EINAUDI INSTITUTE FOR ECONOMICS AND FINANCE
LUISS UNIVERSITY, DEPARTMENT OF ECONOMICS AND FINANCE
ROME MASTERS IN ECONOMICS

OPTIMAL TAXATION IN AN OVERLAPPING GENERATIONS ECONOMY

CANDIDATE
ROBERTO SAITTO

ADVISOR
PROF. FACUNDO PIGUILLEM

CO-ADVISOR
PROF. PIETRO REICHLIN

ACADEMIC YEAR 2019-2020
ACKNOWLEDGEMENTS

I am deeply grateful to my advisor professor Facundo Piguillem for his extraordinary support. Every part of the present work benefited from his insights, which deeply contributed to my intellectual growth. I also owe special thanks to my co-advisor professor Pietro Reichlin for his helpful comments and ideas. I thank the participants to the EIEF presentations for valuable contributions. Finally, I wish to express my gratitude to my classmates and friends Felipe Berrutti and Marco Castelluccio, for careful reading and fruitful discussions.

I dedicate this work to my family.
Abstract. I study optimal taxation within the framework of a two-period overlapping generations economy. I consider an economic environment which preserves a standard Ramsey problem approach while allowing the government to impose non-linear taxes on labour income. I prove that this modification enables the government to achieve the first best when agents are homogeneous, provided that limits on public debt and on labour tax progressivity are sufficiently loose. Then, I argue that debt limits are extremely powerful rationales for optimal capital taxation to be significantly positive in the long run. Taking Italy as the benchmark economy, I run numerical simulations to assess optimal taxation both from a steady state and a dynamic perspective. My quantitative findings are similar to those in Conesa et al. (2009). Implied optimal capital taxation is positive for all the non-negative levels of government debt and increasing in the debt-to-GDP ratio. The optimal labour tax is almost linear in all scenarios.

1. Introduction

Optimal taxation has been hotly debated by economists. A large number of relevant contributions have been written to answer the fundamental question of how fiscal policy should be set over the business cycle and the long run. A classical framework for the problem may be found in the work of Ramsey (1927), where the government has to raise an exogenously given revenue by the means of distortionary linear taxes, with the aim of minimizing the utility loss consequently experienced by the economic agents. The Ramsey problem has then been extended to the dynamic maximization of a given welfare function using distortionary taxes to induce a particular competitive equilibrium. That is, the economic literature has tried to identify the fiscal policy characterizing the second best achievable by a government when non distortionary lump transfers are not feasible. In particular, should capital be taxed in the long run? Which is the optimal progressivity level for labour taxation?

The present work contributes to the literature debate on optimal taxation. As noticed by Peterman (2013), optimal capital taxation models are fairly complicated and it is important to disentangle the implications of each of the assumptions in determining the results. Following Piguillem and Shakhnov (2020), I show that government debt plays a fundamental role in determining capital optimal taxation. In particular, I highlight that government debt limits are among the main rationales for positive capital taxation to be optimal in the long run for an overlapping generations economy. Furthermore, I characterize the full optimal dynamics of the economy for some standard preference specifications, a result which is not common in the literature. To this end, I study a simple two-period overlapping generation environment, where agents may be heterogeneous in their labour productivity and are
subject to a mortality risk, with no annuity market. In this environment, the government is allowed to impose non linear taxes on labour income, according to the functional form proposed by Feldstein (1969) and Heathcote et al. (2017). Interestingly, Heathcote and Tsujiyama (2019) argue that this specification is able to well approximate the optimal solution for the case in which no functional restriction is imposed on the tax schedule. Following Conesa et al. (2009), capital taxation is linear.

First, I show that even if the assumption of a linear labour tax is relaxed, it is still possible to frame the Ramsey problem preserving the so-called primal approach.\footnote{See Chari and Kehoe (1999).} This consists in characterizing the allocations implementable as competitive equilibria by some government fiscal policy only using agents’ budget constraints and first order optimality conditions. This approach has the advantage to handle directly with allocations rather than with the set of taxes which induce them, that means making the problem more intuitive.

Second, I prove that in an homogenous agent economy, if the government is free to choose the labour tax progressivity and does not face binding debt limits, then it is able to achieve the first best in steady state.\footnote{The first best is defined as the allocation which would be implemented by a utilitarian social planner, who does not necessarily share the same time discount factor of the agents.} Furthermore, under this assumptions the capital tax is in general negative in steady state, perfectly compensating the agents for the mortality risk they are subject to.

Third, I provide numerical exercises, taking Italy as benchmark economy, to argue that constraints on the government debt accumulation are fundamental rationales for largely positive capital taxation in the long run. I show that the simple economic environment I adopt is able to closely replicate the results of Conesa et al. (2009).\footnote{Assuming the government has to keep debt equal to zero, as they did for their main model specification, I get an optimal steady state capital tax of 0.37, where they get 0.36.} In particular, for debt-to-GDP ratios between 0 and 0.6, the optimal tax rate on capital ranges from 0.37 to 0.56.

Fourth, building on Piguillem and Shakhnov (2020), I develop a recursive formulation of the problem which allows for aggregation to numerically study the dynamics of the economy. I show that the economy always converges to a unique steady state. The implied optimal labour tax is almost linear.

**Literature.** Regarding capital taxation, Chamley (1986) has proved that, for a wide class of infinitely lived agent models, it should be zero in steady state, a position supported also by Judd (1985) and Lucas (1990). Furthermore, Jones et al. (1997) and Chari and Kehoe (1999) proved the result of Chamley (1986) is robust to the relaxation of a set of assumptions. However, Straub and Werning (2020) argue that convergence to such a steady state may be suboptimal for some preference specifications. Furthermore, the literature has identified at least two rationales for capital taxation to be positive in the long run. Aiyagari (1995) has shown that for the Bewley (1986) class of incomplete market models capital taxation is always positive; interestingly, the result holds not only in steady state, but also for the whole business cycle. Second, Erosa and Gervais (2000) and Garriga (2017) argue that in life cycle models where taxes cannot depend of agents’ age, capital tax is in general different from zero.
Conesa et al. (2009) quantitatively address the issue, finding that representing the U.S. economy in a multi-period overlapping generations model with stochastic life cycle, mortality risk and incomplete markets, optimal capital taxation is significantly positive in steady state.\footnote{They get an optimal capital tax of 0.36 in their main model specification.}

Conesa et al. (2009) also find that the optimal tax on labour is essentially flat. In fact, the classical framework inspired by Ramsey (1927) in general restricts taxes to be linear, but progressivity in labour income taxation has emerged as a central topic in the literature. Several authors, as Benabou (2002) and Heathcote et al. (2017) have claimed for the optimality of a progressive labour tax. Other contributions, like Altig et al. (2001), support flat labour income taxation.

Section 2 presents the economic environment in which the analysis is conducted. Section 3 defines the competitive equilibrium. Section 4 introduces the social planner problem. Sections 5 and 6 contain the main theoretical contributions and discuss the second best achievable by the government. Section 7 describes the specification used for the numerical analysis of the steady state. Sections 8 and 9 introduce the recursive formulation of the problem and the aggregation model used for the numerical simulation of the economics dynamics. Section 10 comments the numerical results. Section 11 concludes.

\section*{2. Economic Environment}

I consider an infinite horizon two-period overlapping generations economy. Time is discrete. At all periods $t \geq 0$, a generation of agents is born. For any given period $t$, I will refer to the newborns as the young and to the members of the previous generation as the old. Young agents survive to the next period with a probability $\psi \in (0, 1]$. They are heterogenous in their production efficiency. Labour supply is endogenous and it is assumed that only the young can work, while the old are retired. Moreover, for simplicity all the generations are assumed to be identical to each other\footnote{For example, I abstract from population growth.}.

In order to formally define the economy, let me introduce some notation. Let $\mathcal{I}$ and $\mathcal{B}$ denote a non-empty separable metric space and the Borel $\sigma$–algebra defined over it, respectively. Further, let $i : \mathcal{B} \to \mathbb{R}$ be a finite measure such that $i (\mathcal{I}) = 1$. Finally, let $L^2$ denote the Hilbert space of real-valued functions defined on the measure space $\mathcal{G} \equiv (\mathcal{I}, \mathcal{B}, i)$,\footnote{To be precise, consider the set of all the $\mathcal{B}$–measurable real valued functions $f : \mathcal{I} \to \mathbb{R}$ such that $\int_\mathcal{I} f^2 \, di < \infty$ and consider the quotient space generated by the equivalence relation: $f \sim g$ if and only if $f = g$ $i$–almost everywhere. Then $L^2$ is the Hilbert space defined by the quotient space and the metric induced by the standard norm $\|f\|_2 = \left(\int_\mathcal{I} f^2 \, di\right)^{\frac{1}{2}}$.}

\textbf{Generations and Idiosyncratic Efficiency.} A generation of agents has is defined as the measure space $\mathcal{G} \equiv (\mathcal{I}, \mathcal{B}, i)$ and thus has measure one. For all $t \geq 0$, let the real functions $c^t_1, l_1, c^{t+1}_1 \in L^2$ denote young consumption, labour supply and old consumption of the generation of agents born in period $t$, respectively. There is intergenerational heterogeneity in agents’ idiosyncratic efficiency. The
productivity of each agent is determined by a function \( e \in \mathcal{L}^2 \) such that \( \int_x e \, dx = 1 \) and \( e \geq e \) almost everywhere for some \( e \in \mathbb{R}^{++} \).

**Consumption Set and Preferences.** All agents of all generations share the same consumption set and preferences. Let \( X \) be the Cartesian product of the two intervals \((a, +\infty)\) and \([0, l] \) with \( a \geq 0, l > 0 \). Then, the consumption set \( X \) is defined as the Cartesian product of the set \( X \) and the interval \((a, +\infty)\). Let \( u : X \rightarrow \mathbb{R} \) be a concave function, twice continuously differentiable on its domain. I assume that consumers enjoy consumption and dislike working, i.e. \( u \) is increasing in its first argument and decreasing in its second one. Furthermore, let \( u \) satisfy the Inada conditions and denote as \( u_e \) and \( u_l \) its first order partial derivatives with respect to its first and second argument\(^7\). The future discount factor is \( \beta > 0 \) such that \( \beta \psi < 1 \). Then, for each generation \( t \), agents’ preferences are represented by the time-separable utility function \( U : \mathbb{R} \rightarrow \mathbb{R} \):

\[
U(c^t_e, l_t, c^t_{t+1}) = u(c^t_e, l_t) + \beta \psi u(c^t_{t+1})
\]

where with a slight abuse of notation I wrote \( u(c^t_{t+1}) \) in place of the more precise \( u(c^t_{t+1}, 0) \). The preferences of the old generation in \( t = 0 \) are represented by the utility function \( \beta \psi u(\cdot) \).

**Technology and Markets.** Markets are competitive. Assume there is a representative firm equipped with a concave, homogeneous of degree one and twice continuously differentiable production function \( F : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+ \) increasing in both its arguments. Further, let the cross derivative be positive and \( F(K, 0) = F(0, L) = 0 \) for all \( K, L \in \mathbb{R}_+ \). Capital is subject to an exogenous depreciation rate \( \delta \in [0, 1] \). Let \( r_t, w_t \in \mathbb{R} \) denote the return on capital and the labour wage at \( t \). At each period, the firm chooses capital \( K_t \) and labour \( L_t \) to maximize its profit \( \pi_t \equiv F(K_t, L_t) - r_t K_t - w_t L_t \).

Markets are sequential. At each period \( t \), the members of the newborn generation can buy a safe asset which pays in the next period. Let the function \( a_{t+1} \in \mathcal{L}^2 \) denote the savings of the generation of agents born in \( t \). Then, at each period \( t \) the pre-tax capital income of the old generation is defined by the function \( R_t a_t \) where \( R_t \equiv 1 - \delta + r_t \) is the pre-tax return on capital assets. Old agents receive the payment just before they have the chance to die. There is no annuity market and at all \( t \) the unintended bequests of the old generation are given to the young through a transfer \( T^p_t \in \mathbb{R} \). The pre-tax labour income of the young generation agents is given by the product of the real number \( w_t \) with the functions \( e \) and \( l_t \), for all \( t \). Assume \( e l_t \in \mathcal{L}^2 \).

**Government.** The government is benevolent, discounts future generations at a rate \( \phi \in (0, 1) \) and faces an exogenously given sequences of public expenditures \( g \equiv \{g_t\}_{t=0}^{+\infty} \) with \( g_t \in \mathbb{R} \) for all \( t \). At every period, it can tax gross capital and labour income and raise public debt. Capital taxation, denoted by \( \tau_c \), is restricted to be linear. Initial taxation on capital \( \tau_0 \) is exogenously given. Further, as in Conesa et al. (2009), the government taxes net pre-tax capital income \((r_t - \delta) a_t \). On the other hand,

\(^7\)Similarly, let \( u_{ee} \) and \( u_{ll} \) denote the second order partial derivatives and \( u_{el} \) or \( u_{le} \) the cross derivative of \( u \).
following Heathcote et al. (2017), labour income taxation is not necessarily linear. For each \( t \geq 0 \) is defined by the function \( T_t : \mathbb{R}_+ \to \mathbb{R}_+ \):

\[
T_t(x) = x - \lambda_t x^{1-\rho_t}
\]

The government can choose \( \lambda_t \in \mathbb{R}_+ \) and \( \rho_t \in (-\infty, 1) \) at every period. Notice that this is a generalization for the classical linear labour income taxation, which is obtained as a special case for \( \rho_t = 0 \). Labour income taxation is progressive for \( \rho_t > 0 \) and regressive for \( \rho_t < 0 \). Let me remark that in a period \( t \geq 0 \), given a pre-tax labour income \( x \geq 0 \), the disposable income is \( \lambda_t x^{1-\rho_t} \).

Furthermore, the government is committed to provide old generations with a transfer equal to the aggregate value of the capital income not enjoyed by old agents due to the mortality risk

\[
T_o^\omega = \frac{1-\psi}{\psi}R^\tau \int_{I} a_i di
\]

where \( R^\tau_t \equiv 1 + (1 - \tau_t)(r_t - \delta) \) is the after-tax return on capital assets. The transfer is given to the agents just before they have the chance to die. In each period \( t \), the government can choose the level of government debt \( b_{t+1} \) subject to the intertemporal budget constraint:

\[
b_{t+1} = R^\tau_t b_t + g_t + T_o^\omega - \int_{I} T_t(w_t e_t) di - \tau_t r_t K_t
\]

To ensure the constraint is well defined, assume the function \((e_t)^{1-\rho_t}\) is in \( L^2 \). Observe that the transfer to the old is financed by the government, and thus it enters the intertemporal budget constraints. Conversely, the transfer to the young does not appear in (2.4) since it is not an expense for the government, being only composed of the unintended bequests of the previous generation. I further assume that at all \( t \geq 0 \) the government is subject to the following two policy constraints:

\[
bF(K_t, L_t) \leq b_t \leq \bar{b}F(K_t, L_t)
\]

\[
\underline{\rho} \leq \rho_t \leq \bar{\rho}
\]

where \( \{\rho, \bar{\rho}, \underline{b}, \bar{b}\} \equiv \Theta \subset \mathbb{R} \cup \{-\infty, +\infty\} \) are such that \( \underline{\rho} \leq \bar{\rho} \) and \( \underline{b} \leq \bar{b} \). The constraints (2.5) on public debt can be interpreted as binding fiscal rules. The progressivity constraints (2.6) may be constitutional limits imposed to the policy makers. A government policy is a collection \( \{\lambda_t, \rho_t, \tau_t\}_{t \geq 0} \) such that \( \underline{\rho} \leq \rho_t \leq \bar{\rho} \) for all \( t \geq 0 \) and where \( \tau_0 \) is the one exogenously specified.

**Budget Sets and the Consumer Problem.** Agents take the tax system, transfers, return on capital and prices as given. Since there is inter-generational heterogeneity in labour productivity, the budget

---

\[^8\]That is, given the transfer, the survived fraction \( \psi \) of the old generation enjoys (as an aggregate) the whole return \( R^\tau_t \int_{I} a_i di \), independently of \( \psi \).

\[^9\]The Italian constitution provides a good example of a lower bound on progressivity. Indeed, Article 53 decrees:

*Every person shall contribute to public expenditure in accordance with their capabilities. The tax system shall be progressive.*
constraint may vary across agents. With a slight abuse of notation\(^\text{10}\), the sequential budget constraints which an agent born in \(t\) is subject to may be written down as

\[
\begin{align*}
c_t^y + a_{t+1} &\leq \lambda_t (w_t e_t l_t)^{1-\rho_t} + T_t^y \\
c_{t+1} &\leq R_{t+1}^o a_{t+1} + T_{t+1}^o
\end{align*}
\]

Notice that \(\lambda_t (w_t e_t l_t)^{1-\rho_t}\) is the after-tax labour income of the agents of a generation born in \(t\). Then, the present value budget constraint is

\[
(c_t^y - T_t^y) + \frac{1}{R_{t+1}^o} (c_{t+1} - T_{t+1}^o) \leq \lambda_t (w_t e_t l_t)^{1-\rho_t}
\]

In analogy with the problem of an agent who maximizes her utility function subject to the above constraints, I give the following definition.

**Definition.** For all \(t \geq 0\), given taxes \(\{\lambda_t, \rho_t, \tau_t\}\), transfers \(\{T_t^y, T_{t+1}^o\}\) and prices \(\{w_t, r_t\}\), a collection of functions \(\{c_t^y, l_t, c_{t+1}^o\}\) in \(\mathcal{L}^2\) solves the consumer problem for the generation born in \(t\) if:

1. The collection \(\{c_t^y, l_t, c_{t+1}^o\}\) lies in the consumption set \(\mathcal{X}\), i.e.
   \[
   c_t^y > c_t, \quad 0 \leq l_t < 1, \quad c_{t+1}^o > c_t \quad i.e.
   \]
2. The Euler equation and the intratemporal labour condition hold, i.e.
   \[
   \beta \psi \frac{u_c}{u} \frac{c_{t+1}^o}{c_t^y, l_t} = \frac{1}{R_{t+1}^o} \quad i.e.
   \]
3. The present value budget constraint is satisfied with equality, i.e.
   \[
   (c_t^y - T_t^y) + \frac{1}{R_{t+1}^o} (c_{t+1} - T_{t+1}^o) = \lambda_t (w_t e_t l_t)^{1-\rho_t} \quad i.e.
   \]

Given the solution to the consumer problem, a function \(a_{t+1} \in \mathcal{L}^2\) represents the optimal savings for the generation born in \(t\) if

\[
a_{t+1} = \frac{1}{R_{t+1}^o} (c_{t+1} - T_{t+1}^o) \quad i.e.
\]

Notice that the definition of optimal savings is derived from the second sequential budget constraint of the consumer problem I discussed above. Then, I can define the economy for given initial conditions for public debt \(b_0 \in \mathcal{R}\), asset distribution \(a_0 \in \mathcal{L}^2\) and period zero capital taxation \(\tau_0 \in \mathcal{R}\). Initial capital stock is defined as \(K_0 \equiv \int_a^Z a_0 d\nu - b_0\). Initial conditions must be such that \(K_0 > 0\).

**Definition.** An economy is a collection \(\mathcal{E} \equiv \{\mathcal{G}, e, \psi\}, \{\mathcal{X}, U\}, \{F, \delta\}, \{\Theta, g, \phi\}, \{b_0, a_0, \tau_0\}\), where all the symbols have the meaning specified above in this section.

\(^{10}\)In fact, \(c_t^y, l_t, c_{t+1}^o\) are functions, but here I am treating them as if they were real numbers. However, proceeding this way should give a clear intuition for the next definition.
Given an economy $E$, let $E(E)$ be the family of economies which shares all the elements of $E$ but (possibly) the initial conditions and the sequence of public expenditures $g'$, which is nonetheless restricted to be defined as $g' = g_{t \geq t_0}$ for some nonnegative integer $t_0$. Finally, I characterize an allocation for a given economy $E$.

**Definition.** Given an economy $E$, an allocation is a collection \( \{c_y^t, l_t, c_0^t, K_t\}_{t \geq 0} \) of positive reals $K_t$ and functions $c_y^t, l_t, c_0^t$ in $L^2$ such that for all $t \geq 0$ condition (2.7) is satisfied and $el_t \in L^2$.

Let $A(E)$ be the set of all the possible allocations for the economy $E$. I say that a collection, family or set of time-indexed objects is constant (over time) to denote the fact that each of its elements is the same for all $t$. For a constant collection, I ignore time indexing. For example, I denote a constant allocation $A$ simply as \( \{c_y, l, c_0, K\} \).

### 3. Competitive Equilibrium

In this section I introduce the competitive equilibrium for the economy and discuss some of its properties. To this end, let $Z$ be the set of all the sequences \( \{T_0^t, T_y^t, r_t, w_t, L_t, b_{t+1}, a_{t+1}\}_{t \geq 0} \) where all the elements - but the functions $a_{t+1} \in L^2$ - are real numbers.

**Definition.** A competitive equilibrium (CE) for an economy $E$, given a government policy $\theta \equiv \{\rho_t, \lambda_t, \tau_t\}_{t \geq 0}$, is a collection $(A, z)$ of an allocation $A \equiv \{c_y^t, l_t, c_0^t, K_t\}_{t \geq 0}$ in $A(E)$ and a sequence $z \equiv \{T_0^t, T_y^t, r_t, w_t, L_t, b_{t+1}, a_{t+1}\}_{t \geq 0}$ in $Z$ such that $c_0^t = T_0^0 + R_0^t \int_a^d a_0 di$ and for all $t \geq 0$:

1. Given taxes, transfers and prices, the collection of functions \( \{c_y^t, l_t, c_0^t+1\} \) solves the consumer problem for the generation born in $t$
2. The function $a_{t+1}$ represents the optimal savings for the generation born in $t$
3. The representative firm maximizes profits, i.e.

\[
(3.1) \quad r_t = F_1(K_t, L_t)
\]
\[
(3.2) \quad w_t = F_2(K_t, L_t)
\]

4. The old transfer eliminates the saving mortality risk in the aggregate and the unintended bequests are completely redistributed to the newborns, i.e.

\[
(3.3) \quad T_0^t = \frac{1 - \psi}{\psi} R_t^\psi \int_a^d a_t di
\]
\[
(3.4) \quad T_y^t = (1 - \psi) \left( T_0^t + R_t^\psi \int_a^d a_t di \right)
\]
(5) The resource feasibility constraint is satisfied and all the markets clear, i.e.
\[
\int_{t} c^e_i di + \psi \int_{t} c^o_i di + K_{t+1} + g_t = F(K_t, L_t) + (1 - \delta) K_t
\]
\[
\int_{t} a_t di = K_t + b_t
\]
\[
\int_{t} c_t di = L_t
\]

(6) The government debt lies within the debt limits, i.e.
\[
bF(K_t, L_t) \leq b_t \leq \bar{b}F(K_t, L_t)
\]

When a collection \( C = (A, z) \) which is a competitive equilibrium given some government policy \( \theta \), I say that \( C \) it is induced by \( \theta \). Let \( C(\mathcal{E}) \) be the set of all the competitive equilibria for the economy \( \mathcal{E} \) induced by some government policy \( \theta \). When needed, I write \( C(\theta) \) to stress that a competitive equilibrium \( C \) in \( C(\mathcal{E}) \) is induced by a government policy \( \theta \).

\textbf{Lemma 1.} In a competitive equilibrium, the intertemporal government budget constraint always holds. Further, \( T_t^\theta = T_t^\phi = (1 - \psi) \int_{t} c_t^e di \) for all \( t \geq 0 \).

\textit{Proof.} The intertemporal government budget constraint given by (2.4) holds at every period as a consequence of Walras' law, as one can verify. To see \( T_t^\theta = T_t^\phi = (1 - \psi) \int_{t} c_t^e di \) it suffices to use (2.11) in (3.3) and (3.4). \( \square \)

\textbf{Definition.} Given an economy \( \mathcal{E} \), a steady state competitive equilibrium is a constant competitive equilibrium \( C(\theta) \in C(\mathcal{E'}) \) for some \( \mathcal{E'} \in E(\mathcal{E}) \) such that \( \theta \) is a constant government policy.

Observe that a steady state competitive equilibrium requires public expenditures to be constant for all \( t \).

4. Social Optimum

In this section, I discuss the problem of a benevolent utilitarian social planner who shares the same discount factor \( \phi \) of the government. The social planner maximizes the discounted sum of the generations’ average utilities. Formally, given an economy \( \mathcal{E} \), the planner has to choose an allocation

\textsuperscript{11}Let me remark that when I write \( C(\theta) \), I am not defining a competitive equilibrium as function of \( \theta \). Indeed, in general, a government policy may induce more than one equilibrium (or none). Rather, I am proceeding the opposite way. Namely, given \( C \in C(\mathcal{E}) \), this must be implemented by some policy \( \theta \). When needed, I write \( C(\theta) \) simply to stress this fact.
\( \{ c^y_t, l_t, c^o_t, K_t \} \in A(E) \) to solve the problem:

\[
\max_{A(E)} \left\{ \phi^{-1} \beta \int_{I} u(c^o_0) \, dt + \sum_{t \geq 0} \phi^t \int_{I} U(c^y_t, l_t, c^o_{t+1}) \, dt \right\}
\]

subject to

\[
\int_{I} c^y_t \, dt + \psi \int_{I} c^o_t \, dt + K_{t+1} + g_t \leq F(K_t, L_t) + (1 - \delta) K_t \quad \text{for all } t \geq 0
\]

\[
K_0 = \int_{I} a_0 \, dt - b_0
\]

where \( L_t \equiv \int_{I} e_l \, dt \). I refer to the above maximization as the social planner problem.

**Definition.** An allocation \( \{ c^y_t, c^o_t, l_t, K_t \}_{t \geq 0} \) is optimal for the economy \( E \) if it solves the social planner problem.

**Proposition 1.** An allocation \( \{ c^y_t, c^o_t, l_t, K_t \}_{t \geq 0} \) is optimal for the economy \( E \) if and only if it satisfies the following conditions at all \( t \geq 0 \).

\( (4.1) \quad u_c(c^y_t, l_t) \) and \( c^o_t \) are constant i.a.e.

\( (4.2) \quad (1 + F_1(K_{t+1}, L_{t+1}) - \delta) = \phi^{-1} \frac{u_c(c^y_t, l_t)}{u_c(c^o_{t+1}, l_{t+1})} \quad i \text{- a.e.} \)

\( (4.3) \quad \frac{\beta u_c(c^o_{t+1})}{u_c(c^y_t, l_t)} = \frac{1}{1 + F_1(K_{t+1}, L_{t+1}) - \delta} \quad i \text{- a.e.} \)

\( (4.4) \quad -u_l(c^y_t, l_t) = u_c(c^o_t, l_t) F_2(K_t, L_t) e \quad i \text{- a.e.} \)

\( (4.5) \quad F(K_t, L_t) + (1 - \delta) K_t - \int_{I} c^o_t \, dt - \psi \int_{I} c^y_t \, dt - K_{t+1} - g_t = 0 \)

**Proof.** Conditions (4.1)-(4.4) are derived from the first order optimality conditions of the Lagrangian problem associated to the social planner problem. The last equation is the feasibility constraint satisfied with equality by non-satiation of preferences. Sufficiency follows from the concavity of the maximization problem.

Notice that, under the (in fact not so weak) assumption that all the agents have the same preferences, as a consequence of the utility function concavity, the social planner has an implicit concern for equality. Indeed, condition (4.1) reveals that optimal allocations can be regarded as fairly egalitarian. Old consumption does not vary across agents, independently of their idiosyncratic efficiency. Moreover, all the young agents reach the same marginal utility from consumption.\(^{12}\)

**Definition.** Given an economy \( E \), a steady state optimal allocation is a constant allocation, optimal for some \( E' \in E(E) \).

\(^{12}\)Under the further assumption of separable \( u \), it immediately follows that also the optimal young consumption function is constant and independent of idiosyncratic efficiency.
Notice that as for competitive equilibria, an allocation $A$ can be steady state optimal only if government public expenditures are constant over time.

**Corollary 1.** Suppose $\{c^y, c^o, l, K\}$ is a steady state optimal allocation for the economy $\mathcal{E}$. Then, the following conditions hold.

(4.6) \[ (1 + F_1(k, 1) - \delta) = \phi^{-1} \]

(4.7) \[ \frac{\beta u_c(c^o)}{u_c(c^y, l)} = \phi \ i \ a.e. \]

(4.8) \[ -u_l(c^y, l) = u_c(c^y, l) F_2(k, 1) e \ i \ a.e. \]

(4.9) \[ (F(k, 1) - k\delta) L - \int_I c^y di - \psi \int_I c^o di - g = 0 \]

where $k \equiv \frac{K}{L}$.

**Proof.** The conditions are obtained by rewriting the conditions (4.2)-(4.4) and (4.5), using the definition of steady state optimal allocation and the fact that the partial derivatives of the production function $F$ are homogeneous of degree zero. $\square$

The properties of $F$ assure that there exists one and only one positive $k$ satisfying (4.6). An immediate consequence of the corollary and the Euler equation (2.8) is that, considering a steady state competitive equilibrium $C = (A, z)$, the allocation $A$ satisfies condition (4.7), i.e. the optimal marginal condition on savings, if and only if $R^* = (\phi \psi)^{-1}$. It is useful to study the comparative statics for optimal allocations given two economies, identical in all their components but the probability of survival.

**Corollary 2.** Assume $e = 1$ $i$–almost everywhere and $u$ separable. Let $\{c^y, c^o, l, K\}$ and $\{c'^y, c'^o, l_1, K_1\}$ be steady state optimal allocations corresponding respectively to two economies $\mathcal{E}$ and $\mathcal{E}_1$ identical in all their elements but $\psi$ and $\psi_1$, with $\psi < \psi_1$. Then:

(4.10) $c^y > c'^y, c^o > c'^o$ $i$–a.e. and $L < L_1, K < K_1$

(4.11) $\int_I (c^y + \psi c^o) di - \int_I (c'^y + \psi_1 c'^o) di = (F_2(k, 1) + k(\phi^{-1} - 1))(L - L_1)$

**Proof.** Firstly, notice that separability of $u$ and (4.1) imply young consumption is constant $i$–almost everywhere. Similarly, $e$ constant $i$–a.e. and (4.8) give $l, l_1$ constant $i$–almost everywhere. Then, it results $l = L$ and $l_1 = L_1$ $i$–almost everywhere. Further, since the two economies share the same $F$, condition (4.6) implies that the two allocations share the same $k$. Now, suppose by contradiction $L \geq L_1$. Then, (4.8) implies $c^o \leq c'^o$ $i$–almost everywhere. In turn, this would require by (4.7) that $c^o \leq c'^o$ $i$–almost everywhere. But this leads to, being $\psi < \psi_1$: \[(F(k, 1) - k\delta) L - \int_I c^y di - \psi \int_I c^o di - g > (F(k, 1) - k\delta) L_1 - \int_I c'^y di - \psi \int_I c'^o di - g = 0\]
so that (4.9) is not satisfied by \( \{c^y, c^o, l, K\} \), leading to a contradiction. Therefore, it must be \( L_1 < L_2 \), which by (4.8) and (4.7) gives \( c^y_1 > c^y_2 \) and \( c^o_1 > c^o_2 \) \( i.e. \) almost everywhere. Finally, since \( k \) is equal across the two allocations, it must be \( K_1 < K_2 \). Hence, (4.10) is proved. Equation (4.11) is obtained by taking the difference between the resource feasibility constraints of the two economies, using the homogeneity of degree one of \( F \) and (4.6) to suitably rewrite it.

This comparative statics result shows that as the probability of survival decreases, optimal steady state output decreases as well. The underlying intuition is straightforward. As the survival probability decreases, the mass of population alive in each period goes down as well. Hence, output per capita increases. By concavity of preferences, it is optimal for young agents to reduce their labour supply and enjoy more leisure. In turn, this leads aggregate consumption to fall as it can be seen in (4.11). Notice that the survival probability plays an important role for aggregate consumption. Indeed, even if \( c^o > c^o_i \) \( i.e. \) implies \( \int_t c^o \text{di} > \int_t c^o_1 \), one may have \( \psi \int_t c^o \text{di} < \psi_1 \int_t c^o_1 \text{di} \) since \( \psi < \psi_1 \). I use this result in section 6 to study the comparative statics of the government problem.

5. Ramsey Problem

The government is benevolent and maximizes the same object as the social planner. However, its actions are limited. The government would like to use lump sum transfers conditioned on idiosyncratic efficiency to achieve optimality. In fact, the government cannot observe idiosyncratic efficiency, but only labour income. This is the intuition underlying the assumption that lump sum transfers are not feasible and the government can only choose a policy \( \theta \) to implement a particular competitive equilibrium. This is the key element of a Ramsey problem. Furthermore, Recall that the government faces an exogenously given sequence \( g \) of public expenditures and is committed to give transfers \( T^r_i \) to old agents such that aggregate savings are not subject to mortality risk. It also takes initial conditions of the economy as given. In order to formally introduce the Ramsey problem faced by the government, let me state the following definitions.

**Definition.** Given an economy \( \mathcal{E} \), an allocation \( A \in \mathcal{A}(\mathcal{E}) \) is said to be implementable as a competitive equilibrium if there exists a competitive equilibrium in \( C \in \mathcal{C}(\mathcal{E}) \) such that \( A \in C \).

**Definition.** Given an economy \( \mathcal{E} \), an allocation \( A \in \mathcal{A}(\mathcal{E}) \) is said to be implemented as a competitive equilibrium by a government policy \( \theta \) if there exists a competitive equilibrium \( C(\theta) \in \mathcal{C}(\mathcal{E}) \) such that \( A \in C(\theta) \).

---

13See Chari and Kehoe (1999); Ramsey (1927)
Then, the problem can be characterized as the choice of an allocation \( \{c_t^y, c_t^o, l_t, K_t\}_{t \geq 0} \) that solves:

\[
\max_{A(\mathcal{E})} \left\{ \phi^{-1} \beta \psi \int_{\mathcal{I}} u(c_0^y) dI + \sum_{t \geq 0} \phi^t \int_{\mathcal{I}} U(c_t^y, l_t, c_{t+1}^o) dt \right\}
\]

subject to

\[
\{c_t^y, c_t^o, l_t, K_t\}_{t \geq 0} \text{ is implementable as a CE}
\]

The above maximization is the Ramsey planner problem for the economy \( \mathcal{E} \). This formulation, even though formally correct, is not really comfortable to work with. Indeed, it may be difficult to say, for a given allocation, whether it is implementable as a competitive equilibrium. Following a similar strategy\(^{14}\) to the one adopted by the classical literature, it is possible to rewrite the problem according to the so-called primal approach, finding necessary and sufficient conditions for an allocation to be implementable as a competitive equilibrium.\(^{15}\)

**Proposition 2.** Given an economy \( \mathcal{E} \), an allocation \( A \equiv \{c_t^y, c_t^o, l_t, K_t\}_{t \geq 0} \) is implementable as a competitive equilibrium for some government policy if and only if there exists a sequence \( \{\rho_t\}_{t \geq 0} \) with \( \underline{\rho} \leq \rho_t \leq \overline{\rho} \) and \( \rho_t < 1 \) for all \( t \) such that the following conditions hold true, defined \( L_t \equiv \int_{\mathcal{I}} e_t dt \).

1. The feasibility constraint is satisfied with equality for all \( t \geq 0 \), i.e.

\[
(5.1) \quad \int_{\mathcal{I}} c_t^y dI + \psi \int_{\mathcal{I}} c_t^o dI + K_{t+1} - (1 - \delta) K_t + g_t = F(K_t, L_t)
\]

2. The allocation satisfies, for all \( t \geq 0 \)

\[
(5.2) \quad \beta \psi \frac{u_c(c_{t+1}^o)}{u_c(c_t^o, l_t)} = P_{t+1} \quad i.e. \quad \text{for some } P_{t+1} \in \mathfrak{R}
\]

\[
(5.3) \quad -\frac{u_l(c_t^y, l_t)}{u_c(c_t^y, l_t)(1 - \rho_t)} = \Lambda_t \left( F_2(K_t, L_t) e_t \right)^{1 - \rho_t} \quad i.e. \quad \text{for some } \Lambda_t \in \mathfrak{R}
\]

3. The implementability constraint is satisfied for all \( t \geq 0 \), i.e.

\[
(5.4) \quad u_c(c_t^y, l_t) \left( c_t^y - (1 - \psi) \int_{\mathcal{I}} c_t^o dI \right) + \beta \psi u_c(c_t^o, l_t) \left( c_{t+1}^o - (1 - \psi) \int_{\mathcal{I}} c_{t+1}^o dI \right) + \frac{u_l(c_t^y, l_t)}{1 - \rho_t} = 0
\]

\( i.e. \text{almost everywhere.} \)

4. The debt limits are not violated for all \( t > 0 \), i.e.

\[
(5.5) \quad \underline{b} F(K_{t+1}, L_{t+1}) + K_{t+1} \leq P_{t+1} \left( \psi \int_{\mathcal{I}} c_{t+1}^o dI \right) \leq \overline{b} F(K_{t+1}, L_{t+1}) + K_{t+1}
\]

---

\(^{14}\) handled the conditions given by binding public debt limits building on Piguillem and Shakhnov (2020)

\(^{15}\) See, for example, Chari and Kehoe (1999)
Further, at period zero

\[ b F (K_0, L_0) \leq b_0 \leq \hat{b} F (K_0, L_0) \]  
\[ c_0^\circ = R_0^\circ a_0 + (1 - \psi) \int_I c_0^\circ di \quad i - a.e. \]
\[ K_0 = \int_I a_0 di - b_0 \]
where \( R_0^\circ \equiv 1 + (1 - \tau_0) (F_1 (K_0, L_0) - \delta) \).

**Proof.** First, let me prove the necessity of the above conditions. If an allocation is implementable as a competitive equilibrium \( C \in C (\mathcal{E}) \), it shall be implemented by a government policy which by definition has \( \bar{\rho} \leq \rho_t \leq \rho \) and \( \rho_t < 1 \) for all \( t \). Notice that labour market clearing condition (3.7) gives
\[ L_t = \int_I c_t di \]
Condition (5.1) is implied by (3.5). Similarly, imposing \( P_{t+1} = \frac{1}{R_{t+1}^\circ} \) and \( \lambda_t = \lambda_t \) and recalling that \( w_t = F_2 (K_t, L_t) \) by (3.2), conditions (5.2) – (5.3) follow from (2.8) – (2.9). Then, (5.4) is obtained substituting \( R_{t+1}^\circ, \lambda_t \) in the present value budget constraint (2.10) using (2.8) – (2.9), appropriately rearranging it, and relying on Lemma 1 to replace \( T_{y_t}^o \) and \( T_{o_t+1}^o \). Finally, observe that \( C \) must be such that
\[ a_{t+1} = \frac{1}{R_{t+1}^\circ} (c_{t+1}^o - T_{t+1}^o) \quad i - a.e. \]
\[ \int_I a_{t+1} di = R_{t+1}^\circ \left( \int_I c_{t+1}^o di - T_{t+1}^o \right) \]
Using the asset market clearing condition (3.6), Lemma 1, the Euler equation (2.8) and recalling
\[ P_{t+1} = \frac{1}{R_{t+1}^\circ} \] it results
\[ b_{t+1} = P_{t+1} \left( \psi \int_I c_{t+1}^o di \right) - K_{t+1} \]
which, together with the public debt constraints (3.8), proves the necessity of (5.5). Conditions at \( t = 0 \) are trivially satisfied. Now, I show sufficiency. That is, given an arbitrary allocation \( A \) which satisfies all the above conditions for some \( \{ \rho_t \}_{t \geq 0} \), I need to find a sequence
\[ \{ T_{y_t}^o, T_{o_t}^o, r_t, w_t, L_t, b_{t+1}, a_{t+1} \}_{t \geq 0} \equiv z \in Z \]
and positive real numbers \( \{ \lambda_t, \tau_{t+1} \}_{t \geq 0} \) such that \( C = (A, z) \) is a competitive equilibrium for \( \mathcal{E} \) given the government policy \( \{ \lambda_t, \rho_t, \tau_{t+1} \}_{t \geq 0} \). To this end, I do the following for all \( t \geq 0 \). Take \( L_t = \int_I c_t di \) to satisfy labour market clearing (3.7). Then, set \( \lambda_t = \lambda_t, w_t = F_2 (K_t, L_t) \) and \( r_t = F_1 (K_t, L_t) \) so that the intratemporal labour condition (2.9) holds and the firm is maximizing profits. Then, it is always possible to set \( \tau_{t+1} \) such that \( \frac{1}{R_{t+1}^\circ} = P_{t+1} \) so that the Euler equation (2.8) holds. Further,
impose $T_t^o = T_t^p = (1 - \psi) \int_x c_t^o di$ to satisfy Lemma 1. Then, (5.4) and (5.2)-(5.3) ensure that the present value budget constraint (2.10) holds. Hence, the consumer problem is solved for the generation born in $t$. Consistently, define

$$a_{t+1} = \frac{1}{R_{t+1}} \left( c_{t+1}^o - T_{t+1}^o \right)$$

so that it represents the optimal savings of the generation born in $t$. Then, to satisfy the asset market clearing condition (3.6) choose $b_{t+1} = \int_x a_{t+1} di - K_{t+1}$. Notice that for $t = 0$ the asset market clears by assumption. Then, the debt limits (3.8) are satisfied by (5.5) and (5.6). The condition on time zero old generation consumption holds by (5.7). Finally, (3.1) is satisfied at every period by (5.1). Limits on $\rho_t$ hold by assumption.

The result proves that the present economic environment allows the government to impose non linear labour taxes while maintaining the standard framework of a classical Ramsey problem. Notice that to determine the implementability of an allocation, I introduced some ancillary parameters which do not enter the maximizing object, namely $\{A_t, \rho_t, P_{t+1} \}_{t \geq 0}$. Let $\Pi$ be the set of all the possible ancillary parameters. It is simply a collection of real numbers, with the specification $\rho_t < 1$ for all $t \geq 0$. I remark that $\Pi$ is defined independently of the economy $\mathcal{E}$. Then, denoting as $\Xi(\mathcal{E})$ the set of all the necessary and sufficient conditions\footnote{Notice that $\rho \leq \rho_t \leq \overline{\rho}$ for all $t \geq 0$ is part of $\Xi(\mathcal{E})$.} of Proposition 2, the Ramsey problem can be conveniently rewritten as it follows.

$$\max_{A(\mathcal{E}) \times \Pi} \left\{ \phi^{-1} \beta \psi \int_I u(c_0^o) di + \sum_{t \geq 0} \phi^t \int_I U(c_t^o, l_t, c_{t+1}^o) di \right\}$$

subject to

all the conditions $\Xi(\mathcal{E})$ are satisfied

Observe that the above proof implies that if $(A, x)$ satisfies all the conditions in $\Xi(\mathcal{E})$ and thus can be implemented as competitive equilibrium, then one has $P_{t+1} = \frac{1}{R_{t+1}}$ at all periods. Considering the new formulation of the problem it is possible to write down the associated Lagrangian problem and derive the first order optimality conditions. Nevertheless, I avoid to report them here as they are space consuming. The following definitions are preliminary to the next section.

**Definition.** An allocation $A \equiv \{c_t^o, c_t^a, l_t, K_t\}_{t \geq 0}$ is Ramsey optimal for the economy $\mathcal{E}$ if there exists $x \in \Pi$ such that $(A, x)$ solves the associated Ramsey planner problem. The government policy $\theta$ which implements $A$ as a competitive equilibrium is Ramsey optimal for the economy $\mathcal{E}$.

**Definition.** Given an economy $\mathcal{E}'$, a steady state Ramsey optimal allocation is a constant allocation $A$ which is Ramsey optimal for some $\mathcal{E}' \in E(\mathcal{E})$ and it is implemented by a constant government policy.
Equivalently, I may say that $A$ is steady state Ramsey optimal if it is Ramsey optimal and can be implemented as a steady state competitive equilibrium. As usual, a steady state Ramsey optimal allocation requires public expenditure to be constant over time.

It is worth noticing a technicality in the above definition which may be not immediately evident. Saying that a steady state Ramsey optimal allocation is implemented by a constant government policy, I am only specifying that the capital tax rate is the same for all $t$, including period zero.\footnote{Indeed, even if it is exogenously specified given an economy $\mathcal{E}$, it is free to vary when considering $E(\mathcal{E})$.}

6. RAMSEY OPTIMUM

In this section, I present some results on the second-best achievable by the government. In general, intragenerational heterogeneity in labour productivity prevents the Ramsey planner from achieving an optimal allocation. Indeed, no government policy is able to implement an allocation which satisfies the marginal optimal condition on labour choice (4.4) while keeping the old consumption function constant as required by (4.1). However, in the special case of homogeneous efficiency, the assumed labour tax - non necessarily linear - functional form allows the Ramsey planner to achieve the same allocation of the Social Planner in steady state under mild assumptions. Notice that in the classical literature the homogenous agent assumption is quite common, while the labour tax is usually restricted to be linear.\footnote{See Chari and Kehoe (1999)}

Before proving the main result, it is convenient to state the following lemma.

**Lemma 2.** Consider an economy $\mathcal{E}$, and let $A$ be a Ramsey optimal allocation implemented by a government policy $\theta \equiv \{\lambda_t, \rho_t, \tau_t\}_{t \geq 0}$. Let $x = \{\Lambda_t, \bar{\rho}_t, P_{t+1}\}_{t \geq 0}$ be the element in $\Pi$ such that $(A, x)$ solves the Ramsey problem. Then, for all $t \geq 0$

\begin{align*}
(1) \quad & \lambda_t = \Lambda_t \quad \text{and} \quad \rho_t = \bar{\rho}_t \\
(2) \quad & \tau_{t+1} = 1 - \frac{1}{(1-\lambda_t)(K_{t+1},L_{t+1})} \neq P_{t+1}
\end{align*}

Proof. (1) is evident from the proof of Proposition 2. To obtain (2), it suffices to recall from the proof of Proposition 2 that for a government policy to implement $A$, one has to set $\tau_{t+1}$ so that $P_{t+1} = \frac{1}{R_{t+1}}$ for all $t \geq 0$. Then, recalling the definition of $R_t^*$ and the equilibrium condition $r_t = F_1(K_t, L_t)$ I obtain:

$$P_{t+1} = \frac{1}{1 + (1 - \tau_{t+1}) (F_1(K_{t+1}, L_{t+1}) - \delta)}$$

which - suitably rearranged - completes the proof. \hfill $\square$

The above result defines a map from a Ramsey problem solution $(A, x)$ to the policy which implements the allocation $A$ as a competitive equilibrium. That is, determining the element $x$ in $\Pi$ such that $(A, x)$ solves the Ramsey planner problem is equivalent to solving for the policy $\theta$ which implements $A$. In particular, Lemma 2 implies that the condition for a steady state Ramsey optimal allocation to be implemented by a constant policy is satisfied if and only if the economy $\mathcal{E}' \in E(\mathcal{E})$ for which the constant allocation $A$ is Ramsey optimal specifies $\tau_0 = \tau$ where $\tau$ is defined by (2).
Theorem 1. Consider an economy $E$ such that $e = 1$ $i$–almost everywhere. Assume at least one between

1. $\phi \psi^2 + \psi \geq 1$
2. $\phi \leq \beta$ and $u_{ct} \leq 0$

Further, let $\Theta$ be such that constraints on debt (2.5) and progressivity (2.6) can be ignored. If $A \equiv \{c^y, c^o, l, K\}$ is a steady state optimal allocation, then it is also a steady state Ramsey optimal allocation.

Proof. Since $A$ is steady state optimal, it solves the social planner problem for some set of initial conditions. By steady state definition, $K$ must be the initial capital defined by this set of initial conditions. Notice that initial conditions enter the social planner problem only through $K$. Further, the sequence of public expenditures must be constant over time at some value $g \geq 0$. Thus, for all the economies in $E(\delta')$, as long as initial conditions are such that the initial capital is $K$ and public expenditure is $g$ for all $t$, the Ramsey planner is maximizing the same object of the social planner. However, the constrained choice set of the Ramsey problem is contained in the one of the social planner. Hence, to prove the statement, it is sufficient to find $x \in \Pi$ such that its elements do not change over time and $(A, x)$ satisfies all the conditions $\Xi(\delta')$ for some $\delta' \in E(\delta')$ where the implied initial capital is $K_0 = K$ and $g_t = g$ constant over time. Furthermore, recall that $A$ must be implemented by a constant government policy. To this end, I rely on Proposition 1 and Corollary 1. To be explicit, let $x = \{\Lambda, \rho, P\}$. Debt and progressivity constraints hold by assumption. Observe that any optimal allocation will satisfy with equality the feasibility constraint (5.1) by (4.9). The optimal allocation satisfies (4.7), thus setting:

\[ P = \psi \phi \]

condition (5.2) is satisfied. Moreover, $A$ satisfies (4.8). Further, given the assumption on $e$, (4.8) and (4.1) imply $c^o, l$ are constant $i$–almost everywhere. Then, using the definition of $L$, I can write $l = L$ $i$–almost everywhere. Hence, using (4.4) in (5.3) I have that imposing

\[ (A, x) \text{ satisfies (5.3). Now, it is possible to rewrite the implementability constraint (5.4) as it follows.} \]

\[ \left( c^y - (1 - \psi) \int_I c^o di \right) + P \left( c^o - (1 - \psi) \int_I c^o di \right) + \frac{u_t(c^y, l)}{(1 - \rho) u_e(c^y, l)} = 0 \quad i - a.e. \]

Using the fact that $c^y$ is constant $i$–a.e. - as it is deduced above - and (4.1), it is possible to write $c^y = \int_I c^y di$ and $c^o = \int_I c^o di$ $i$–almost everywhere. Hence, using (4.8) and the other results I showed in the proof it is possible to rearrange the previous equation as:

\[ \int_I c^y di - (1 - \psi - \psi^2 \phi) \int_I c^o di - \frac{1}{1 - \rho} F_2(K, L) L = 0 \]
Thus, the implementability constraint is satisfied setting:

\[ \rho = 1 - \frac{F_2(K, L)}{\int_T c^\rho di - (1 - \psi - \psi^2 \phi) \int_T c^\sigma di} \]

Since the progressivity bounds can be ignored by assumption, I only have to ensure \( \rho < 1 \). To this end, it suffices to show that the denominator \( \int_T c^\rho di - (1 - \psi - \psi^2 \phi) \int_T c^\sigma di \) is positive. If (1) holds, then it is trivially positive. If (2) is true, then it is easy to show its positivity using (4.7) and the concavity of \( u \). Then, Lemma 2 implies \( A \) is implemented by a constant government policy if and only if

\[ \tau'_0 = \tau \equiv 1 - \frac{1 - F}{(F_1(K, L) - \delta) \rho} \] .

Hence, I have to properly choose \( a'_0, b'_0 \) such that \( K'_0 = K \) and

\[ c^\rho = (1 + (F_1(K, L) - \delta) (1 - \tau)) a'_0 + (1 - \psi) \int_T c^\sigma di \ i.e. \ a.e. \]

Recalling \( c^\sigma = \int_T c^\sigma di \ i.e. \text{almost everywhere} \), the above equation can be satisfied defining \( a'_0 \) as a constant function, according to

\[ a'_0 \equiv (1 + (F_1(K, L) - \delta) (1 - \tau))^{-1} \psi \int_T c^\sigma di \]

Finally, to satisfy the condition on initial capital stock set \( b'_0 = \int_T a^\sigma di - K \). Notice that \( A \) is implemented by \( \{\lambda, \rho, \tau\} \) as a steady state competitive equilibrium. \( \square \)

The theorem is highly interesting since it reveals that, if the agents are homogeneous and there are no tight public debt limits, the Ramsey planner is able to achieve the social optimum in steady state as long as the government is free to choose labour tax progressivity.\(^{19}\) In fact, notice that the classical case where the government is restricted to tax labour income linearly may be obtained as a special case imposing \( \rho = \bar{\rho} = 0 \). Under this setting, the assumption of Theorem 1 does not hold since in general progressivity bounds cannot be ignored. In other words, even tough the other assumptions of Theorem 1 hold true, in the classical setup the Ramsey planner would not be able to achieve first best optimality. Thus the proposed generalization for the labour tax function appears to be rather significant. Indeed, while maintaining the set up of a standard Ramsey problem, the planner is now able to reach optimality in steady state. The intuition is that the Ramsey planner can mimic lump sum transfers through the labour tax progressivity. This is possible since the agents are assumed to be homogeneous, thus progressivity has no effect on intragenerational consumption distribution. All the agents are going to work the same, get the same income and consume the same.

**Corollary 3.** Suppose \( \mathcal{E} \) is such that the assumptions of the Theorem hold true. Let \( A \equiv \{c^\rho, c^\sigma, l, K\} \) be a steady state optimal allocation and \( \mathcal{E}' \in E(\mathcal{E}) \) be the economy for which \( A \) is a steady state Ramsey optimal allocation. Then, \( A \) is implemented as a steady state competitive equilibrium for \( \mathcal{E}' \)

\(^{19}\) Notice that progressivity may be negative, i.e. the tax function may be regressive.
by the constant policy \( \{ \lambda, \rho, \tau \} \) where \( \rho \) is given by (6.3) and \( \lambda, \tau \) are defined as follows.

\[
(6.4) \quad \lambda = \frac{(F_2(k, 1)L)^\rho}{1 - \rho}
\]

\[
(6.5) \quad \tau = 1 - \frac{1 - \psi \phi}{(1 - \delta) \psi}
\]

where \( k \) is defined as in Corollary 1.

**Proof.** The result follows from Lemma 2 and (6.1)-(6.3), using condition (4.6) of Proposition 1 to obtain (6.5). \( \square \)

Equation 6.5 implies that under the assumptions of the Theorem the steady state Ramsey optimal capital tax is always nonpositive and negative whenever \( \psi < 1 \). The intuition for this result is that to make agents save optimally in a competitive equilibrium, the government needs to subsidize capital in order to perfectly compensate the mortality risk \( 1 - \psi \) they are subject to. Notice that this result holds true even if the Ramsey planner does not share the same time discount factor of the agents.

In general, the government needs to accumulate a relevant portion of the assets in the economy to make satisfy its steady state Ramsey optimal intertemporal budget constraints. As a consequence - given an economy \( \mathcal{E} \) - it may be possible for a Ramsey optimal allocation to never converge to such a steady state, even if the assumptions of the Theorem hold true.\(^{20}\) I address this issue in the numerical analysis of section 10.

Considering the labour tax (2.2), (6.4) give that in equilibrium, where \( wL = wL \) i.e. and \( w = F_2(K, L) \), the government gets nonnegative revenues if and only if \( \rho \) is nonpositive, i.e. if the tax system is regressive. Intuitively, this establishes a connection between \( \rho \) and \( g \): an increase in public expenditure implies the government needs more money to satisfy its intertemporal budget constraint, calling for a more regressive labour taxation. Indeed, determining the sign of \( \rho \) is of a certain relevance, since it provides an important efficiency benchmark when I relax the assumption of \( e \) constant \( i \) – almost everywhere. If \( e \) is constant, the progressivity of labour tax is driven only by efficiency concerns. When this assumption is relaxed, the problem is more complicated since the Ramsey planner also faces inequality concerns. However, the sign of \( \rho \) under the assumption of a constant \( e \) gives important insights to interpret the numerical solutions of the more general case. In particular, for reasonably calibrated parameters, when agents are homogenous the optimal labour tax is regressive. Furthermore, building on Corollary 2 I can show the following result.

**Corollary 4.** Let the assumptions of Theorem 1 hold and \( u \) be separable. Consider a family of economies \( \{ \mathcal{E}_\psi \} \) identical in all their elements but the probability of survival \( \psi \in (0, 1] \). Let \( A_\psi \) denote a steady state optimal allocation for the economy \( \mathcal{E}_\psi \). Let \( \psi_1 \leq 1 \) be such that \( A_{\psi_1} \equiv \{ c^1_t, c^1_t, l_t, K_t \} \)

\(^{20}\) I say that an allocation \( A \equiv \{ c^\alpha_t, l_t, c^\alpha_t, K_t \}_{t \geq 0} \) converges to a constant allocation \( A' \equiv \{ c^\alpha, l, c^\alpha, K \} \) if the sequence \( \{ c^\alpha_t, l_t, c^\alpha_t, K_t \}_{t \geq 0} \) converges to \( \{ c^\alpha, l, c^\alpha, K \} \) in the product topology defined on \( (L^2)_\alpha \times \mathbb{R}_+ \).
OPTIMAL TAXATION IN AN OVERLAPPING GENERATIONS ECONOMY

is implemented by a steady state Ramsey optimal government policy with \( \rho_1 \leq 0 \). Then, any \( A_{\psi} \) such that \( \psi < \psi_1 \) is implemented as a steady state competitive equilibrium by a Ramsey optimal government policy with negative progressivity parameter \( \rho_{\psi} \).

**Proof.** Notice that by Theorem 1, there is always a Ramsey optimal government policy that implements \( A_{\psi} \) as a steady state competitive equilibrium. For some \( \psi < \psi_1 \), let \( A_{\psi} = \{ c^0, c'^0, l, K \} \) be implemented by a government policy \( \{ \lambda, \rho, \tau \} \). Arguing as in Theorem 1, by the implementability constraints (5.4) for \( A_{\psi} \) and \( A_{\psi_1} \) I may write the following.

\[
\int_I c^0 di - (1 - \psi_1 \psi_1^2) \int_I c'^0 di \geq F_2(k, 1) L_1
\]

\[
\int_I c^0 di - (1 - \psi - \psi^2 \phi) \int_I c'^0 di = \frac{1}{1 - \rho} F_2(k, 1) L
\]

where \( k \) is defined as in Corollary 1. Notice that the two allocations share the same \( k \) by (4.6). Subtracting the first equation to the second one and rearranging one obtains

\[
\int_I (c^0 + \psi c'^0) di - \int_I (c'^0 + \psi_1 c'^1) di + (\psi^2 \phi - 1) \int_I c'^0 di + (1 - \psi_1^2 \phi) \int_I c'^1 di \geq \frac{F_2(k, 1) L}{1 - \rho} - F_2(k, 1) L_1
\]

which using Corollary 2 leads to write

\[
1 - \rho > 1
\]

which requires \( \rho < 0 \). \( \blacksquare \)

Notice that under these assumptions consumption functions are constant \( i \)–almost everywhere. The result may be helpful in determining when the tax on labour will be regressive looking to the survival probability. It is sufficient to find \( \psi_1 \leq 1 \) such that the steady state Ramsey optimal labour tax is not progressive to claim that it will be regressive for all \( \psi < \psi_1 \). Recall that a negative \( \rho \) under the assumptions of Theorem 1 is an important efficiency benchmark for the cases where idiosyncratic efficiency is not homogenous. That is, because of efficiency reasons the government would like to implement regressive labour taxation. This force is still present and affects the results even in the more general problem.

So far, I have ignored the constraints on public debt. In fact, they may be extremely important in determining capital taxation, as I quantitatively discuss in section 10. As long as the constraints are not binding, the government can use debt to achieve optimal level of capital investment introducing no distortion in the agents saving decision. That is, the government can reach at the same time optimal old consumption and capital investment. The proof of Theorem 1 is helpful to understand this point. Look to equation (6.1): to implement the steady state optimal allocation as a competitive equilibrium, the government set \( \tau \) such that the Euler equation of the agents is the same of the optimality condition (4.7) required by the social planner. Then, debt is determined (at the end of the proof) as a residual such that the allocation can be implemented as a steady state competitive equilibrium. When the
debt channel is crushed by a binding constraint, this procedure does not work and the government may be forced to introduce some distortion in agents saving decisions, as it is claimed by the following proposition.

**Proposition 3.** Consider an economy $\mathcal{E}$ and let the allocation $A = \{c_t, l_t, c_0, K_t\}_{t \geq 0}$ be Ramsey optimal. For all $t \geq 1$, if one of the two debt limits implied by (5.5) is binding at $t$, assuming $F_1(K_t, L_t) \neq \delta$, the policy $\theta$ which implements $A$ as a competitive equilibrium must specify:

$$\tau_t = 1 - \frac{\psi \int \bar{c}_t \, dt - K_t}{K_t (F_1(K_t, L_t) - \delta)}$$

where $\bar{b}$ is either $b$ or $\bar{b}$, depending of which of the two bounds is binding.

**Proof.** By construction, the binding borrowing limit implies:

$$P_t^{-1} = \frac{\psi \int \bar{c}_t \, dt}{\bar{b}F(K_t, L_t) + K_t}$$

Recalling it must be $P_t = 1 + (F_1(K_t, L_t) - \delta)(1 - \tau_t)$ and combining the two one completes the proof.

The interpretation of the result is straightforward when $\bar{b} = 0$. Indeed, in that case it is possible to rewrite (6.6) as

$$\tau_t = 1 - \frac{\psi \int \bar{c}_t \, dt - K_t}{K_t (F_1(K_t, L_t) - \delta)}$$

which, recalling in equilibrium $R_t = 1 + F_1(K_t, L_t) - \delta$ implies a positive capital tax whenever

$$\psi \int \bar{c}_t \, dt < R_t K_t$$

that is, when the Ramsey optimal aggregate pre-tax gross capital income is greater than the optimal aggregate consumption of the old generation in $t$. This is a consequence of the fact that the government cannot hold debt and therefore the only way to keep old consumption away from gross capital income is to tax it. Notice that this may be the case especially public expenditure is positive.

7. **An Example with Logarithmic Utility**

In this section I study an example I use for calibration. Throughout this section I assume, setting $\bar{c} = 0$ and $\bar{I} = 1$, the following preference specification for $\gamma \in (0, 1)$:

$$u(c, l) \equiv \gamma \log(c) + (1 - \gamma) \log(1 - l)$$

Furthermore, for analytical tractability I deviate from the assumptions of the main model, abstracting from transfers to the old and assuming that unintended bequests are lost. Notice that as the probability $\psi$ goes to one, this deviation tends to disappear - since the transfers go to zero - and in the calibrated economy I have $\psi \simeq 0.915$. Note that the function $\epsilon$ is not assumed to be constant. Indeed, the function
is estimated directly from the data. Considering the conditions for an allocation to be implementable as a competitive equilibrium defined by Proposition 2, one can show that at all periods consumption and labour functions of the newborn generation are uniquely determined (i.a.e.) by the ancillary parameters \( \{ \Lambda_i, \rho_i, P_i \} \) and the capital \( K_i \) as it follows.

\[
\begin{align*}
  l_t &= \frac{\tilde{\gamma} (1 - \rho_t) (1 + \beta \psi)}{1 + \tilde{\gamma} (1 - \rho_t) (1 + \beta \psi)} \\
  c_t^\psi &= (1 + \beta \psi)^{-1} \Lambda_t (F_2 (K_t, L_t) \epsilon L_t)^{1 - \rho_t} \\
  c_{t+1}^\rho &= \frac{\beta \psi}{P_{t+1}} c_t^\psi
\end{align*}
\]

where \( \tilde{\gamma} \equiv \frac{\gamma}{1 - \tilde{\gamma}} \). Therefore, the Ramsey problem can be characterized as a maximization with respect to the collection \( \{ \Lambda_t, P_{t+1}, K_{t+1}, \rho_t \}_{t \geq 0} \) taking \( K_0 \) as given. Further, constraints can be reduced to the feasibility constraint (2.5), the debt limits implied by (5.5) and the bounds progressivity specified by (2.6). For space reasons and not boring the reader, below I report the first order optimality conditions for the case in which the policy constraints are not binding at \( t \). Clearly, for the numerical simulation I consider the full first order conditions. Since in the numerical analysis I consider a steady state allocation, I ignore time zero.

\[
\begin{align*}
  (\Lambda_t): & \quad \gamma + \beta \psi \gamma - \left( \zeta_t + \phi \zeta_{t+1} \frac{\beta \psi^2}{P_{t+1}} \right) \int_I c_t^\psi \, di = 0 \\
  (P_{t+1}): & \quad -\beta \psi \gamma + \phi \zeta_{t+1} \frac{\beta \psi^2}{P_{t+1}} \int_I c_t^\psi \, di = 0 \\
  (K_{t+1}): & \quad -\zeta_t + \phi \zeta_{t+1} (F_1 (K_{t+1}, L_{t+1}) + 1 - \delta) = 0 \\
  (\rho_t): & \quad \gamma (1 + \beta \psi) \int_I z_t \, di - \frac{1 - \gamma}{1 - l_t} + \zeta_t F_2 (K_t, L_t) - \left( \zeta_t + \phi \beta \psi \frac{\zeta_{t+1}}{P_{t+1}} \right) \int_I z_t \, di = 0
\end{align*}
\]

where \( \zeta_t \) is the Lagrangian multiplier attached to the feasibility constraint and

\[
z_t \equiv -c_t^\psi \log (F_2 (K_t, L_t) \epsilon L_t) (F_2 (K_t, L_t) + F_22 (K_t, L_t) L_t) e
\]

Clearly, in steady state one can forget of the time index. In particular, \( \zeta_t = \zeta_{t+1} = \zeta \).

**Proposition 4.** Let the assumptions of this section hold true. Suppose \( A \) is a steady state Ramsey optimal allocation for \( \Theta \) such the debt limits are not binding, implemented by a constant policy \( \theta \equiv \{ \lambda, \rho, \tau \} \). Then, \( \tau \) is given by (6.5).

**Proof.** Using the first order optimality conditions with respect to \( \Lambda_t \) and \( P_{t+1} \) one deduces \( P = \psi \phi \). The first order condition with respect to \( K_{t+1} \) gives \( R^\tau = \phi^{-1} \). Then, it suffices to recall that in equilibrium it must be \( \frac{1}{P^\tau} = P \), use the definition of \( R^\tau \) and rearrange the terms to complete the proof. \( \square \)
Observe that Proposition 4 is not trivial since here I am not assuming $e$ is constant $i$–almost everywhere. That is, if the government debt budget constraints can be ignored, it is Ramsey optimal to subsidize capital investment to compensate the mortality risk the agents are subject to, independently of intergenerational efficiency heterogeneity. The result relies on the fact that under these preferences the labour choice is uniquely determined by $\rho_t$ and it is the same for all the agents, as it is evident from (7.1). Furthermore, recalling $L_t = \int_x e_l \, dl$ , it follows $L_t = l_t$ $i$–almost everywhere.

### 8. Recursive Approach

To study the dynamics of a Ramsey optimal allocation $A$ for a given economy $\mathcal{E}$, it is of great value to write the problem recursively. To do this, it is convenient to assume the Ramsey planner is subject to an additional constraint, namely the average value $C_0^o > 0$ of the consumption function of the old generation at period zero. Moreover, I shall assume the sequence of public expenditure to be constant over time at the value $g \geq 0$. Given the other initial conditions specified by the economy, the consumption function of the old generation at period zero results to be exogenously given. Consistently, I rewrite the Ramsey problem as it follows.

$$\max_{A(\mathcal{E}) \times \Pi} \sum_{t \geq 0} \phi^t \int_I U \left( c_t^y, l_t, c_t^{o+1} \right) \, dl$$

subject to

all the conditions in $\Xi(\mathcal{E})$ are satisfied

$$\int_I c_t^o \, dl = C_0^o$$

To find the recursive functional equation, it is necessary to identify the relevant state variables of the problem. Considering its structure, it is possible to see that the relevant information provided by initials conditions are now the integral value $C_0^o$, the initial capital stock $K_0$, and the level of public debt at zero $b_0$ to take into account the debt limits. However, since debt at the next periods is not directly chosen by the Ramsey planner, I need to introduce a different state variable with the same informational content. To this end, let me define $P_0$ as the positive real such that $P_0 \psi C_0^o - K_0 = b_0$. Furthermore, let $C_t^y$ and $C_t^o$ denote the average value of young and old consumption consumption functions at period $t$, respectively. Then, the state variables at $t$ are given by the vector $s_t = (C_t^y, K_t, P_t)$. Let $S$ denote the state space. The control variables then are the functions $c_t^y, l_t, c_t^{o+1}$ and the ancillary parameters $\Lambda_t, \rho_t$. Let $h_t = (c_t^y, l_t, c_t^{o+1}, \rho_t, \Lambda_t)$ denote the vector of control variables at time $t$. Lastly, I have to define the choice set at $t$. Let $\Gamma : S \supset H \times S$ be the choice set correspondence, such that for all $s_t \in S$ one has $(h_t, s_{t+1}) \in \Gamma (s_t)$ if and only if

$$C_t^y + \psi C_t^o + K_{t+1} - (1 - \delta) K_t + g = F(K_t, L_t) \tag{8.1}$$

$\footnote{Then, aggregate old consumption is $\psi C_0^o$.}$
\( u(c^y_t, l_t) (c^y_t - (1 - \psi) C^o_t) + \beta \psi u_c(c^o_t, l_t) (c^o_t - (1 - \psi) C^o_{t+1}) + \frac{u_l(c^y_t, l_t) l_t}{1 - \rho_t} = 0 \) \( i \) a.e.

\( (\text{8.3}) \)

\[ V(s_t) = \max \left\{ \int_U U(c^y_t, l_t, c^o_{t+1}) \, di + \phi V(s_{t+1}) \right\} \]

where \( V \) is a real valued function defined on the state space \( S \).

9. An Example with Aggregation

This section heavily relies on Piguillem and Shakhnov (2020). The aim is to make the recursive problem numerically tractable, that is, to reduce the dimensionality of the problem, making consumption and labour functions to be fully determined by their integral values. Furthermore, this allows to forget the constraint (8.2) and simplify the implementability constraint (8.2). To this end I assume \( \xi = 0, \eta = 1 \) and the following instantaneous utility function:

\[ u(c, l) \equiv c^\gamma (1 - l)^{1-\gamma} \frac{1}{1 - \sigma} \]  \( \text{where } \sigma > 1, \gamma \in (0, 1) \)

Notice that the instantaneous utility specification (7.1) used for the section 7 steady state analysis can be viewed as a special case of (9.1).\(^{22}\) As for section 7, analytical tractability requires a deviation from the assumptions of the main model on transfers. In particular, in this section I assume that the transfers are not equal for all agents, but are distributed according to a weight function \( \omega_t \equiv e^{1-\rho_t} (\int_x e^{1-\rho_t} \, di)^{-1} \). That is, transfers are assumed to be proportional to the idiosyncratic efficiency of the agents. A possible justification for this assumption is intergenerational persistence: more efficient individuals are likely to come from efficient families. Notice that \( \omega_t \) endogenously depends of \( \rho_t \) and assuming \( e^{1-\rho_t} \in L^2 \) one has \( \int_x \omega_t \, di = 1 \). The only difference with the previous setup is that the

\(^{22}\)In fact, to have pointwise convergence as \( \sigma \downarrow 1 \) one would need to subtract 1 from the numerator of (9.1). However, the aggregation results would hold the same.
implementability constraint (8.3) now reads as

\[(9.2) \quad u_c (c^o_t, l_t) (c^o_t - (1 - \psi) \omega_t C^o_t) + \beta \psi u_c (c^o_{t+1}) (c^o_{t+1} - (1 - \psi) \omega_t C^o_{t+1}) + \frac{u_t (c^o_t, l_t) L_t}{1 - \rho_t} = 0 \quad i - \text{a.e.} \]

Then, it is possible to show that the recursive Ramsey problem can be written only using aggregate variables. To this, I introduce two preparatory lemmas.

**Lemma 3.** Let the assumptions of this section hold true and consider the vector \(s_t \equiv (C^o_t, K_t, P_t)\) in the state space \(S\). Suppose the control variables \(c^o_t, l_t, c^o_{t+1}, \rho_t\) and the state variable \(K_{t+1}\) are such that (8.1), (8.4), (8.5) holds and the following conditions are satisfied.

\[(9.3) \quad l_t = L_t \quad i - \text{a.e.} \]

\[(9.4) \quad c^o_t = \omega_t C^o_t \quad \text{and} \quad c^o_{t+1} = \omega_t C^o_{t+1} \quad i - \text{a.e.} \]

\[(9.5) \quad u_c (C^o_t, L_t) (C^o_t - (1 - \psi) C^o_t) + \beta \psi^2 u_c (C^a_{t+1}) C^a_{t+1} + \frac{u_t (C^o_t, L_t) L_t}{1 - \rho_t} = 0 \]

Then, there exists \(P_{t+1}, \Lambda_t \in R^+\) such that for \(h_t \equiv (c^o_t, l_t, c^o_{t+1}, \rho_t, \Lambda_t)\) and \(s_{t+1} \equiv (C^o_{t+1}, K_{t+1}, P_{t+1})\) one has \((h_t, s_{t+1}) \in \Gamma (s_t)\).

**Proof.** See the Appendix.

Hence, there are sufficient conditions on aggregate values to determine if a pair \((h_t, s_{t+1})\) is feasible given a state variables vector \(s_t \in S\). To reduce the problem to aggregate values, I need to show that satisfying the conditions of Lemma 3 is necessary for \((h_t, s_{t+1})\) to be in \(\Gamma (s_t)\). However, the reverse of Lemma 3 needs some additional assumption to prove it. The complication arises due to the presence of transfers. Of course, it is possible to fix that assuming \(\psi = 1\), eliminating transfers to the agents. However, as I state in the next Lemma, another possibility is to assume a zero lower bound for labour tax progressivity. Since the numerical simulation takes Italy as the benchmark economy, this appears as a quite reasonable assumption.\(^{23}\)

**Lemma 4.** Let the assumptions of this section hold true and suppose \(\rho \geq 0\). Then for all \(s_t \in S\), one has \((h_t, s_{t+1}) \in \Gamma (s_t)\) only if conditions of Lemma 3 are satisfied.

**Proof.** See the Appendix.

It is clear that the two lemmas allow to get rid of (8.2), and simplify the control variables vector to \(q_t \equiv (C^o_t, L_t, \rho_t)\). I remark that \(\Lambda_t\) is an ancillary parameter, i.e. it does not enter the object maximized by the Ramsey planner therefore it can be dismissed, since it does not enter any more in the problem constraints. Furthermore, observe that using the law of \(u_9\) (9.5) can be rearranged as

\[(9.6) \quad C^o_{t+1} = C^o_t (1 - L_t) \frac{1 + \gamma}{1 - \gamma} \left\{ \frac{1 - \gamma}{\beta \psi^2} \frac{L_t}{1 - \rho_t} - \left( 1 - (1 - \psi) \frac{C^o_t}{C^o_t} \right) \right\} \]

\(^{23}\)See footnote 9 in section 2.
Let $\mathcal{M}(s_t)$ denote the simplified choice set correspondence, defined such that $(q_t, s_{t+1}) \in \mathcal{M}(s_t)$ if and only if given $s_t$ the pair $(q_t, s_{t+1})$ satisfies conditions (8.1), (8.4), (8.5) and (9.6). Then, I can state the following result.

**Proposition 5.** Let the assumptions of this section hold true and suppose $\rho \geq 0$. Then, $V : S \to \mathbb{R}$ is a solution for the functional equation (8.6) if and only if it solves

$$V(s_t) = \max_{\mathcal{M}(s_t)} \left\{ \mathcal{K}(\rho_t) U(C_t^y, L_t, C_{t+1}^o) + \phi V(s_{t+1}) \right\}$$

where $\mathcal{K}(\rho_t) \equiv \int_{\mathcal{X}} \omega_t^{1-\sigma} \gamma^\gamma \, d\xi$, recalling $\omega_t \equiv e^{1-\rho_t} \left( \int_{\mathcal{X}} e^{1-\rho_t} \, d\xi \right)^{-1}$.

**Proof.** The equivalence is an immediate consequence of Lemmas 3 and 4 and the above discussion. Notice $\mathcal{K}(\rho_t) U(C_t^y, L_t, C_{t+1}^o) = \int_{\mathcal{X}} U(\omega_t C_t^y, L_t, \omega_t C_{t+1}^o) \, d\xi$. \(\square\)

The functional equation is now similar to a standard recursive social planner problem. The main difference is that here the consumption of the old enters the problem as an additional state variable. The underlying intuition is that for an allocation to be implementable as a competitive equilibrium, for given labour supply and young consumption, a certain amount of old consumption must be promised to the agents of a certain generation. However, when the next period comes, old consumption is an overhead cost on production and the Ramsey planner has an incentive to break that promise. Thus, it is necessary to add a state variable to keep memory of the promise made to old agents.

Notice that the problem is not convex, hence I need to use value function iteration to identify the solution to (9.7). Furthermore, I remark that the state variable $P_t$ can be eliminated when debt limits are either ignored by assumption or fixed to zero.\(^{24}\)

10. **Calibration and Numerical Results**

This section presents a set of numerical simulations for the Italian economy. First, section 7 provides a framework for a steady state analysis. Then, the aggregation result from section 9 allows to study the Ramsey optimal dynamics of the economy for given initial conditions. Throughout, I refer to a Ramsey optimal government policy simply as optimal policy. In the appendix I give a detailed description of the variables construction and of the Matlab code.

The numerical findings are very close to the results in Conesa et al. (2009). They enlighten the fundamental role played by government debt limits in determining positive optimal capital taxation. Furthermore, optimal labour taxation is almost flat. The results are robust to two different utility specifications, namely (7.1) and (9.1).

**Data.** The main dataset is EU-SILC 2015, the European Union reference source for comparative statics on income distribution. I focus on cross sectional Italian data on gross and net income subject

\(^{24}\)As it is the case in the main specification of Conesa et al. (2009).
to IRPEF taxation\textsuperscript{25}. The sample is composed of 18,716 observations. Besides, I employed 2019 ISTAT and 2015 OECD data to calibrate probability of survival, public expenditure and debt-to GDP ratio.

**Calibration.** Table 1 summarizes the calibrated parameters. I assume the government shares the same discount factor of the agents. The results are robust to different specifications. Time discounting is set to 0.5 as a consequence of the fact that I am considering a two-period overlapping generation economy, that is, one period approximately corresponds to 40 years. The production function is Cobb-Douglas, where the capital share is assumed to be the standard 0.36. Survival probability is calibrated to 0.915 using ISTAT data and capital depreciation is set to 0.88, consistently with an annual depreciation of approximately 0.05. Public expenditure is calibrated to match OECD data on government production costs as percentage of Italian GDP. Furthermore, the benchmark debt to GDP ratio is set to 1.35, as reported by ISTAT data.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Benchmark Calibration</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.50</td>
<td>Time discount factor</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.50</td>
<td>Planner’s time discount factor</td>
</tr>
<tr>
<td>$\psi$</td>
<td>0.915</td>
<td>Survival probability</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.36</td>
<td>Capital share</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.88</td>
<td>Depreciation</td>
</tr>
<tr>
<td>$\frac{g}{Y}$</td>
<td>0.20</td>
<td>Public expenditure to GDP</td>
</tr>
<tr>
<td>$\frac{b}{Y}$</td>
<td>1.35</td>
<td>Government debt to GDP</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.21</td>
<td>Labour tax progressivity</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.26</td>
<td>Capital tax</td>
</tr>
</tbody>
</table>

**Table 1. Benchmark Economy**

I use income data from EU-SILC to calibrate the labour tax function benchmark parameters. My strategy builds on Heathcote et al. (2017), who estimated the progressivity parameter to be 0.18 for the U.S. economy. I consider agents who are between 25 and 65 years old. Taxable income is constructed subtracting social contributions to gross income. Then I construct disposable income adding government transfers to net income. Finally I regress the logarithm of disposable income on the logarithm of taxable income. The regression fits well the data with $R^2 = 0.89$ and gives $\rho = 0.20$.\textsuperscript{26} This finding suggests that labour taxation is more progressive in Italy than in the United States. Capital taxation is set to 0.26 accordingly to the Italian flat tax on dividend income. This is

\textsuperscript{25}IRPEF (Imposta sul Reddito delle PErsona Fisiche) is a tax on personal income. The main component of IRPEF taxable income is labour income. However, some types of capital income are also subject to IRPEF.

\textsuperscript{26}Heathcote et al. (2017) regression has $R^2 = 0.92$. 

a simplifying approximation since, for example, part of capital income is subject to IRPEF taxation together with labour income. On the one hand, I remark that some degree of approximation seems to be unavoidable. In particular, Italian legislation does not always separate capital from labour income, as it is the case for IRPEF. The part of capital income subjected to IRPEF is treated exactly as if it were labour compensation, and it is common for one agent to have her taxable income composed of both labour and capital remuneration. In these cases, due to IRPEF progressivity, it is not even possible, up to make some discretionary assumption, to determine the tax rate the capital income is subject to. Second, I stress that there is no consequence for the optimal results except the fact that the policy maker should be aware that the model assumes that there is only one form of (linear) capital taxation. That is, the policy implications of the model would require the Italian policy maker to suitably simplify the current capital taxation regime.

Finally, I use gross labour income data to construct the idiosyncratic efficiency function $\varepsilon$. I express the efficiency of an agent as the fraction of his gross labour income over the average gross labour income. Observe that the function $\varepsilon$ has unit mean by construction. The efficiency distribution resembles a lognormal. The median is approximately 0.88. The Gini coefficient is 0.39 and the the top one percentile has an efficiency of 4.15. The estimated distribution is plotted in Figure A1 in the appendix.

**Steady State Analysis.** The steady state analysis is conducted within the framework presented in section 7. I calibrate the weight parameter $\gamma$ in (7.1) to 0.31 such that in the benchmark economy labour supply is 0.33. Without government debt budget constraint, the steady state optimal policy define a negative capital tax $\tau^* = -0.18$ which, as stated by Proposition 4, perfectly compensates the agents for the mortality risk they are subject to. The optimal labour tax is defined by a linear term $\lambda^* = 1.13$ and the optimal progressivity is $\rho^* = -0.12$. That is, optimal labour taxation is slightly regressive. Optimal labour supply is 0.42. The implied level of public debt necessary to satisfy the steady state government budget constraint is extremely negative, the debt-to-GDP ratio resulting to be $-1.32$. Table 2 summarizes the main results.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Optimal Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varphi$</td>
<td>0.13</td>
<td>Public expenditure-to-GDP</td>
</tr>
<tr>
<td>$h$</td>
<td>-1.32</td>
<td>Debt-to-GDP</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>1.13</td>
<td>Labour tax, linear component</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.12</td>
<td>Labour tax, progressivity</td>
</tr>
<tr>
<td>$\tau$</td>
<td>-0.18</td>
<td>Capital tax</td>
</tr>
</tbody>
</table>

**Table 2. Unconstrained Ramsey Optimal Steady State**
When debt limits are considered, the optimal policy is substantially affected. While optimal labour taxation tends to become flat, limiting government borrowing leads to highly positive capital taxation. Assuming the government debt to be constrained to zero, the implied optimal capital tax is 0.37. This result is particularly interesting when compared to the results of Conesa et al. (2009), who have a more realistic but rather complex economic environment.\(^{27}\) Indeed, in their main model, where government debt is assumed to be zero, they numerically find an optimal capital taxation of 0.36. Furthermore, I found that my framework is able to closely replicate the results of Conesa et al. (2009) for optimal capital taxation for different levels of government debt and that optimal capital taxation is increasing in the debt-to-GDP ratio, as reported in Table 3.

\[
\begin{array}{cccc}
\text{Debt-to-GDP Ratio} & \text{Optimal Capital Tax} & \text{Optimal Progressivity} \\
\hline
0.00 & 0.37 & 0.36 & -0.09 \\
0.20 & 0.43 & 0.43 & -0.08 \\
0.60 & 0.56 & \sim & -0.07 \\
1.00 & 0.67 & 0.55 & -0.05 \\
\end{array}
\]

**Table 3. Ramsey Optimal Steady State with Debt Limits**

As an illustrative example, for a debt-to-GDP ratio of 0.6, the optimal capital taxation result to be approximately 0.56.\(^{28}\) The intuition for this finding may be found in Proposition 3. That is, if debt cannot be used to separate old consumption from capital remuneration, then the government is forced to introduce distortionary taxes in agents’ saving decisions. Observe that in the simulations the binding constraint is always the lower bound. This is intuitive since as I discussed above, the unconstrained steady state optimal government policy requires negative debt to be sustainable. Hence, the numerical analysis suggests that government debt limits are one of the main rationales for optimal capital taxation to be positive in the long run. Table 4 shows that optimal progressivity approaches zero as the debt-to-GDP ratio increases.

One may reasonably argue that when considering the dynamic problem, the government may find optimal to never converge to such a steady state, even in the absence of government debt constrains. Indeed, the benchmark economy has an extremely high debt-to-GDP ratio and reaching the steady state optimal level of debt may result excessively costly. To assess this important concern, I propose a numerical exercise to study the dynamics of the economy for arbitrary initial conditions.

\(^{27}\)They consider a multi period overlapping generations model with stochastic life cycle.

\(^{28}\)A debt-to-GDP ratio of 0.6 is part of the convergence criteria of the Treaty on European Union (1992), also know as the Maastricht Treaty.
Dynamic Analysis. To study the dynamics of the economy I rely on the set up discussed in section 9. For computational feasibility reasons, I approximated the distribution of idiosyncratic efficiency with a lognormal with mean one. This way, I have a closed form expression for the term $\delta_0(t)$. Furthermore, in order to satisfy the assumptions of Lemma 4 and thus ensuring the existence of a representative agent for the economy I assume a zero lower bound for the progressivity parameter. As I already observed, this seems a reasonable assumption for Italy, since the constitution decrees the tax system to be progressive. Moreover, in order to have only two state variables for the recursive problem, I focus on the case where there are no debt limits. I calibrate $\sigma = 1.5$ and $\gamma = 0.31$.

For any initial conditions, the Ramsey optimal path for the economy converges to a unique steady state. Recall that each period amounts to approximately 40 years. Optimal progressivity is very close to zero at every period. When the government faces no restriction on public debt, it uses high capital taxation in the first periods to reach a steady state with asset accumulation. This finding reassures from the concern raised above, suggesting that for any initial conditions, the Ramsey plan actually converges to a steady state where savings are subsidized and the govern controls a relevant part of the assets in the economy. Figure 1 describes the dynamics of the main variables until convergence in the case debt constraints are not binding: recall that each time period on the x-axis corresponds to roughly 40 years.

![Figure 1. Ramsey Optimal Dynamics](image)

The scenario that most closely approximate the Italian economy is represented by the non-dotted red line, which specifies an initial debt-to-GDP ratio of approximately 1.35. Capital taxation is positive.
for the first three generations. The capital tax is particularly high in the second period. This way, the government reduces the consumption of the old (and of the young, since the unintended bequests are proportional to old consumption) and accumulates assets to reach the Ramsey optimal steady state. There, the government controls a relevant portion of the economy and subsidizes capital. The dynamics of additional variables, such as the consumption of the young, is reported in Figure A2 in the appendix.

The policy implications of these results suggest that Italian policy makers should raise capital taxation, e.g. on dividend income and housing and drastically reduce labour taxation progressivity. Then, once enough assets are accumulated, the government should subsidize capital in the long run: approximately 120 years from now.

However, if the government aims to satisfy some benchmark steady state nonnegative debt-to-GDP ratio, such as the 0.6 ratio prescribed by the convergence criteria of Maastricht treaty, then optimal capital taxation is consistently positive even in the long run.

11. Conclusion

My work presents a two-period overlapping generations model, that allows agents to be heterogenous in their productive efficiency and assumes they are subject to some mortality risk, with no annuity market. This concluding section briefly summarizes the main results.

I show that the functional form proposed by Feldstein (1969) and Heathcote et al. (2017) is a significant generalization for standard Ramsey problems, allowing the government to impose non linear taxes on labour income while maintaining the so called primal approach, that is, handling directly with allocations rather than taxes, significantly simplifying the framework.\textsuperscript{29}

I prove that in an homogeneous agent economy, if the government can freely choose progressivity of labour taxation and does not face binding debt limits, then it is able to achieve the first best in steady state. Under this assumption, capital tax is non positive and perfectly compensates the agents for the mortality risk they are subject to.

I argue that government debt limits are a fundamental rationale for positive capital taxation in the long run. The presented framework, using Italy as a benchmark economy, is able to closely replicate the optimal taxation numerical results of Conesa et al. (2009). Furthermore, I develop a numerical study of the dynamics of the economy showing that, for any initial conditions, there is always convergence to a unique steady state, where, if the government does not face tight debt limits, there is asset accumulation.

\textsuperscript{29}See Chari and Kehoe (1999)
References


Appendix: Proofs of Lemmas 3 and 4

This appendix presents the proofs of lemmas 3 and 4, stated in section 9 and used in the proof of Proposition 5.

Lemma 3.

Proof. I need to show that (8.2), (9.2) hold. Given (9.3)-(9.4), the first constraint implied by condition (8.2) can be satisfied defining

\[ P_{t+1} = \beta \psi c_{t+1} \]

Notice I am relying on the law of \( u \) to get rid of \( \omega_t \). For what concerns the second one, notice that using (9.3)-(9.4), the assumed law for \( u \), and the definition of \( \omega_t \) to elide the function cone realizes it is satisfied for

\[ \Lambda_t = - \frac{u_t (C_t^\psi, L_t) L_t}{u_c (C_t^\psi, L_t) (1 - \rho_t) (F_2 (K_t, L_t) L_t)} \left( \int \psi e^{1 - \rho_t} d_i \right)^{-1} \]

Finally, I need to check (9.2) is satisfied. Conditions (9.3)-(9.4) let me rewrite it as

\[ \omega_t u_c (\omega_t C_t^\psi, L_t) (C_t^\psi - (1 - \psi) C_t^\psi) + \omega_t \beta \psi^2 u_c (\omega_t C_t^\psi) C_t^{\psi+1} + \frac{u_t (\omega_t C_t^\psi, L_t) L_t}{(1 - \rho_t)} = 0 \text{ a.e.} \]

and given the law of \( u \) I can elide the weight function \( \omega_t \) and conclude that the above condition is equivalent to (9.5), which holds by assumption. \( \square \)

Lemma 4.

Proof. It has to be proved that (9.3)-(9.5) hold. I use (8.2) and the law of \( u \) to express \( C_{t+1}^\psi \) and \( C_t^\psi \) as functions of \( l_t \). Then I use these findings and (8.2) in the implementability constraint (9.2) to find another expression for \( C_t^\psi \) as a function of \( l_t \). Equating the two expressions for \( C_t^\psi \), and dividing both sides by \( e^{1 - \rho_t} \) one gets

\[ C_{t+1}^\psi = \eta_t C_t^\psi \text{ a.e.} \]

\[ C_t^\psi = \Lambda_t \frac{\gamma}{1 - \gamma} \left( \frac{1 - l_t}{l_t} \right) (1 - \rho_t) (F_2 (K_t, L_t) L_t) e^{1 - \rho_t} \text{ a.e.} \]

where \( \eta_t \equiv \left[ \frac{P_{t+1}}{\beta \psi} (1 - l_t)^{(1 - \gamma)(1 - \sigma)} \right]^{(1 - \rho_t) / (1 - \rho_t)} \). Furthermore using (8.2) in (9.2) and rearranging I obtain

\[ C_t^\psi = (1 + P_{t+1} \eta_t)^{-1} \left[ \Lambda_t (F_2 (K_t, L_t) L_t) e^{1 - \rho_t} + \omega_t \hat{T}_t \right] \text{ a.e.} \]

for \( \hat{T}_t \equiv (1 - \psi) (C_t^\psi + P_{t+1} C_{t+1}^\psi) \). Thus, it must be, defining \( \Omega_t \equiv \int \psi e^{1 - \rho_t} d_i \)

\[ \Lambda_t (F_2 (K_t, L_t) L_t)^{1 - \rho_t} + \hat{T}_t \Omega_t^{-1} = \Lambda_t \frac{\gamma}{1 - \gamma} (1 - \rho_t) (F_2 (K_t, L_t))^{1 - \rho_t} (1 + P_{t+1} \eta_t) \frac{1 - l_t}{l_t} \text{ a.e.} \]
Notice I divided both sides by $e^{1-m}$. For any set of aggregate values, $l_t$ is the only function free to move in order to satisfy (11.6). However, there is only one value $l$ which can solve the equation. Indeed the left hand side is increasing in $l_t$. Speculatively, under the assumption $\rho \geq 0$, recalling it must be $\rho_t \geq \rho$, the left hand side is decreasing in $l_t$. Notice that for the result we used the fact that $\sigma > 1$, which assures $\eta_t$ is decreasing in $l_t$. Thus $l_t$ must be constant $i.e.$ and by construction it then results $l_t = L_t$ $i.e.$ almost everywhere, which is (9.3). Then (9.4) follows from (11.3)-(11.4) and (9.5) is derived similarly as I did in the proof of Lemma 3.

Appendix: Dataset and Variables

I use the dataset EU-SILC 2015 (36,602 observations) to construct the idiosyncratic efficiency function $e$ and estimate the progressivity of the benchmark labour tax. EU-SILC is the EU reference source for comparative statistics on income distribution. It provides personal and household-level annual data.

To construct the function $e$, I use gross labour income data. I compute the gross labour income of an agent as the sum of his gross employed (PY010G) and self employed (PY050G) labour income. Then, I define the idiosyncratic efficiency as the ratio between the individual gross labour income and the average. I obtain the density distribution using the Matlab function $\text{ksdensity}$.

To estimate the benchmark progressivity, I proceed as it follows. I compute the gross income of each individual as the sum of all the types of income subject to IRPEF. Namely, among the gross personal-level data I consider: employee (PY010G), company car (PY021G) and self employed (PY050G) labour income; unemployment benefits (PY090G); old age (PY100G) and survivors benefits (PY110G). Then, I add household-level gross income data to the income of the householder (whose personal ID’s last figure is 1). To be specific, they are income from rental of a property (HY040G), regular inter-household cash transfers (HY080G) and profits from incorporated business (HY090G). Considering these forms of capital income is necessary to correctly address the IRPEF progressivity, which treats them as if they were labour income. Then, I obtain the taxable income subtracting social contributions (CSdi, CSda) to gross income. I obtain the disposable income summing the net counterparts of the mentioned variables, (PY010N, PY021N, PY050N, PY090N, PY100N, PY110N and HY040N,HY080N,HY090N) and adding the social transfers. Personal-level data on transfers are disability benefits (PY130N), education-related allowances (PY140N). Household-level data on transfers are children-related allowances (HY050N), social exclusion benefits (HY060N) and housing allowances (HY070N). All the household-level transfers are added to the disposable income of the householder, with the exception of children-related allowances. In fact, to approximate the Italian legislation, when it is possible I split this type of transfer between the first and second individual associated to the household.
Appendix: Matlab

To characterize the steady state Ramsey optimal allocation I use the Matlab function `fmincon` to solve the system of equations given by the first order conditions identified in section 7, adding the lagrangian multipliers associated to (5.5) and (2.6) to consider the possibility of binding policy constraints. Due to the non-regularity of the problem, it is very important to make the correct guess to get the solution, paying special attention to the initial values of the Lagrangian multipliers. In particular, when looking for a solution where some constraint is not binding, it is necessary to use `fmincon` specifications to impose the associated Lagrangian multiplier is zero.

To address the dynamic problem I needed to solve the functional equation (9.7). To this end, I proceed as it follows.

(1) First, I solve the intratemporal maximization problem. That is, I find the optimal choice for the vector of control variables \((C^n_t, L_t, \rho_t)\) taking as given the pair \((s_t, s_{t+1})\) of vectors of state variables.

(a) I construct grids for capital and consumption of the old. Moreover, I define a grid of possible values for \(f_{lt}\).

(b) I use the feasibility constraint 5.1 to express \(C^n_t\) as a function of \(L_t, \rho_t\). Then, taking \(\rho_t\) as given, obtain \(L_t\) using the Matlab function `fsolve` to have (9.6) satisfied.

(c) Then, for each pair \((s_t, s_{t+1})\) I compute \(\frac{\partial}{\partial f_{lt}} U(C^n_t, L_t, C^{o}_{t+1})\) for all the possible \(\rho_t\) in the grid and take the \(\rho_t\) which maximizes it.

(d) I select the corresponding optimal value of \(L_t\) for each \((s_t, s_{t+1})\)

(2) Second, I use the Matlab function `griddedinterpolant` to construct \(L_t\) as a function of \((s_t, s_{t+1})\).

Chosen \(L_t\), I recover \(C^n_t\) using the feasibility constraint and \(\rho_t\) from (9.6).

(3) I solve for the value function \(V\) using the method of value function iteration.

(a) I take a matrix of zeros as initial guess for \(V\) as a function of \(s_t = (C^n_t, K_t)\). 

(b) I compute the value of \(\mathbb{E}(\rho_t) U(C^n_t, L_t, C^{o}_{t+1}) + \phi V(s_{t+1})\) for all the pair \((s_t, s_{t+1})\).

(c) I update \(V(s_t)\) finding the maximum value of \(\mathbb{E}(\rho_t) U(C^n_t, L_t, C^{o}_{t+1}) + \phi V(s_{t+1})\) for each \(s_t\) in the grid.

(d) I iterate until convergence.

(4) I find the policy function using the obtained \(V\).

(a) For each \(s_t\), I find the corresponding \(s_{t+1}\) which maximizes \(\mathbb{E}(\rho_t) U(C^n_t, L_t, C^{o}_{t+1}) + \phi V(s_{t+1})\).

(b) I recover the corresponding \(L_t, C^n_t\), and \(\rho_t\).

Notice that the code handles with four dimensional objects. As a consequence, to obtain an acceptable degree of approximation high computational power is needed. Furthermore, when solving for \(L_t\) in point (1), the `fsolve` part must be handled with care, controlling for the results to be actual real roots for (9.6).
Appendix: Additional Figures

Figure A1: Cross Sectional Efficiency Distribution

Figure A2: Additional Ramsey Optimal Dynamics
SUMMARY

Optimal taxation has been hotly debated by economists. A large number of relevant contributions have been written to answer the fundamental question of how fiscal policy should be set over the business cycle and the long run. A classical framework for the problem may be found in the work of Ramsey (1927), where the government has to raise an exogenously given revenue by the means of distortionary linear taxes, with the aim of minimizing the utility loss consequently experienced by the economic agents. The Ramsey problem has then been extended to the dynamic maximization of a given welfare function using distortionary taxes to induce a particular competitive equilibrium. That is, the economic literature has tried to identify the fiscal policy characterizing the second best achievable by a government when non distortionary lump transfers are not feasible. In particular, should capital be taxed in the long run? Which is the optimal progressivity level for labour taxation?

The present work contributes to the literature debate on optimal taxation. As noticed by Peterman (2013), optimal capital taxation models are fairly complicated and it is important to disentangle the implications of each of the assumptions in determining the results. Following Piguillem and Shakhnov (2020), I show that government debt plays a fundamental role in determining capital optimal taxation. In particular, I highlight that government debt limits are among the main rationales for positive capital taxation to be optimal in the long run for an overlapping generations economy. Furthermore, I characterize the full optimal dynamics of the economy for some standard preference specifications, a result which is not common in the literature. To this end, I study a simple two-period overlapping generation environment, where agents may be heterogeneous in their labour productivity and are subject to a mortality risk, with no annuity market. In this environment, the government is allowed to impose non linear taxes on labour income, according to the functional form proposed by Feldstein (1969) and Heathcote et al. (2017). Interestingly, Heathcote and Tsujiyama (2019) argue that this specification is able to well approximate the optimal solution for the case in which no functional restriction is imposed on the tax schedule. Following Conesa et al. (2009), capital taxation is linear.

First, I show that even if the assumption of a linear labour tax is relaxed, it is still possible to frame the Ramsey problem preserving the so-called primal approach.¹ This consists in characterizing the allocations implementable as competitive equilibria by some government fiscal policy only using agents’ budget constraints and first order optimality conditions. This approach has the advantage to handle directly with allocations rather than with the set of taxes which induce them, that means making the problem more intuitive.

Second, I prove that in an homogenous agent economy, if the government is free to choose the labour tax progressivity and does not face binding debt limits, then it is able to achieve the first best in steady state.\footnote{The first best is defined as the allocation which would be implemented by a utilitarian social planner, who does not necessarily share the same time discount factor of the agents.} Furthermore, under this assumptions the capital tax is in general negative in steady state, perfectly compensating the agents for the mortality risk they are subject to.

Third, I provide numerical exercises, taking Italy as benchmark economy, to argue that constraints on the government debt accumulation are fundamental rationales for largely positive capital taxation in the long run. I show that the simple economic environment I adopt is able to closely replicate the results of Conesa et al. (2009).\footnote{Assuming the government has to keep debt equal to zero, as they did for their main model specification, I get an optimal steady state capital tax of 0.37, where they get 0.36.} In particular, for debt-to-GDP ratios between 0 and 0.6, the optimal tax rate on capital ranges from 0.37 to 0.56.

Fourth, building on Piguillem and Shakhnov (2020), I develop a recursive formulation of the problem which allows for aggregation to numerically study the dynamics of the economy. I show that the economy always converges to a unique steady state. The implied optimal labour tax is almost linear.

**Literature.** Regarding capital taxation, Chamley (1986) has proved that, for a wide class of infinitely lived agent models, it should be zero in steady state, a position supported also by Judd (1985) and Lucas (1990). Furthermore, Jones et al. (1997) and Charli and Kehoe (1999) proved the result of Chamley (1986) is robust to the relaxation of a set of assumptions. However, Straub and Werning (2020) argue that convergence to such a steady state may be suboptimal for some preference specifications. Furthermore, the literature has identified at least two rationales for capital taxation to be positive in the long run. Aiyagari (1995) has shown that for the Bewley (1986) class of incomplete market models capital taxation is always positive; interestingly, the result holds not only in steady state, but also for the whole business cycle. Second, Erosa and Gervais (2000) and Garriga (2017) argue that in life cycle models where taxes cannot depend of agents’ age, capital tax is in general different from zero. Conesa et al. (2009) quantitatively address the issue, finding that representing the U.S. economy in a multi-period overlapping generations model with stochastic life cycle, mortality risk and incomplete markets, optimal capital taxation is significantly positive in steady state.\footnote{They get an optimal capital tax of 0.36 in their main model specification.}

Conesa et al. (2009) also find that the optimal tax on labour is essentially flat. In fact, the classical framework inspired by Ramsey (1927) in general restricts taxes to be linear, but progressivity in labour income taxation has emerged as a central topic in the literature. Several authors, as Benabou (2002) and Heathcote et al. (2017) have claimed for the optimality of a progressive labour tax. Other contributions, like Altig et al. (2001), support flat labour income taxation.
My work presents a two-period overlapping generations model, that allows agents to be heterogenous in their productive efficiency and assumes they are subject to some mortality risk, with no annuity market. This concluding section briefly summarizes the main results.

I show that the functional form proposed by Feldstein (1969) and Heathcote et al. (2017) is a significant generalization for standard Ramsey problems, allowing the government to impose non linear taxes on labour income while maintaining the so called *primal approach*, that is, handling directly with allocations rather than taxes, significantly simplifying the framework.\(^5\)

I prove that in an homogeneous agent economy, *if* the government can freely choose progressivity of labour taxation and does not face binding debt limits, *then* it is able to achieve the first best in steady state. Under this assumption, capital tax is non positive and perfectly compensates the agents for the mortality risk they are subject to.

I argue that government debt limits are a fundamental rationale for positive capital taxation in the long run. The presented framework, using Italy as a benchmark economy, is able to closely replicate the optimal taxation numerical results of Conesa et al. (2009). Furthermore, I develop a numerical study of the dynamics of the economy showing that, for any initial conditions, there is always convergence to a unique steady state, where, if the government does not face tight debt limits, there is asset accumulation.