{t-1}} \\
& + \frac{\lambda^K_{\theta,\theta',t-1}}{R^{\tau,b}_{t-1}}\frac{d\tau^K_{t-1}}{db_{\theta,t-1}} + \lambda^L_{\theta,\theta',t-1}\beta\psi_t\kappa^{ss}_{t-1}\frac{\varepsilon^{R^{\tau,b}}_{b_\theta,t-1}}{R^{\tau,b}_{t-1} b^y_{\theta,t-1}}
\end{aligned}$$

Using Equations (A.1b), (A.6), (A.7), and (A.8) we get:

$$\begin{aligned}
E_{t-1}\left[\lambda^f_t\right] n^o_t \psi_t \Theta_t(\theta) = {} & \frac{\tilde{\mathcal{W}}_{\theta,t-1}}{R^{b,\tau}_{t-1}}\left[1 - \tilde{\lambda}^y_{\theta,t-1}\left[1 + \varepsilon^{R^{\tau,b}}_{b_\theta,t-1}\left(1 - \frac{1}{R^{b,\tau}_{t-1}}\frac{\kappa^{ss}_{t-1}}{b^y_{\theta,t-1}}\left(z_{\theta,t-1}p_{t|t-1} + \psi_t\frac{\lambda^L_{\theta,\theta',t-1}}{R^{\tau,b}_{t-1}E_{t-1}\left[\lambda^o_{\theta,t}\right]}\right)\right)\right.\right. \\
& \left.\left. + \frac{d\tau^K_{t-1}}{db_{\theta,t-1}}\frac{q^y_{\theta,t-1}\beta E_{t-1}\left[\left(\gamma_t F_{K,t}-\delta\right)\lambda^o_{\theta,t}\right] - \lambda^K_{\theta,\theta',t-1}}{\beta R^{\tau,b}_{t-1}E_{t-1}\left[\lambda^o_{\theta,t}\right]}\right]\right]
\end{aligned}$$

where we used the notation $z_{\theta,t-1} = \theta\gamma_{t-1}F_{\tilde{L},t-1}l_{\theta,t-1}$. Finally we can put together the two expressions we found and get:

$$\frac{\tilde{\mathcal{W}}_{\theta,t}}{R_t^{b,\tau}}\left[1-\tilde{\lambda}^y_{\theta,t}\left[1+\varepsilon^{R^{\tau,b}}_{b_{\theta,t}}\left(1-\frac{1}{R_t^{b,\tau}}\frac{\kappa_t^{ss}}{b^y_{\theta,t}}\left(z_{\theta,t}p_{t+1|t}+\psi_{t+1}\frac{\lambda_t^L}{R_t^{\tau,b}E_t\left[\lambda^o_{\theta,t+1}\right]}\right)\right)\right.\right.$$
$$\left.\left.+\frac{d\tau_t^K}{db_{\theta,t}}\frac{q^y_{\theta,t}\beta E_t\left[\left(\gamma_{t+1}F_{K,t+1}-\delta\right)\lambda^o_{\theta,t+1}\right]-\lambda^K_{\theta,\theta',t}}{\beta R_t^{\tau,b}E_t\left[\lambda^o_{\theta,t+1}\right]}\right]\right]$$
$$=\frac{n^o_{t+1}\psi_{t+1}\Theta_{t+1}(\theta)}{n^y_{t+1}\Theta_{t+1}(\theta)}E_t\left[\widetilde{\mathcal{W}}_{\theta,t+1}\left[1-\tilde{\lambda}^y_{\theta,t+1}\left[\eta^y_{\theta,t+1}+\sigma^y_{\theta,t+1}\frac{q^y_{\theta,t+1}}{c^y_{\theta,t+1}}E_{t+1}\left[\frac{\lambda^o_{\theta,t+2}\left(\gamma_{t+2}F_{K,t+2}-\delta\right)}{r^b_{t+1}E_{t+1}\left[\lambda^o_{\theta,t+2}\right]}-1\right]\right.\right.\right.$$
$$\left.\left.\left.-\sigma^y_{\theta,t+1}\frac{\lambda^K_{t+1}}{c^y_{\theta,t+1}}\frac{1}{\beta r^b_{t+1}E_{t+1}\left[\lambda^o_{\theta,t+2}\right]}-\frac{\lambda^L_{\theta,\theta',t+1}}{R^{b,\tau}_{t+1}\beta E_{t+1}\left[\lambda^o_{\theta,t+2}\right]}\left[\frac{u^y_{lc,\theta,t+1}}{\gamma_{t+1}\theta F_{\tilde{L},t+1}u^y_{c,\theta,t+1}}-\sigma_{\theta,t+1}\frac{\left(1-\tau^L_{t+1}\right)}{c^y_{\theta,t+1}}\right]\right]\right]\right] \tag{A.10}$$

This general expression can be then reduced to the case of a separable utility function noting that in that case

$$\eta^y_{\theta,t} = 1-\sigma^y_{\theta,t}$$

□

A.5 Capital taxes and social security

A.5.1 Proof of Proposition 6

Proof. This proposition follows from A.5.4 once we set $\tau_t^{\text{ss,F}}$ and κ_t^{ss} to zero, and noting that in the case of homogeneous agents we do not have terms in $\lambda^K_{\theta,\theta',t}$. □

A.5.2 Proof of Proposition 7

Proof. This proposition follows from A.5.4 once we set $\tau_t^{\text{ss,F}}$ and κ_t^{ss} to zero. □

A.5.3 Proof of Proposition 8

Proof. The proof follows form A.5.4. In particular, we can use (A.12) and note that with homogeneous agents there are no terms in $\lambda^K_{\theta,\theta',t}$. Then we can substitute in (A.14) and realize that almost all terms cancel out and we are left with

$$\tau_t^K = \frac{E_t\left[\beta\tilde{\lambda}^o_{t+1}u^o_{c,t+1}R_{t+1}\right]-\tilde{\lambda}^y_t u^y_{c,\theta,t}}{\beta E_t\left[\tilde{\lambda}^o_{t+1}u^o_{c,t+1}\left(\gamma_{t+1}F_{K,t+1}-\delta\right)\right]}$$

□

A.5.4 Proof of Proposition 9

Proof. We prove the most general case (for two types θ and θ') in which the utility is not necessarily separable and then all the other cases will follow from this. To start, we compute the first order conditions of the Ramsey planner with respect to $q_{\theta,t}$ by taking the appropriate derivative of the Lagrangian (A.2). This first order condition reads:

$$
\begin{aligned}
0 = & -\lambda^{y}_{\theta,t} u^{y}_{c,\theta,t} + E_t \left[\left(\lambda^{o}_{\theta,t+2} - \psi_{t+2} \lambda^{y}_{\theta,t+1} E_{t+1} \left[u^{o}_{c,\theta,t+2} \right] \right) \beta \kappa^{ss}_{t+1} \theta \gamma_{t+1} F_{\tilde{L}K,t+1} n^{y}_{t} \Theta_t(\theta) \, l_{\theta,t+1} \right] \\
& + E_t \left[\beta \lambda^{o}_{\theta,t+1} \left(1 + \left(u^{y}_{c,\theta,t} - \beta \psi_{t+1} E_t \left[u^{o}_{c,\theta,t+1} \right] \right) \frac{(\gamma_{t+1} F_{K,t+1} - \delta)}{\beta \psi_{t+1} E_t \left[(\gamma_{t+1} F_{K,t+1} - \delta) \, u^{o}_{c,\theta,t+1} \right]} \right) \right] \\
& + E_t \left[\beta \lambda^{o}_{\theta,t+1} q^{y}_{\theta,t} \left(u^{y}_{c,\theta,t} - \beta \psi_{t+1} E_t \left[u^{o}_{c,\theta,t+1} \right] \right) \frac{n^{y}_{t} \Theta_t(\theta)}{\beta \psi_{t+1}} \right. \\
& \left. \left(\frac{\gamma_{t+1} F_{KK,t+1} E_t \left[(\gamma_{t+1} F_{K,t+1} - \delta) \, u^{o}_{c,\theta,t+1} \right] - (\gamma_{t+1} F_{K,t+1} - \delta) E_t \left[\gamma_{t+1} F_{KK,t+1} u^{o}_{c,\theta,t+1} \right]}{E_t \left[(\gamma_{t+1} F_{K,t+1} - \delta) \, u^{o}_{c,\theta,t+1} \right]^2} \right) \right] \\
& - \lambda^{f}_{t} n^{y}_{t} \Theta_t(\theta) + E_t \left[\lambda^{f}_{t+1} \Theta_t(\theta) \left(\gamma_{t+1} F_{K,t+1} n^{y}_{t} + (1-\delta) n^{y}_{t} \right) \right] \\
& + E_t \left[\left(\lambda^{f}_{t+2} (1-\delta) - \lambda^{f}_{t+1} \right) n^{y}_{t+1} \tau^{ss,F}_{t+1} \gamma_{t+1} F_{\tilde{L}K,t+1} n^{y}_{t} \Theta_t(\theta) \sum_{\theta} \theta l_{\theta,t+1} \Theta_{t+1}(\theta) \right] \\
& + \lambda^{K}_{\theta,\theta',t} \left[\frac{u^{y}_{c,\theta,t} - \beta \psi_{t+1} E_t \left[u^{o}_{c,\theta,t+1} \right]}{\beta \psi_{t+1} E_t \left[(\gamma_{t+1} F_{K,t+1} - \delta) \, u^{o}_{c,\theta,t+1} \right]^2} E_t \left[\gamma_{t+1} F_{KK,t+1} u^{o}_{c,\theta,t+1} \right] n^{y}_{t} \Theta_t(\theta) \right. \\
& \left. - \frac{u^{y}_{c,\theta',t} - \beta \psi_{t+1} E_t \left[u^{o}_{c,\theta',t+1} \right]}{\beta \psi_{t+1} E_t \left[(\gamma_{t+1} F_{K,t+1} - \delta) \, u^{o}_{c,\theta',t+1} \right]^2} E_t \left[\gamma_{t+1} F_{KK,t+1} u^{o}_{c,\theta',t+1} \right] n^{y}_{t} \Theta_t(\theta') \right] \\
& - E_t \left[\lambda^{L}_{\theta,\theta',t+1} \frac{\kappa^{ss}_{t+1} \gamma_{t+1} \theta F_{\tilde{L}K,t+1} E_{t+1} \left[u^{o}_{c,\theta,t+2} \right]}{\gamma_{t+1} \theta F_{\tilde{L},t+1} u^{y}_{c,\theta,t+1}} n^{y}_{t} \Theta_t(\theta) - \lambda^{L}_{\theta,\theta',t+1} \frac{\kappa^{ss}_{t+1} \gamma_{t+1} \theta' F_{\tilde{L}K,t+1} E_{t+1} \left[u^{o}_{c,\theta',t+2} \right]}{\gamma_{t+1} \theta' F_{\tilde{L},t+1} u^{y}_{c,\theta',t+1}} n^{y}_{t} \Theta_t(\theta) \right. \\
& - \lambda^{L}_{\theta,\theta',t+1} \frac{u^{y}_{l,\theta,t} + \beta \psi_{t+2} \kappa^{ss}_{t+1} \gamma_{t+1} \theta F_{\tilde{L},t+1} E_{t+1} \left[u_{c,\theta,t+2} \right]}{\gamma_{t+1} \theta u^{y}_{c,\theta,t+1} F^{2}_{\tilde{L},t+1}} F_{\tilde{L}K,t+1} n^{y}_{t} \Theta_t(\theta) \\
& \left. + \lambda^{L}_{\theta,\theta',t+1} \frac{u^{y}_{l,\theta',t} + \beta \psi_{t+2} \kappa^{ss}_{t+1} \gamma_{t+1} \theta' F_{\tilde{L},t+1} E_{t+1} \left[u_{c,\theta',t+2} \right]}{\gamma_{t+1} \theta' u^{y}_{c,\theta',t+1} F^{2}_{\tilde{L},t+1}} F_{\tilde{L}K,t+1} n^{y}_{t} \Theta_t(\theta) \right]
\end{aligned}
$$

Using (4.4), (4.5), and (A.7) we get:

$$0 = -\lambda^y_{\theta,t} u^y_{c,\theta,t} + E_t \left[\left(\lambda^o_{\theta,t+2} - \psi_{t+2} \lambda^y_{\theta,t+1} E_{t+1} \left[u^o_{c,\theta,t+2} \right] \right) \beta \kappa^{ss}_{t+1} \theta \gamma_{t+1} F_{\tilde{L}K,t+1} n^y_t \Theta_t(\theta) l_{\theta,t+1} \right] + E_t \left[\beta \lambda^o_{\theta,t+1} R^{K,\tau}_{t+1} \right]$$
$$+ E_t \left[\beta \lambda^o_{\theta,t+1} q^y_{\theta,t} \left(1 - \tau^K_t\right) \frac{n^y_t \Theta_t(\theta)}{\beta \psi_{t+1}} \right.$$
$$\left. \left(\frac{\gamma_{t+1} F_{KK,t+1} E_t \left[\left(\gamma_{t+1} F_{K,t+1} - \delta \right) u^o_{c,\theta,t+1} \right] - \left(\gamma_{t+1} F_{K,t+1} - \delta \right) E_t \left[\gamma_{t+1} F_{KK,t+1} u^o_{c,\theta,t+1} \right]}{E_t \left[\left(\gamma_{t+1} F_{K,t+1} - \delta \right) u^o_{c,\theta,t+1} \right]} \right) \right]$$
$$- \lambda^f_t n^y_t \Theta_t(\theta) + E_t \left[\lambda^f_{t+1} \Theta_t(\theta) \left(\gamma_{t+1} F_{K,t+1} n^y_t + (1-\delta) n^y_t \right) \right]$$
$$+ E_t \left[\left(\lambda^f_{t+2}(1-\delta) - \lambda^f_{t+1} \right) n^y_{t+1} \tau^{ss,F}_{t+1} \gamma_{t+1} F_{\tilde{L}K,t+1} n^y_t \Theta_t(\theta) \sum_\theta \theta l_{\theta,t+1} \Theta_{t+1}(\theta) \right]$$
$$+ \lambda^K_{\theta,\theta',t} \left(1 - \tau^K_t\right) n^y_t \left(\frac{E_t \left[\gamma_{t+1} F_{KK,t+1} u^o_{c,\theta,t+1} \right]}{E_t \left[\left(\gamma_{t+1} F_{K,t+1} - \delta \right) u^o_{c,\theta,t+1} \right]} \Theta_t(\theta) - \frac{E_t \left[\gamma_{t+1} F_{KK,t+1} u^o_{c,\theta',t+1} \right]}{E_t \left[\left(\gamma_{t+1} F_{K,t+1} - \delta \right) u^o_{c,\theta',t+1} \right]} \Theta_t(\theta) \right)$$
$$- E_t \left[\lambda^L_{\theta,\theta',t+1} n^y_t \kappa^{ss}_{t+1} \frac{F_{\tilde{L}K,t+1}}{F_{\tilde{L},t+1}} \left(\frac{E_{t+1} \left[u^o_{c,\theta,t+2} \right]}{u^y_{c,\theta,t+1}} \Theta_t(\theta) - \frac{E_{t+1} \left[u^o_{c,\theta',t+2} \right]}{u^y_{c,\theta',t+1}} \Theta_t(\theta) \right) \right.$$
$$\left. - \lambda^L_{\theta,\theta',t+1} \left(1 - \tau^L_t\right) \frac{F_{\tilde{L}K,t+1}}{F_{\tilde{L},t+1}} n^y_t \left(\Theta_t(\theta) - \Theta_t(\theta) \right) \right] \tag{A.11}$$

Rearranging some terms and using (A.1b) we get:

$$\tau^K_t = \frac{E_t \left[\left(\lambda^f_{t+1} R_{t+1} - \lambda^f_t \right) \right] n^y_t \Theta_t(\theta)}{\mathcal{W}_{\theta,t} \beta E_t \left[\tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) \right]} + \frac{E_t \left[\beta \tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} R_{t+1} \right] - \tilde{\lambda}^y_{\theta,t} u^y_{c,\theta,t}}{\beta E_t \left[\tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) \right]}$$
$$+ \frac{\beta n^y_t \Theta_t(\theta) E_t \left[- \left(1 - \tau^K_t\right) q^y_{\theta,t} \gamma_{t+1} F_{KK,t+1} u^o_{c,\theta,t+1} \right]}{\beta E_t \left[\tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) \right]}$$
$$\left(\frac{Cov_t \left[\tilde{\lambda}^o_{\theta,t+1}, q^y_{\theta,t} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) u^o_{c,\theta,t+1} \right]}{E_t \left[q^y_{\theta,t} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) u^o_{c,\theta,t+1} \right]} - \frac{Cov_t \left[\tilde{\lambda}^o_{\theta,t+1}, q^y_{\theta,t} \gamma_{t+1} F_{KK,t+1} u^o_{c,\theta,t+1} \right]}{E_t \left[q^y_{\theta,t} \gamma_{t+1} F_{KK,t+1} u^o_{c,\theta,t+1} \right]} \right)$$
$$+ \frac{E_t \left[\left(\lambda^f_{t+2}(1-\delta) - \lambda^f_{t+1} \right) n^y_{t+1} \tau^{ss,F}_{t+1} \gamma_{t+1} F_{\tilde{L}K,t+1} n^y_t \Theta_t(\theta) \left(\sum_\theta \theta l_{\theta,t+1} \Theta_{t+1}(\theta) \right) \right]}{\mathcal{W}_{\theta,t} \beta E_t \left[\tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) \right]} \tag{A.12}$$
$$+ \frac{E_t \left[\mathcal{W}_{\theta,t+1} \left(\tilde{\lambda}^o_{\theta,t+2} u^o_{c,\theta,t+2} - E_{t+1} \left[\tilde{\lambda}^o_{\theta,t+2} u^o_{c,\theta,t+2} \right] \right) \beta \kappa^{ss}_{t+1} \theta \gamma_{t+1} F_{\tilde{L}K,t+1} n^y_t \Theta_t(\theta) l_{\theta,t+1} \right]}{\mathcal{W}_{\theta,t} \beta E_t \left[\tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) \right]}$$
$$+ \frac{\lambda^K_{\theta,\theta',t} n^y_t \Theta_t(\theta)}{\mathcal{W}_{\theta,t} \beta E_t \left[\tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) \right]} \frac{\left(1 - \tau^K_t\right)}{r^b_t}$$
$$\left(\frac{Cov_t \left[\gamma_{t+1} F_{KK,t+1}, u^o_{c,\theta,t+1} \right]}{E_t \left[u^o_{c,\theta,t+1} \right]} - \frac{Cov_t \left[\gamma_{t+1} F_{KK,t+1}, u^o_{c,\theta',t+1} \right]}{E_t \left[u^o_{c,\theta',t+1} \right]} \right)$$

where $\tilde{\lambda}^y_{\theta,t} = \frac{\lambda^y_{\theta,t}}{\mathcal{W}_{\theta,t}}$, $\tilde{\lambda}^o_{\theta,t+1} = \frac{\lambda^o_{\theta,t+1}}{\mathcal{W}_{\theta,t} u^o_{c,\theta,t+1}}$ and $\mathcal{W}_{\theta,t} = \frac{\tilde{W}_{\theta,t}}{u^y_{c,\theta,t}}$. Now we can simplify this expression by using the first order condition of the Ramsey planner with respect to $\tau^{\text{ss,F}}_t$:

$$
\begin{aligned}
0 = &-E_t\left[\sum_\theta\left[\lambda^y_{\theta,t+1}\beta\psi_{t+2}\kappa^{ss}_{t+1}E_{t+1}\left[u^o_{c,\theta,t+2}\right]\theta\gamma_{t+1}F_{K\tilde{L},t+1}l_{\theta,t+1}n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\right]\right]\\
&+E_t\left[\beta\sum_\theta\left[\lambda^o_{\theta,t+1}\left(\frac{u^y_{c,\theta,t}-\beta\psi_{t+1}E_t\left[u^o_{c,\theta,t+1}\right]}{\beta\psi_{t+1}E_t\left[\left(\gamma_{t+1}F_{K,t+1}-\delta\right)u^o_{c,\theta,t+1}\right]}\right)\gamma_{t+1}F_{KK,t+1}q^y_{\theta,t}n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\right]\right]\\
&-E_t\left[\beta\sum_\theta\left[\lambda^o_{\theta,t+1}\left(\frac{u^y_{c,\theta,t}-\beta\psi_{t+1}E_t\left[u^o_{c,\theta,t+1}\right]}{\beta\psi_{t+1}E_t\left[\left(\gamma_{t+1}F_{K,t+1}-\delta\right)u^o_{c,\theta,t+1}\right]^2}\right)\right.\right.\\
&\left.\left.E_t\left[\gamma_{t+1}F_{KK,t+1}u^o_{c,\theta,t+1}\right]\left(\gamma_{t+1}F_{K,t+1}-\delta\right)q^y_{\theta,t}n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\right]\right]\\
&+E_t\left[\beta\sum_\theta\left[\lambda^o_{\theta,t+2}\kappa^{ss}_{t+1}\theta\gamma_{t+1}F_{\tilde{L}K,t+1}l_{\theta,t+1}n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\right]\right]\\
&-\lambda^f_t n^y_t\gamma\left(s^t\right)F_{\tilde{L},t}\left(\sum_\theta\theta l_{\theta,t}\Theta_t(\theta)\right)-n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)E_t\left[\lambda^f_{t+1}n^y_{t+1}\tau^{ss,F}_{t+1}\gamma_{t+1}F_{\tilde{L}K,t+1}\left(\sum_\theta\theta l_{\theta,t+1}\Theta_{t+1}(\theta)\right)\right]\\
&+n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)E_t\left[\lambda^f_{t+1}\gamma_{t+1}F_{K,t+1}\right]+(1-\delta)n^y_t\gamma_tF_{\tilde{L},t}\left(\sum_\theta\theta l_{\theta,t}\Theta_t(\theta)\right)E_t\left[\lambda^f_{t+1}\right]\\
&+(1-\delta)n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)E_t\left[\lambda^f_{t+2}n^y_{t+1}\tau^{ss,F}_{t+1}\gamma_{t+1}F_{\tilde{L}K,t+1}\left(\sum_\theta\theta l_{\theta,t+1}\Theta_{t+1}(\theta)\right)\right]\\
&+\lambda^K_{\theta,\theta',t}\left[\frac{u^y_{c,\theta,t}-\beta\psi_{t+1}E_t\left[u^o_{c,\theta,t+1}\right]}{\beta\psi_{t+1}E_t\left[\left(\gamma_{t+1}F_{K,t+1}-\delta\right)u^o_{c,\theta,t+1}\right]^2}E_t\left[\gamma_{t+1}F_{KK,t+1}u^o_{c,\theta,t+1}\right]n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\right.\\
&\left.-\frac{u^y_{c,\theta',t}-\beta\psi_{t+1}E_t\left[u^o_{c,\theta',t+1}\right]}{\beta\psi_{t+1}E_t\left[\left(\gamma_{t+1}F_{K,t+1}-\delta\right)u^o_{c,\theta',t+1}\right]^2}E_t\left[\gamma_{t+1}F_{KK,t+1}u^o_{c,\theta',t+1}\right]n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\right]\\
&-E_t\left[\lambda^L_{\theta,\theta',t+1}\frac{\kappa^{ss}_{t+1}\gamma_{t+1}\theta F_{\tilde{L}K,t+1}E_{t+1}\left[u^o_{c,\theta,t+2}\right]}{\gamma_{t+1}\theta F_{\tilde{L},t+1}u^y_{c,\theta,t+1}}n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\right.\\
&-\lambda^L_{\theta,\theta',t+1}\frac{u^y_{l,\theta,t}+\beta\psi_{t+2}\kappa^{ss}_{t+1}\gamma_{t+1}\theta F_{\tilde{L},t+1}E_{t+1}\left[u_{c,\theta,t+2}\right]}{\gamma_{t+1}\theta u^y_{c,\theta,t+1}F^2_{\tilde{L},t+1}}F_{\tilde{L}K,t+1}n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\\
&-\lambda^L_{\theta,\theta',t+1}\frac{\kappa^{ss}_{t+1}\gamma_{t+1}\theta' F_{\tilde{L}K,t+1}E_{t+1}\left[u^o_{c,\theta',t+2}\right]}{\gamma_{t+1}\theta' F_{\tilde{L},t+1}u^y_{c,\theta',t+1}}n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\\
&\left.+\lambda^L_{\theta,\theta',t+1}\frac{u^y_{l,\theta',t}+\beta\psi_{t+2}\kappa^{ss}_{t+1}\gamma_{t+1}\theta' F_{\tilde{L},t+1}E_{t+1}\left[u_{c,\theta',t+2}\right]}{\gamma_{t+1}\theta' u^y_{c,\theta',t+1}F^2_{\tilde{L},t+1}}F_{\tilde{L}K,t+1}n^y_t\left(\sum_\theta w_{\theta,t}l_{\theta,t}\Theta_t(\theta)\right)\right]
\end{aligned}
\tag{A.13}
$$

Using similar steps to the ones that let us obtain (A.12), we get:

$$\begin{aligned}
&\beta E_t \left[\sum_\theta \left(\lambda^o_{\theta,t+2} - E_{t+1}\left[\lambda^o_{\theta,t+2} \right] \right) \kappa^{ss}_{t+1} \gamma_{t+1} F_{\tilde{L}K,t+1} \theta l_{\theta,t+1} \right] \\
&+ \left[\beta \left(1-\tau^K_t\right) \sum_\theta q^y_{\theta,t} \mathcal{W}_{\theta,t} E_t \left[\gamma_{t+1} F_{KK,t+1} u^o_{c,\theta,t+1} \right] \right. \\
&\left. \left(\frac{Cov\left(\tilde{\lambda}^o_{\theta,t+1}, u^o_{c,\theta,t+1} \gamma_{t+1} F_{KK,t+1} \right)}{E_t\left[\gamma_{t+1} F_{KK,t+1} u^o_{c,\theta,t+1} \right]} - \frac{Cov\left(\tilde{\lambda}^o_{\theta,t+1}, u^o_{c,\theta,t+1} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) \right)}{E_t \left[\left(\gamma_{t+1} F_{K,t+1} - \delta \right) u^o_{c,\theta,t+1} \right]} \right) \right] \\
&+ E_t \left[\left(\lambda^f_{t+2} (1-\delta) - \lambda^f_{t+1} \right) \tau^{ss,F}_{t+1} z_{t+1} \frac{F_{\tilde{L}K,t+1}}{F_{\tilde{L},t+1}} \right] + E_t \left[\lambda^f_{t+1} \left(\gamma_{t+1} F_{K,t+1} + (1-\delta) \right) - \lambda^f_t \right] \\
&+ \lambda^K_{\theta,\theta',t} \left[\frac{\left(1-\tau^K_t\right)}{r^b_t} \left(\frac{Cov_t\left[\gamma_{t+1} F_{KK,t+1}, u^o_{c,\theta,t+1} \right]}{E_t\left[u^o_{c,\theta,t+1} \right]} - \frac{Cov_t\left[\gamma_{t+1} F_{KK,t+1}, u^o_{c,\theta',t+1} \right]}{E_t\left[u^o_{c,\theta',t+1} \right]} \right) \right] = 0
\end{aligned} \tag{A.14}$$

Now, we can sum (A.12) across all θ and substituting in (A.14) we arrive to:

$$\begin{aligned}
&\tau^K_t \sum_\theta \frac{\mathcal{W}_{\theta,t} \beta E_t \left[\tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} \left(\gamma_{t+1} F_{K,t+1} - \delta \right) \right]}{n^y_t \Theta_t(\theta)} = \left(N^\Theta - 1 \right) E_t \left[\left(\lambda^f_{t+1} R_{t+1} - \lambda^f_t \right) \right] \\
&\quad + \sum_\theta \frac{\mathcal{W}_{\theta,t} \left[E_t \left[\beta \tilde{\lambda}^o_{\theta,t+1} u^o_{c,\theta,t+1} R_{t+1} \right] - \tilde{\lambda}^y_{\theta,t} u^y_{c,\theta,t} \right]}{n^y_t \Theta_t(\theta)} \\
&\quad + \left(N^\Theta - 1 \right) E_t \left[\left(\lambda^f_{t+2} (1-\delta) - \lambda^f_{t+1} \right) n^y_{t+1} \tau^{ss,F}_{t+1} \gamma_{t+1} F_{\tilde{L}K,t+1} \left(\sum_\theta \theta l_{\theta,t+1} \Theta_{t+1}(\theta) \right) \right] \\
&\quad + \lambda^K_{\theta,\theta',t} \left(N^\Theta - 1 \right) \frac{\left(1-\tau^K_t\right)}{r^b_t} \left(\frac{Cov_t\left[\gamma_{t+1} F_{KK,t+1}, u^o_{c,\theta,t+1} \right]}{E_t\left[u^o_{c,\theta,t+1} \right]} - \frac{Cov_t\left[\gamma_{t+1} F_{KK,t+1}, u^o_{c,\theta',t+1} \right]}{E_t\left[u^o_{c,\theta',t+1} \right]} \right)
\end{aligned} \tag{A.15}$$

□