



**HAL**  
open science

## Scale-dependent and risky returns to savings: Consequences for optimal capital taxation

Eddy Zanoutene

► **To cite this version:**

Eddy Zanoutene. Scale-dependent and risky returns to savings: Consequences for optimal capital taxation. *Journal of Public Economic Theory*, 2023, 25 (3), pp.532-569. 10.1111/jpet.12633. hal-03891225

**HAL Id: hal-03891225**

**<https://amu.hal.science/hal-03891225>**

Submitted on 9 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Scale-dependent and risky returns to savings: Consequences for optimal capital taxation

Eddy Zanutene<sup>1,2</sup> 

<sup>1</sup>Aix-Marseille School of Economics (AMSE), CNRS, EHESS, Centrale Marseille, Aix-Marseille University, Marseille, France

<sup>2</sup>Centre de recherches en économie et droit (CRED) (TEPP), Université Paris-Panthéon-Assas, Paris, France

## Correspondence

Eddy Zanutene, Aix-Marseille School of Economics (AMSE), CNRS, EHESS, Centrale Marseille, Aix-Marseille University, Marseille, France.

Email: [eddy.zanutene@u-paris2.fr](mailto:eddy.zanutene@u-paris2.fr)

## Abstract

I present a model of optimal capital taxation where agents with heterogeneous labor productivity randomly draw their rate of return to savings. Because of scale dependence, the distribution of rates of returns can depend on the amount saved. Uncertainty in returns to savings yields an insurance rationale for taxing capital on top of labor income. I first show that, because of scale dependence, agents making the same saving decision should access the same rate of return at the optimum. I then constrain the information set of the government and show that, as soon as return are uncertain, positive capital income taxation is needed at the optimum. The optimal linear tax on capital income trades off insurance with distortions to both savings and to the rate of return in a context of scale dependence. Eventually, I argue that scale dependence in and of itself is not sufficient to justify capital taxation on top of labor income taxes. These results are still valid when agents can optimize between a risk-free and a risky-asset that can both exhibit scale dependence.

## 1 | INTRODUCTION

In developed economies, capital taxes are part of the fiscal tools used by the government to finance its expenditures. For instance, according to a calculation based on statutory rates at both the corporate and the shareholder level provided by Harding (2013), dividends were on average taxed at 41.8% in 2012 in the OECD. Yet delivering a clear theoretical justification for taxing capital is challenging as seminal contributions by Atkinson and Stiglitz (1976), Judd (1985) and Chamley (1986) can provide a rationale

against capital taxation.<sup>1</sup> Among the assumptions required to establish these zero-capital tax results, the homogeneity and certainty of the rate of return has been recently challenged on both theoretical and empirical grounds. First, in an attempt to replicate the dynamics of wealth concentration documented for instance in Saez and Zucman (2016), a wave of papers in quantitative macroeconomics insists on the volatility of returns to savings. Indeed, life-cycle models can match the empirical properties of wealth distributions only by assuming stochastic returns, violating the homogeneity assumption made in standard optimal tax models.<sup>2</sup> Second, recent contributions in household finance by Fagereng et al. (2020) and Bach et al. (2020) present empirical evidence of an important dispersion in returns to savings.<sup>3</sup> On top of this volatility, rates of return are also likely to exhibit a positive correlation with wealth, what is often described in the literature as *scale dependence*. From a theoretical point of view, scale dependence has been introduced by Gabaix et al. (2016) to replicate the fast transition in observed wealth dynamics. Besides, both Fagereng et al. (2020) and Bach et al. (2020) show that the empirical distribution of returns does positively correlate with wealth.

On these grounds, I develop a two-period model with agents making work and saving decisions in a context where the rate of return on savings can be risky and scale-dependent. Following Mirrlees (1971), agents are ex ante heterogeneous with respect to their labor productivity. Given this labor productivity heterogeneity, they make their labor supply and saving decisions. However, departing from traditional optimal tax models, I suppose that the rate of return they get from their savings is randomly drawn from a wealth-dependent distribution. This modeling strategy captures both the changes in expected return and in riskiness caused by changes in the saving decision in a context where returns can exhibit scale dependence. Agents know that the amount they save determines the distribution in which the rate of return is drawn. Assuming separable preferences between consumption and work effort as in Atkinson and Stiglitz (1976), I discuss the properties of optimal capital taxation when nonlinear labor income taxation is available.<sup>4</sup>

The main results of the paper are the following. First, when the government can freely tax any component of capital, be it capital income or savings, the only role for capital taxation is to perfectly insure agents against the volatility of the rate of return. Without scale dependence, this result implies that the government combines confiscatory capital income taxes with a wealth subsidy to grant access to the (unconditionnal) average rate of return to every taxpayers. However, as soon as returns display scale dependence, the optimal policy has to take into account that agents making different saving decision do not expect the same rate of return. Hence in this context, capital taxes are used to grant access to the average rate of return, conditional on the saving decision made by taxpayers. Eventually, when returns are deterministic, there is no tax on capital at the optimum, even in presence of scale dependence.

---

<sup>1</sup>The Chamley-Judd result rules out capital taxation on the long run. Interpreting capital taxation as taxes on future consumption, the Atkinson-Stiglitz theorem can also be used to assert the irrelevance of capital taxation.

<sup>2</sup>See for instance Benhabib et al. (2011), Gabaix et al. (2016), or Benhabib and Bisin (2018).

<sup>3</sup>For instance, Fagereng et al. (2020) documents a 22% standard deviation in returns to net wealth, using Norwegian data.

<sup>4</sup>Applying Atkinson and Stiglitz (1976) to a two-period economy, capital taxation is zero at the optimum in presence of an optimal labor income tax schedule. Laroque (2005) and Kaplow (2006) extend Atkinson-Stiglitz theorem to nonoptimal labor income tax schedule. A review of this standard approach to capital taxation can be found in chapter 9 of Kaplow (2010).

Second, I study the optimal capital tax policy in contexts where the government does not observe all components of capital. For instance, when the government is constrained to tax only savings and labor income, the logic of Atkinson and Stiglitz (1976) applies and there is no capital taxation at the optimum, even in presence of scale dependence. However, when the constraint is to only rely on labor and capital income taxation, there exists an insurance rationale for taxing capital income at the optimum when returns are risky. In this context, I derive an optimal linear tax formula that highlights the trade-off between providing insurance, distorting savings, and distorting the rate of return in a context of scale dependence.

Third, I show that the logic behind these results extends to a setting where agents can invest in a risk-free and a risky asset, both potentially exhibiting scale dependence. In this two-asset framework, I show that taxing the risk-free asset is irrelevant in every configuration of the model, even when there exists scale dependence on this safe asset. However, taxing the capital income arising from investment in a risky asset is always Pareto-improving. In particular, when the government can perfectly distinguish returns of the risk-free asset from those attached to the risky one, capital taxation is used to provide to every agents the average rate of return on the risky asset, conditional on the amount invested in this specific asset. All these results are proven under a general social welfare function aggregating individual expected utilities.

## 1.1 | Related literature

This paper first relates to previous works analyzing the consequences of uncertainty on capital taxation. A seminal contribution has been provided by Domar and Musgrave (1944), showing that agents completely offset the impact a tax on capital income has on expected utility by adjusting the riskiness of their portfolio. Extended and balanced by Stiglitz (1975), Bulow and Summers (1984), or Gordon (1985), this paper has drawn major attention to the effect of capital taxation on risk-taking and portfolio choices. An optimal tax approach with endogenous labor supply has been recently provided by Boadway and Spiritus (2021). I especially relate to this strand of the literature in the extension of the model where a simple portfolio choice decision is introduced: agents can choose between a risk-free and a risky asset, in the tradition of Domar and Musgrave (1944). My main contribution to this literature is to include scale dependence in a two-assets environment. To the best of my knowledge, this paper is the first to explicitly consider scale dependence on the risk-free asset, as empirically documented by Fagereng et al. (2020). Regarding risky assets, introducing scale dependence allows me to derive two new results on the way to tax capital in uncertain environment. First, I discuss how the full insurance property of capital taxation in presence of idiosyncratic risk described in Gordon (1985) and recently extended by Boadway and Spiritus (2021) evolves when scale effects are taken into account. In particular, I show that because of scale dependence, full insurance does not imply that agents access the same rate of return at the optimum. With scale dependence, capital taxation should be used to provide to every taxpayer the average rate of return, conditional on the amount they have invested in the risky asset. Second, I provide an optimal (linear) tax formula on (risky) capital income in presence of both uncertainty and scale dependence. As in Boadway and Spiritus (2021), this tax rate is positive at the optimum because of the (partial) insurance it provides. Yet, coherent with the logic of Ramsey (1927), the tax rate decreases when the tax base reacts strongly to tax reforms. Hence the tax rate would be lower if investment in the risky asset responds strongly to capital income taxation. This effect would be amplified if, because of scale effects, the rate of return also declines when investment is reduced.

Departing from the portfolio choice analysis, the consequences of uncertainty for capital taxation can also be analyzed by simply adding a random component to future income. Although less realistic, this route allows capturing the key trade-off between providing social insurance and reducing incentives. In this sense, the majority of my paper relates to the contribution of Varian (1980) who studies a two-period economy where the dispersion in second-period observed income only comes from exogenous differences in luck. In this context, Varian (1980) shows that if the government can distinguish between the random and the deterministic part of income, perfect insurance should be provided at the optimum. Besides, even when this distinction is no longer feasible, positive taxes on income are still part of the optimal policy. I extend these results in two ways. First, I consider a Mirrleesian economy with ex ante heterogeneous agents who can adapt to taxes on two margins: savings and labor supply. In Varian (1980), agents are ex ante homogeneous and efficiency losses can only occur through incentives to save. Second, modeling scale dependence in the rate of return allows me to discuss how the optimal tax policy is affected when the random component can actually depend on an endogenous decision, namely the saving decision an agent make. This adds further distortions to savings that were not taken into account in Varian (1980) and can mitigate the insurance rationale for taxing capital.

Including scale dependence in my analysis allows me to contribute to the optimal tax literature with return heterogeneity in deterministic environments. Studying two-type economies, Kristjánsson (2016) and Gahvari and Micheletto (2016) solve the optimal capital tax problem when returns can exhibit scale dependence. My framework nests these models as I study both scale dependence and risk in a continuous-type Mirrleesian economy. In every optimal tax exercise considered in my paper, I reach the conclusion that the primary role of capital taxation is to provide insurance against risky returns while scale dependence in and of itself does not provide a rationale for taxing capital. This conclusion is coherent with the zero capital tax results of both Kristjánsson (2016) and Gahvari and Micheletto (2016) in presence of scale dependence. However, this is at odds with the optimal tax analysis provided by Gerritsen et al. (2020) where scale dependence can justify positive capital income taxation in a deterministic framework. The main reason for this divergence lies in our assumptions regarding the timing of taxes: while I, along with Kristjánsson (2016) and Gahvari and Micheletto (2016), assume that both labor income and capital taxes are levied at time 2, Gerritsen et al. (2020) assume that labor income is taxed at time 1 while capital income is taxed at time 2. In this case, capital income taxation can be desirable as it alleviates the market failure preventing households to pool their savings to maximize scale effects. My main contribution to this strand of the literature is to go beyond the deterministic environment and see how the optimal tax exercise evolves when returns are scale dependent and uncertain. In particular, I show that while scale dependence does not break zero capital tax results in my framework, it affects the optimal level of insurance against risky returns capital taxation should provide. Deriving an optimal linear capital income tax formula, I for instance show that scale effects, by making the tax base more elastic, reduce the optimal provision of insurance against risky returns and lower the capital income tax rates.

Eventually, I contribute to the debate analyzing the desirability of capital income and wealth taxation. While either form of capital taxation is often ruled out by hypothesis (in Gerritsen et al., 2020; Boadway & Spiritus, 2021), I allow for a relatively flexible combination of wealth and capital income taxation to design the optimal policy. Besides to isolate the respective properties of wealth and capital income taxation, I study their relevance when either one these two tools is the only form of capital taxation available. Performing an exercise similar to Kristjánsson (2016) but with a continuous type economy, I show that the riskiness in returns rather advocate for capital income taxation as wealth taxation is of no use in this Atkinson-Stiglitz framework, even with scale dependence. This rationale for capital income taxation

depends heavily on the assumption of unidimensional ex ante heterogeneity considered in my model. The argument would no longer hold if individuals differ in their initial wealth, as for instance in Cremer et al. (2003) and Piketty and Saez (2013), or in their entrepreneurial productivity, as in Guvenen et al. (2019).

The remainder of the paper is organized as follows: Section 2 sets up the model and the method used for finding the optimal capital tax. Section 3 derives the optimal capital tax schedule when the government can freely tax labor income, wealth, and capital income. Section 4 derives properties of the optimal capital tax schedule when the government has limited information on some components of capital. Section 5 discuss these results in an extended two-asset version of the model. Section 6 concludes.

## 2 | THE MODEL

I consider a population of heterogeneous agents living for two periods. In the first period, agents consume and save out of a homogeneous endowment  $A$ .<sup>5</sup> In the second period, they work, receive labor and capital incomes, pay taxes and consume. Similar to Gahvari and Micheletto (2016) and Kristjánsson (2016), this timing implies that taxes on labor and capital are levied at the same time, namely in Period 2. This matches real-life situations where individuals do receive labor and capital incomes during the same year, both being part of the taxable income of a given year.<sup>6</sup> Contrary to Gahvari and Micheletto (2016) and Kristjánsson (2016), I assume that labor is also earned at time 2 while they consider a model where labor is earned at time 1, although being taxed at time 2. I consider their modeling strategy in Appendix G and show that my conclusions extend to the case where labor is earned at time 1 and taxed at time 2. Gerritsen et al. (2020) follows a different path by assuming that labor is earned and taxed at time 1 while capital income is earned and taxed at time 2. As extensively discussed in Section 3.2, this difference in the timing of taxes does affect the optimal policy in a context of scale dependence.

At the beginning of the first period each individual draws a skill endowment  $\theta \in \Theta \subset \mathbb{R}_+^*$ . Skills are distributed according to a cumulative distributive function (c.d.f)  $G : \theta \mapsto G(\theta)$ . Given the draw of  $\theta$ , individuals choose pretax labor income  $y \in \mathbb{R}_+$  (for short labor income hereafter) and savings  $s \in \mathbb{R}_+$ . I augment this standard two-period version of Mirrlees (1971) with two capital market failures, namely uninsurable risk and scale dependence in the rate of return to savings.

*Risky returns.* I consider a framework where agents face an idiosyncratic risk to their rate of return to savings. For exogenous reasons, the private sector cannot provide insurance against these risky returns. Hence I assume that at the beginning of the second period, agents draw a rate of return  $r \in \mathbb{R}$  from some exogenously given probability distribution with c.d.f  $F : r \mapsto F(r)$ . I call capital income the product of savings  $s$  and the rate of return  $r$ .

*Scale dependence.* To account for the empirical relationship between the rate of return and the amount invested described by Bach et al. (2020) and Fagereng et al. (2020), I assume that the draw of  $r$  can depend on the chosen saving amount  $s$ .<sup>7</sup> I therefore study the conditional

---

<sup>5</sup>Here I stick to the standard Mirrlees (1971)'s framework where ex ante heterogeneity is unidimensional and boils down to differences in labor productivity. Introducing heterogeneity in initial endowments  $A$  would provide an immediate rationale for taxing capital on top of labor income, as established for instance by Cremer et al. (2003) and more recently by Piketty and Saez (2013).

<sup>6</sup>This modeling strategy also allows me to capture situations where agents have to pay labor and capital income taxes at the same time as wealth taxes.

distribution of  $r$  given  $s$ , described by the conditional density function  $f: (r, s) \mapsto f(r|s)$ . Scale dependence, therefore, arises as soon as the distribution of  $r$  is not independent from  $s$  or equivalently when  $f(r|s) \neq f(r)$  for some  $s$ . Analyzing the conditional distribution of  $r$  on  $s$  allows me not only to capture the relationship between the average rate of return and savings but also the potential influence of savings on the dispersion of returns, as documented by Bach et al. (2020) and Fagereng et al. (2020). Note that this modeling strategy implies a second capital market failure preventing agents to pool their savings to maximize scale effects.

## 2.1 | Taxpayers

I denote by  $u(\cdot)$  the function measuring utility from first-period consumption, by  $v(\cdot)$  the function measuring utility from second-period consumption, and by  $h(\cdot)$  the function measuring disutility from work effort. I assume that  $u(\cdot)$  and  $v(\cdot)$ <sup>8</sup> are both strictly increasing, twice continuously differentiable functions with  $u(\cdot)$  and  $v(\cdot)$  concave while  $h(\cdot)$  is convex. Besides, both  $u(\cdot)$  and  $v(\cdot)$  satisfy Inada's conditions with  $\lim_{x \rightarrow 0} u'(x) = \lim_{x \rightarrow 0} v'(x) = +\infty$  and

$$\lim_{x \rightarrow +\infty} u'(x) = \lim_{x \rightarrow +\infty} v'(x) = 0.$$

Let  $T: y \mapsto T(y)$  be the labor income tax schedule and  $t: (s, rs) \mapsto t(s, rs)$  be the capital tax schedule.<sup>9</sup> Assuming additively separable preferences for consumption and leisure, an agent with type  $\theta$  solves the following problem:

$$U(\theta) \stackrel{\text{def}}{=} \max_{y, s} u(A - s) + \int_{r \in \mathbb{R}} v(y - T(y) + (1 + r)s - t(s, rs))f(r|s)dr - h(y, \theta) \quad (1)$$

The solution of (1) is denoted by  $\{y(\theta), s(\theta)\}$ . First-period consumption is therefore  $c_1(\theta) = A - s(\theta)$  while second-period consumption for a given draw of  $r$  is denoted by  $c_2(\theta, r) = y(\theta) - T(y(\theta)) + (1 + r)s(\theta) - t(s(\theta), rs(\theta))$ .

Without scale dependence, the assumptions made on utility functions guarantee that absent capital taxes, the saving problem of taxpayer has a unique interior solution. However, this is no longer the case when the rate of return is increasing with savings. I, therefore, assume that the conditional density function  $f(r|s)$  is neither “too” increasing nor “too” convex in  $s$  so that the function  $\psi: s \mapsto u(A - s) + \int_{r \in \mathbb{R}} v(y + (1 + r)s - T(y))f(r|s)dr$  remains concave even with scale dependence.<sup>10</sup>

<sup>7</sup>In the extension of the main model described in Section 5, I introduce a distinction within savings between risk-free and a risky asset, where the rate of return on both assets can depend on the amount invested.

<sup>8</sup>This general formulation for second-period consumption nests the case  $v(\cdot) = \beta u(\cdot)$  with  $\beta$  the discount factor.

<sup>9</sup>Throughout the paper, “capital taxation” is used to refer to this  $t$  function. When  $t$  takes the form  $t: s \mapsto t(s)$  then the term “capital taxation” refers to wealth taxation. When  $t$  takes the form  $t: rs \mapsto t(rs)$ , the term “capital taxation” refers to capital income taxation.

<sup>10</sup>One potential microfoundation for the positive correlation between rate of return and savings are economies of scale in wealth management. In this case, it seems realistic to assume that  $f(r|s)$  is as an increasing, concave function of  $s$  so that  $r$  increases with  $s$  but at a declining rate.

## 2.2 | The government

The government levies taxes to finance an exogenous amount of public good  $E$  and to redistribute resources across agents. To do so, it can either rely on labor income taxation through  $T(\cdot)$  or capital taxation through  $t(\cdot)$ . Both  $T(\cdot)$  and  $t(\cdot)$  can be nonlinear. In the general case where  $t(\cdot)$  can freely depend on both savings  $s$  and capital income  $rs$ , the budget constraint of the government takes the form:

$$B(T, t) = \int_{\theta \in \Theta} T(y(\theta))dG(\theta) + \iint_{\theta \in \Theta; r \in \mathbb{R}} t(s(\theta), rs(\theta))f(rs(\theta))drdG(\theta) \geq E \quad (2)$$

Eventually, I suppose that the objective of the government is to maximize a generalized social welfare function that sums over all types  $\theta$  some transformation  $\Phi(U(\cdot); \theta)$  of individual utility  $U(\cdot)$  defined in (1):

$$SW \stackrel{\text{def}}{=} \int_{\theta \in \Theta} \Phi(U(\theta); \theta)dG(\theta) \quad (3)$$

Hence the problem of the government is to maximize (3) subject to the budget constraint (2). I constrain  $\Phi(\cdot)$  to be increasing in individual utility  $U$ , and to be strictly increasing for at least one type  $\theta$ . Assuming that  $\Phi(\cdot)$  can depend on skill  $\theta$  allows me to cover a wide range of welfare criteria. For instance,  $\Phi(U; \theta) \equiv \phi(\theta)U$ , where weights  $\phi(\theta)$  directly depend on type  $\theta$  embeds *weighted utilitarianism* views of justice in my framework. Hence standard *utilitarianism* is obtained when  $\phi(\theta) = 1$  while a *Rawlsian* objective arises when  $\phi(\theta) = 0$  except for the lowest type  $\underline{\theta}$  with  $\phi(\underline{\theta}) > 0$ . Social welfare as measured by (3) is quite general although as  $U$  is individual expected utility, it rules out justice principles that would value ex post utility, that is, utility measured after the draw of the rate of return.<sup>11</sup> However focusing on the justice principles embedded in (3) allows me to establish a simple criterion to elicit the optimal capital tax schedule: a capital tax schedule  $t^*$  is optimal if it increases government revenue without affecting individual (expected) utility  $U$ . I detail in the next section the method I use to find  $t^*$ .

## 2.3 | A method for finding the optimal capital tax

Throughout the paper, my objective is to characterize the optimal tax on capital, without solving for the optimal labor income tax schedule. To prove that a candidate capital tax schedule is the optimal one, I will show that it generates more government revenue than any other capital tax, without affecting taxpayers' expected utilities. Hence moving from a given capital tax function to the candidate relaxes the government's budget constraint (2) without affecting social welfare as measured by (3). I, therefore, need a method that cancels the impact on the utility of reforming capital taxation. Building on the proof of Atkinson and Stiglitz (1976) given by Konishi (1995), Laroque (2005), and Kaplow (2006),

<sup>11</sup>Social welfare as an aggregation of expected utilities is coherent with the tradition initiated by Harsanyi (1953) and is therefore subject to the discussion of this justice principle provided for instance by Diamond (1967).



I first show that starting from a given tax schedule on labor income  $T^0(y)$  and a given tax schedule on capital  $t^0(s, rs)$ , there always exists a reform of the labor income tax that neutralizes the impact on utility of implementing a candidate capital tax function  $\hat{t}(s, rs)$ .

First, note that separability allows me to split the problem defined in (1) between a subproblem over the saving decision for a given labor income  $y$  and a subproblem over the labor supply choice for a given type  $\theta$ . Hence under a given tax schedule  $\{T^0, t^0\}$  an agent with labor income  $y$  solves:

$$V^0(y) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[v(y - T^0(y) + (1 + r)s - t^0(s, rs))|s] \quad (4)$$

The full problem of an agent with type  $\theta$  under a given tax schedule  $\{T^0, t^0\}$  can therefore be written as:

$$U^0(\theta) = \max_y V^0(y) - h(y, \theta) \quad (5)$$

Denote by  $y^0(\theta)$  the solution of (5).

Now suppose that the government moves from the initial capital tax schedule  $t^0$  to a candidate tax schedule  $\hat{t}$ . Absent any behavioral responses in savings, the reform only affect (4) through a change in after-tax second-period income. This change in after-tax second-period income after a reform of capital taxation can mechanically be compensated by a reform of labor income taxation in the opposite direction. Denoting by  $\hat{T}$  this reformed labor income tax, taxpayers with labor income  $y$  would enjoy the same utility under  $\{\hat{T}, \hat{t}\}$  than under the initial tax schedule  $\{T^0, t^0\}$ . Of course, this reasoning is no longer true when taxpayers can change their saving decision in response to the reform. In this case, the existence of the utility-neutralizing labor income tax  $\hat{T}$  is not guaranteed. To be sure that there always exists such a function  $\hat{T}$ , I put the following constraints on the capital tax function:

**Assumption 1.**

- (i) The initial and the candidate capital tax function  $t^0(\cdot)$  and  $\hat{t}(\cdot)$  are twice continuously differentiable functions.
- (ii) For all  $y$ , the objective function of problem (4) is strictly concave.

The first part of Assumption 1 guarantees that the first-order condition associated to problem (4) is differentiable. Combining (i) and (ii) allows me to apply the implicit function theorem to prove that the optimal saving decision of taxpayers reacts smoothly to tax reforms. These two constraints on capital tax functions roughly correspond to the sufficient conditions for the tax perturbation method described in Jacquet and Lehmann (2021).

Note that Assumption 1 is necessarily verified when the capital tax function is linear as the objective of problem (4) would be strictly concave.<sup>12</sup> Similarly, a convex capital tax function also automatically verifies Assumption 1 as after-tax second-period income  $y + (1 + r)s - T(y) - t(s, rs)$  would be concave in  $s$ .<sup>13</sup> However concave capital tax functions can potentially violate Assumption 1 so that the optimal capital tax analysis is valid whenever the marginal tax rate on capital does not decrease too quickly with  $s$ .

Using the implicit function theorem, I prove in Appendix A that under Assumption 1, there always exists a labor income tax function  $\hat{T}$  such that:

$$V^0(y) = \max_s u(A - s) + \mathbb{E}[v(y - \hat{T}(y) + (1 + r)s - \hat{t}(s, rs))|s] \quad (6)$$

Hence for every labor income  $y$ , subutility from consumption is the same under  $\{T^0, t^0\}$  than under  $\{\hat{T}, \hat{t}\}$ . Denoting by  $\hat{U}(\theta)$  the total utility of an agent with type  $\theta$  under the tax schedule  $\{\hat{T}, \hat{t}\}$ , it follows from (6) that:

$$\begin{aligned} \hat{U}(\theta) &= \max_y V^0(y) - h(y, \theta) \\ &= U^0(\theta) \end{aligned} \quad (7)$$

It follows from (7) that total utility is the same under  $\{\hat{T}, \hat{t}\}$  than under  $\{T^0(\cdot), t^0(\cdot)\}$  and that  $y^0(\theta)$  solves the labor supply problem under both tax regimes, for every type  $\theta$ .

**Lemma 1.** *When a government moves from a given capital tax function  $t^0(s, rs)$  to a candidate  $\hat{t}(s, rs)$ , there always exists a labor income tax function  $\hat{T}(y)$  that offsets the impact of the reform on:*

1.  $V(y)$ , that is, subutility from consumption for a given labor income  $y$ .
2.  $y(\theta)$ , that is, labor income for a given type  $\theta$ .
3.  $U(\theta)$ , that is, (expected) total utility for a given type  $\theta$ .

Using Lemma 1, it is always possible to compare different capital tax regimes that leave (expected) utility, hence social welfare as measured by (3), unchanged. It is therefore possible to rank different capital tax regimes by only looking at government revenue, as defined by the left-hand side of constraint (2). Since Lemma 1 has been proved under the general case where the government can combine wealth and capital income taxation, it nests the specific cases  $t^0 : s \mapsto t^0(s)$  and  $t^0 : rs \mapsto t^0(rs)$  so that I can still rely on this Lemma in the constrained environments studied in Section 4.<sup>14</sup>

<sup>12</sup>Hence a smooth approximation of a dual system that combines a progressive labor income tax with a linear tax on capital income, as it is the case in most European countries, would verify Assumption 1.

<sup>13</sup>Such a convex capital tax function can be related to systems where capital income is taxed at an increasing rate, as it is the case in the US.

<sup>14</sup>Note that this method yields results that do not depend on social welfare weights, which is not generally the case when solving the government problem by directly maximizing (3) subject to the budget constraint (2).

### 3 | OPTIMAL CAPITAL TAX

#### 3.1 | Derivation of the optimal capital tax formula

In this section, I suppose that the government observes both savings  $s$  and capital income  $rs$ , such that only skills  $\theta$  are private information. There are no constraints on the capital tax schedule  $t(s, rs)$ , except that it should verify Assumption 1.

To characterize the optimal capital tax schedule, I rely on Lemma 1 to find a candidate  $\hat{t}$  that yields more government revenue than any initial capital tax  $t^0$  without affecting utility. Indeed from Lemma 1, for every  $t^0$  and every  $\hat{t}$  that verify Assumption 1, there exists a labor income tax function  $\hat{T}$  so that utility  $U^0(\theta)$  and labor income  $y^0(\theta)$  are the same under  $\{t^0, T^0\}$  than under  $\{\hat{t}, \hat{T}\}$ . In this case, the reform only affect social welfare through the government's budget constraint (2). A necessary and sufficient condition for the optimality of  $\hat{t}$  is therefore given by:

$$\hat{T}(y^0(\theta)) + \mathbb{E}[\hat{t}(\hat{s}(\theta), r\hat{s}(\theta))|\hat{s}(\theta)] \geq T^0(y^0(\theta)) + \mathbb{E}[t^0(s^0(\theta), rs^0(\theta))|s^0(\theta)] \quad (8)$$

with  $s^0(\theta)$  and  $\hat{s}(\theta)$  the optimal saving decision respectively under  $\{t^0, T^0\}$  and under  $\{\hat{t}, \hat{T}\}$ . Indeed, integrating (8) overall type  $\theta$  directly implies that government revenue as defined by the left-hand side of (2) is higher under  $\{\hat{t}, \hat{T}\}$  than under  $\{t^0, T^0\}$ , while utility is the same under both tax systems.

Let  $\bar{r}(s) = \mathbb{E}[r|s]$  be the average rate of return conditional on savings. Since an agent with type  $\theta$  chooses the same labor income  $y^0(\theta)$  under  $\{\hat{t}, \hat{T}\}$  than under  $\{t^0, T^0\}$ , one can use the individual's second-period budget constraint to rewrite (8) as:

$$(1 + \bar{r}(\hat{s}(\theta)))\hat{s}(\theta) - \mathbb{E}[\hat{c}_2(\theta, r)|\hat{s}(\theta)] \geq (1 + \bar{r}(s^0(\theta)))s^0(\theta) - \mathbb{E}[c_2^0(\theta, r)|s^0(\theta)] \quad (9)$$

with  $c_2^0(\theta, r)$  and  $\hat{c}_2(\theta, r)$  denoting second-period consumption for a given  $\theta$  and a given  $r$  respectively under  $\{t^0, T^0\}$  and under  $\{\hat{t}, \hat{T}\}$ . Thanks to this transformation, the condition for optimality is no longer expressed in terms of taxes but in terms of saving and second-period consumption. Recall that by definition of  $\hat{T}$ , sub-utility from consumption evaluated at labor income  $y^0(\theta)$  verifies:

$$u(A - \hat{s}(\theta)) + \mathbb{E}[v(\hat{c}_2(\theta, r)|\hat{s}(\theta))] = u(A - s^0(\theta)) + \mathbb{E}\left[v\left(c_2^0(\theta, r)|s^0(\theta)\right)\right] = V^0(y^0(\theta))$$

Hence the condition for optimality (9) is verified if  $\{\hat{s}(\theta), \hat{c}_2(\theta, r)\}$  is the solution of:

$$\begin{aligned} & \max_{s, c_2(r)} (1 + \bar{r}(s))s - \mathbb{E}[c_2(r)|s] \\ & \text{subject to: } u(A - s) + \mathbb{E}[v(c_2(r))|s] = V^0(y^0(\theta)) \end{aligned} \quad (10)$$

Without scale dependence, problem (10) would be well-behaved as the objective would be linear and the constraint set convex. As discussed in Section 2.1, I constrain the density function  $f(r|s)$  so that the constraint set remains convex. I further assume that  $f(\cdot)$  is such that

the objective of problem (10) is at least quasiconcave. In this case, I prove in Appendix B that a necessary and sufficient condition for solving (10) is given by the Euler Equation:

$$u'(A - s) = (1 + \bar{r}(s) + \bar{r}'(s)s)v'(c_2) \quad (11)$$

The question is now to find a candidate  $\hat{t}$  such that  $\{\hat{s}(\theta), \hat{c}_2(\theta, r)\}$ , that is, savings and second-period consumption under  $\{\hat{t}, \hat{T}\}$ , verify the optimality condition (11).

Using the first-order condition associated to problem (6), I know that  $\{\hat{s}(\theta), \hat{c}_2(\theta, r)\}$  verify, for any candidate  $\hat{t}$ :

$$\begin{aligned} u'(A - \hat{s}(\theta)) &= \int_{r \in \mathbb{R}} \left( 1 + r - \frac{d\hat{t}(s, rs)}{ds} \Big|_{s=\hat{s}(\theta)} \right) v'(\hat{c}_2(\theta, r)) f(r|\hat{s}(\theta)) dr \\ &+ \int_{r \in \mathbb{R}} v(\hat{c}_2(\theta, r)) f'_s(r|s) dr \end{aligned} \quad (12)$$

Hence the optimal capital tax function should be such that (12) matches with (11). First, it follows from (11) that at the optimum, second-period consumption of an agent with type  $\theta$  should not depend on the specific draw of  $r$ :  $\hat{c}_2(\theta, r) = \hat{c}_2(\theta)$ . In other words, the optimal tax function induced by (11) provides insurance against risky returns so that second-period consumption becomes deterministic. Second, since the optimal capital tax implied by (11) depends on the average rate of return conditional on savings  $\bar{r}(s)$ , a natural candidate for the optimal policy would combine a confiscatory tax on capital income with some wealth subsidy depending on  $\bar{r}(s)$ . Hence consider the candidate tax function:

$$t^*(s, rs) \mapsto rs - \bar{r}(s)s$$

Under this capital tax function, second-period consumption satisfies the insurance condition  $c_2(r) = c_2$ . Besides the derivative of this specific candidate with respect to  $s$  verifies  $\frac{dt^*}{ds} = r - \bar{r}(s) - \bar{r}'(s)s$ . It therefore follows from (12) that taxpayers' saving decision and second-period consumption under  $t^*$  verify (11) and therefore solves problem (10). Hence for any initial capital tax schedule  $t^0$  that verifies Assumption 1, the government can guarantee the same individual utility level while raising more tax revenue by implementing  $t^*$ .

**Proposition 1.** *In the case where the government can tax both capital income  $rs$  and savings  $s$ , the optimal capital tax function is given by:*

$$t^*(s, rs) = rs - \bar{r}(s)s$$

### 3.2 | Implications of Proposition 1

Proposition 1 implies that when the government can freely tax savings and capital income, the optimal policy consists in providing insurance against the volatility of  $r$  at each saving level  $s$ . Indeed, in presence of scale dependence, taxpayers making different saving decisions have access to different average rate of return and take this into account when choosing the level of

their savings. Hence the optimal policy grants insurance against the volatility of  $r$  at each saving level  $s$ . By doing so, an agent saving  $s$  and knowing that this would yield on average a rate of return  $\bar{r}(s)$  will earn on average the same capital income under the optimal policy than in the laissez-faire economy. But in the laissez-faire economy, consumption smoothing required agents to oversave to insure against the risk of drawing a low rate of return. This “oversaving” motive disappears when  $t^*$  is implemented. Proposition 1 is fairly general and include a wide class of distributions of the rate of return. Some special cases are nevertheless worth mentioning.

### 3.2.1 | Risky rates of return, increasing with $s$

In this setting, scale dependence triggers a “rich-get-richer” effect as high savers would also earn higher (average) rate of return. Nevertheless, this rich-get-richer effect does not call for redistributive capital taxes, as Proposition 1 only implies transfers between lucky and unlucky taxpayers who made the same saving decision. In particular, there is no transfers between groups of savers through capital taxation, as the only redistributive tool remains labor income taxation. However, the rich-get-richer effect can actually affect the shape of the *isotax curves*, understood here as the combination of  $r$  and  $s$  that yields the same capital tax liability. It indeed follows from the optimal tax formula of Proposition 1, that at a given capital tax level  $t$ , the relation between  $r$  and  $s$  is given by:  $r = \frac{t}{s} + \bar{r}(s)$ . The shape of the isotax curves can therefore be described by:

$$\frac{\partial r}{\partial s} = \frac{-t}{s^2} + \bar{r}'(s)$$

In the case of risky returns without scale dependence, the isotax curves for those earning above average  $r$ , thus with  $t > 0$ , would be unambiguously decreasing. This means that a high capital tax liability is associated either to a lucky draw of  $r$  or to a high level of savings, but not both. However, when the average rate of return increases with  $s$ , the isotax curves can actually be increasing. Hence if the “rich-get-richer” effect is strong enough, a high capital tax liability would be associated to both a high rate of return and a high level of savings. In other words, the burden of capital taxes would systematically fall on the lucky rich.

### 3.2.2 | The deterministic case with $r = \bar{r}(s)$

In this setting, Proposition 1 implies  $t^*(.) = 0$ , even if the rate of return does depends on  $s$ . This is not surprising as in absence of risky returns, agents are heterogeneous only along one dimension, which is the skill heterogeneity captured by the parameter  $\theta$ . Therefore, as soon as the government observes and taxes labor income  $y(\theta)$ , savings  $s(\theta)$  does not reveal any additional information on the unobserved parameter  $\theta$ . Capital taxation on top of labor income taxation is therefore useless in this Atkinson and Stiglitz (1976) economy. Scale dependence, even when it results in higher returns for richer individuals, does not alter this intuition.

Both Gahvari and Micheletto (2016) and Kristjánsson (2016) already argue that scale dependence, understood as an increasing relationship between savings and the rate of return, does not advocate for capital income taxation. In this sense, Proposition 1 extends their results

obtained in a two-type economy to a continuous type environment in the tradition of Mirrlees (1971). Note however that my modeling strategy is slightly different than the one used in these two papers, since they both consider a two-period model where labor income is earned at time 1 while being taxed at time 2. To ensure that this difference in the timing of labor income does not affect my conclusions, I prove in Appendix G that Proposition 1 extends to the case where labor income is earned at time 1 while being taxed at time 2, as in Gahvari and Micheletto (2016) and Kristjánsson (2016). This emphasizes the idea that scale dependence does not per se provide a strong rationale for redistributive capital taxes.<sup>15</sup>

The absence of capital taxation in the deterministic case with scale dependence however differs from the results of Gerritsen et al. (2020), where capital income taxation is part of the optimal policy when rates of return increase with savings. The key reason for this difference in results is due to the assumptions on the timing of taxes. Gerritsen et al. (2020) consider a two-period model where labor income is earned and taxed at time 1 while capital income is earned and taxed at time 2. In this case, capital income taxes can be used to alleviate the market failure that prevents the poor from lending to the rich to benefit from their higher marginal rates of return in a context of scale dependence. This missing transaction can be replicated by lowering taxes at time 1 for the rich, which is equivalent to transferring funds from the poor to the rich at time 1, and by increasing taxes for the rich at time 2 so that the government can transfer part of the higher returns of the rich to the poor at time 2. In Gerritsen et al. (2020), taxes at time 1 are labor income taxes and taxes at time 2 are capital income taxes. Eventually, the optimal capital income tax is the result of a trade-off between alleviating market failure and distorting savings. Unless the elasticity of savings with respect to the capital income tax is infinite, this trade-off results in strictly positive capital income taxation at the optimum.<sup>16</sup> Assuming that labor income taxes are also levied at time 2, as in Gahvari and Micheletto (2016), Kristjánsson (2016), and myself, implies that labor income taxes are as good as capital income taxes to alleviate the market failure preventing the rich to save on behalf of the poor. Hence, without risk, capital income taxes only add distortion to savings in our framework so it should be zero at the optimum.<sup>17</sup> Note that Gerritsen et al. (2020) rule out wealth taxation, that is, a tax on savings, by hypothesis while I show that both capital income and wealth taxes are not needed in deterministic environments with scale dependence.

To the best of my knowledge, Proposition 1 is the first to characterize the optimal capital tax in a context of both uninsurable risk and scale dependence on the rate of return. Yet, some remarks on the limit of my analysis are needed to better assess the scope of Proposition 1.

First, note that except for risky returns and scale dependence, the framework considered here is the same as Atkinson and Stiglitz (1976) and is therefore prone to the standard critics regarding the separability and preference homogeneity assumptions that drive Atkinson-Stiglitz's theorem.<sup>18</sup> Although I depart from the return homogeneity implied by Atkinson and Stiglitz (1976), I do not allow individuals to differ in their innate ability to obtain higher

---

<sup>15</sup>Assuming that agents are ex ante heterogeneous with respect to their wealth endowment  $A$  would provide a rationale for redistributive capital taxation. This could however be considered as a problem of optimal inheritance taxation, as in Piketty and Saez (2013), which is beyond the scope of this paper.

<sup>16</sup>see proposition 3 of Gerritsen et al. (2020).

<sup>17</sup>Eventually a careful discussion of the timing of taxes would benefit from modeling of public debt, which is not straightforward in a context of scale dependence, as rightly pointed out by Gerritsen et al. (2020).

<sup>18</sup>Christiansen (1984) for instance shows how positive indirect taxation can be part of the optimal policy when separability does not hold. Saez (2002) and Diamond and Spinnewijn (2011) focus on cases where taste homogeneity for savings does not hold. Interestingly, a fierce critic of the application of Atkinson and Stiglitz (1976) to capital taxation can be found in Section 2 of Stiglitz (2018).

returns. In other words, ex ante heterogeneity boils down to labor productivity heterogeneity as in Mirrlees (1971) so that differences in entrepreneurial skills are not considered in the present analysis. As a consequence, the rationale for using capital taxation to allocate funds toward the most productive *entrepreneurs* is ruled out by hypothesis here. Since this mechanism has been highlighted as quantitatively relevant by Guvenen et al. (2019), it is worth having in mind this limit of my analysis when discussing optimal policy. I stick to the Atkinson and Stiglitz (1976)'s framework as my objective is to pin down the specific implications of risk and scale dependence in a world where I know that, absent these two mechanisms, capital taxation is ruled out at the optimum.

Second, the optimal capital tax function  $t^*$  assigns the role of insurance provider to the government. This role emerges because it is assumed that the private sector cannot provide this insurance. Yet the underlying reason behind this absence of private insurance can have important policy implications. When I assume that the distribution of the rate of return is exogenously given, I actually suppose that market incompleteness arises because of exogenous reasons. This implies that government intervention, here through capital taxation, does not affect private insurance provision. However, if there exists endogenous reasons for market incompleteness, then capital taxation can actually crowd-out private insurance, as described by Krueger and Perri (2011). Perhaps more importantly for welfare analysis, if idiosyncratic risk is actually used by private markets to provide incentives, then the policy described in Proposition 1 can actually be detrimental to welfare.<sup>19</sup>

Eventually, note that the perhaps most obvious limit of the tax schedule described in Proposition 1 is its feasibility. Indeed, such a tax schedule requires that the government perfectly observes individual savings and capital income, from which the rate of return can be inferred. This is why in the rest of the paper, I consider alternative settings that are less demanding in terms of government's information set, assuming that either savings or capital income are only privately observed.

## **4 | CAPITAL TAXATION WITH IMPERFECT INFORMATION ON CAPITAL**

As shown in the previous section, riskiness and scale dependence, two components often studied separately, have to be both taken into account when designing the optimal capital tax policy. The objective in this section is to understand how riskiness and scale dependence jointly affect policy when the government has constraints on the available information on capital. Since the optimal tax described in Proposition 1 is a function of both savings and capital income, I study here the consequences of assuming that either one of these two dimensions of capital is only privately observed. This exercise can be insightful first to clarify whether heterogeneity in savings or heterogeneity in capital income can on their own justify government intervention in a context of risky, scale-dependent rates of return. Second, in real-world economies there exist several constraints, either informational or institutional, on what tax the government can actually impose on capital. Hence assuming that some components of capital are only privately observed allows me to determine whether there is still room for capital taxation in these constrained but perhaps more realistic settings.

---

<sup>19</sup>Section 4 of Buchholz and Konrad (2014) documents how insurance provision through taxes can prevent markets from achieving better resource allocations when returns depend on an unobservable effort by entrepreneurs.

## 4.1 | Optimal capital taxation when capital income is private information

Suppose that the government cannot observe capital income  $rs$  so that capital can only be taxed through savings  $s$ . In other words, the capital tax function is now defined as  $t : s \mapsto t(s)$ , which in my framework is equivalent to a tax on (ex ante) wealth. Following the same route as in Section 3, I rely on Lemma 1 to compare government revenue under tax systems that otherwise guarantee the same level of individual utility. Lemma 1 applies for any initial capital tax schedule  $t^0(s, rs)$ , and therefore nests the case  $t^0 : s \mapsto t^0(s)$ , as long as  $t^0$  verifies Assumption 1. I show in Appendix C that moving from an economy with some tax  $t^0(s)$  to a zero-capital tax economy is actually welfare improving.

**Proposition 2.** *If the government observes savings but capital income remains private information, then there should be no capital income tax at the optimum:  $t^*(s) = 0$ .*

The intuition behind Proposition 2 is first that a tax on savings does not have interesting properties in terms of insurance against risky returns. Indeed, taxing only savings  $s$  independently of the realization of the shock on  $r$  rules out by hypothesis the possibility of reallocating resources from the lucky to the unlucky ones. Hence the key rationale for government intervention in the unconstrained setting disappears. Nevertheless, a tax  $t(s)$  could still affect the draw of  $r$  since distorting  $s$  changes the distribution of the rate of return in the context of scale dependence. One could think of taxing or subsidizing savings to deter agents from drawing their rate of returns within a distribution with undesirable properties, be it in terms of volatility or average returns. Yet again, under homogeneous and separable preferences, the saving decision mechanically stems from the labor income decision. Hence taxing labor income is sufficient to reallocate savings in a direction deemed socially desirable. This is why Proposition 2 is similar to the nondesirability of indirect taxation in Atkinson and Stiglitz (1976).

From a more practical point of view, the assumption that savings are observed and can be taxed while capital income cannot, is at odds with actual tax systems since to the best of my knowledge, taxes on savings, which can be seen here as wealth taxes, never occur without capital income taxation. It is however often the case that capital income is taxed while savings are not. This is the setting studied in the next section.

## 4.2 | Optimal capital income taxation when savings are private information

Here I conduct the opposite exercise compared to the one done in the previous section: assume that savings are only privately observed but the government can tax capital income. The capital tax function now is constrained to take the form  $t : rs \mapsto t(rs)$ . To fully understand the consequences of scale dependence on optimal policy in my model, I first depart from the general case described in (1) by assuming that  $r$  is scale dependent but certain. An individual with type  $\theta$  would therefore solve:

$$U(\theta) \stackrel{\text{def}}{=} \max_{y,s} u(A - s) + v(y - T(y) + (1 + r(s))s - t(r(s)s)) - h(y, \theta) \quad (13)$$



Using the method described in Section 2.3, I show in Appendix D that in this particular setting, capital income taxation is not desirable.

**Proposition 3.** *When the government observes capital income but not savings, in a context where the rate of return is scale dependent but deterministic, there is no capital tax at the optimum:  $t_D^*(rs) = 0$ .*

This is not surprising as without uncertainty, observing capital income does not provide more information than just observing savings. Hence the rationale against taxing capital even in presence of scale dependence used for explaining Proposition 2 extends to Proposition 3. However, with uncertain returns, taxing capital income has different properties than taxing savings as the tax occurs after the realization of the shock. Hence a tax on capital income is a tax on risky returns, although the government is unable to distinguish between the risky component of capital income  $r$  and the deterministic component  $s$ . This prevents the government from achieving the perfect insurance allocation described in Section 3 but this could leave room for positive capital taxation. To examine this, I go back to the general problem exposed in (1) with  $t : rs \mapsto t(rs)$ . My objective here is not to pin down the optimal capital tax as done previously but only to examine if there is positive capital income taxation in this constrained environment. To do so, I measure the impact on an individual's fiscal contribution (hence on government revenue) of introducing a small linear tax with rate  $\tau$  on capital income after neutralizing its impact on individual utility. To do so, I measure the impact on an individual's fiscal contribution (hence on government revenue) of introducing a small linear tax on capital income, with rate  $\tau$ . Note that the capital income tax  $t : rs \mapsto \tau \cdot rs$  necessarily verifies Assumption 1 so that I can always apply Lemma 1 to offset the impact on utility of reforming capital income taxation. In Appendix E.1, I show that this first-order impact verifies the following equation:

$$\left. \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} \right|_{\tau=0} = -s^0 \left( \frac{\text{Cov}[r; v'(y^0 - T^0(y) + (1+r)s^0) | s^0]}{\mathbb{E}[v'(y^0 - T^0(y) + (1+r)s^0) | s^0]} \right) \quad (14)$$

with  $\mathcal{B}(y, \tau)$  the expected fiscal contribution of an individual with labor income  $y$  while  $s^0$  and  $T^0(y)$  are respectively savings and labor income tax payments when there is no capital tax ( $\tau = 0$ ). The concavity of the utility function  $v(\cdot)$  guarantees that the covariance term is negative so that the left-hand side of (14) is positive.

**Proposition 4.** *When savings are private information, introducing a strictly positive linear tax on capital income in an economy with risky, scale-dependent rates of return generates a first-order welfare improvement.*

Proposition 4 establishes the desirability of capital income taxation even when the government does not observe initial savings. It proves that risky returns justify capital income taxation in a general framework that includes scale dependence. The key intuition is that capital income taxation not only reduces the expected return but also the variance of risky returns. Since agents are assumed risk-averse, the insurance provided by capital income taxation reduces the compensation needed to offset the impact  $\tau$  has on utility. This is why the difference between what the government on average gets when implementing  $\tau$ , that is,  $\bar{r}(s)$ , and what the government loses, through the decrease in labor income taxes needed to

compensate agents, is positive. Note that this reasoning extends to the case of an ex post *wealth tax*  $t : (1 + r)s \mapsto t((1 + r)s)$ , as proved in Appendix E.2.

The desirability of capital income taxation in risky environment has been established in various settings previously in the literature. The closest one to the exercise made in this section is Varian (1980), that I extend by considering scale dependence and by studying a pure Mirrleesian economy with endogenous labor supply. I therefore show that the desirability of capital income taxation is maintained, even when it triggers distortions to labor supply choices that were ruled out by hypothesis in Varian (1980). Using a portfolio choice model Domar and Musgrave (1944) also provides a justification for positive capital income taxation in risky environment. This approach has been recently introduced in a continuous-type Mirrleesian economy by Boadway and Spiritus (2021), although again none of these two papers featured scale dependence. In the following section, I introduce a distinction within savings between a safe and a risky asset to relate more directly to this branch of the literature and see how Propositions 1, 2, and 4 can be extended to such portfolio choice settings. Proposition 4 only establishes the desirability of capital income. In Appendix E.3, I show that the actual optimal linear tax on capital income verifies the following formula:

**Proposition 5.** *The optimum linear tax on capital income  $\tau^*$  verifies:*

$$\frac{\tau^*}{1 - \tau^*} = - \left( \frac{\text{Cov}[rv'(\cdot)|s]}{\bar{r}(s)E[v'(\cdot)|s]} \right) \cdot \frac{1}{\epsilon_{1-\tau}^s} \cdot \frac{1}{1 + \epsilon_s^{\bar{r}}} \quad (15)$$

where  $\epsilon_{1-\tau}^s = \frac{\partial s}{\partial 1-\tau} \frac{1-\tau}{s}$  is the elasticity of savings with respect to the net of capital income tax rate and  $\epsilon_s^{\bar{r}} = \frac{\partial \bar{r}}{\partial s} \frac{s}{\bar{r}}$  is the elasticity of the average rate of return with respect to savings.

Proposition 5 helps clarifying the role of (linear) capital income taxation in a context of scale dependence and risky returns. Note that absent risk, the covariance term in (15) is zero so that there is no capital income tax at the optimum, even with scale dependence. This emphasizes the idea that the primary role of capital taxation in my framework is to provide insurance against risky returns. Hence the more risk-averse agents are, the stronger this insurance motive and the higher the capital income tax. Second, the  $\frac{1}{\epsilon_{1-\tau}^s}$  term illustrates the standard distortion to savings capital income taxation can create. Following the logic of the inverse elasticity rule of Ramsey (1927), the higher the elasticity of savings with respect to taxes, the lower should be the tax rate. Absent scale effects, the  $\epsilon_s^{\bar{r}}$  is zero and (15) simply illustrates the trade-off between providing insurance and distorting savings. However, in presence of scale dependence, capital income taxation not only affects savings but also the average rate of return. Hence for a given drop of  $s$ , the drop in capital income will be more important in an economy where scale effects are stronger. In other words, the more the rate of return increases with savings, the less desirable capital income taxation is. This amplification of the response of the tax base due to scale dependence is reminiscent of Scheuer and Werning (2017) where superstar effects increase the responsiveness of earnings to taxes. So if anything, scale dependence makes the tax base more elastic, hence pushing optimal tax rates down. However, to the best of my knowledge, there exist no estimates of the  $\epsilon_s^{\bar{r}}$  parameter in the empirical literature so far so the importance of this amplification response in practice is yet to be determined.

## 5 | EXTENSION: A TWO-ASSET VERSION OF THE MODEL

I now introduce a distinction within savings between what is invested in a risk-free asset and what is invested in a risky one. Such a distinction allows me to explicitly state the optimal capital tax results in terms usually used by the literature initiated by Domar and Musgrave (1944) and by the policy recommendations expressed in the Mirrlees and Adam (2010).

So now suppose that individuals can diversify their portfolio by choosing how much they invest in risk-free bonds  $b$  and risky assets  $z$ . The rate of return on  $b$  is known with certainty and denoted by  $r(b)$ . Indeed, as documented by Fagereng et al. (2020), even safe assets can be exposed to scale effects. Expressing the risk-free rate of return as a function of  $b$  allows me to include scale dependence on this specific asset.<sup>20</sup> Eventually I denote by  $x$  the return on the risky asset, drawn from a distribution with p.d.f  $f(x|z)$ , the conditional term capturing potential scale effects in investment in the risky asset. In general, the capital tax function can freely depend on the capital income generated by both the risk-free and the risky asset but also on the amounts initially invested in these two assets. I, therefore, consider the capital tax function  $t : (b, rb, z, xz) \mapsto t(b, rb, z, xz)$ . Hence an individual with type  $\theta$  under some tax schedule  $\{T, t\}$  solves:

$$U(\theta) \stackrel{\text{def}}{\equiv} \max_{y,b,z} u(A - b - z) + \mathbb{E}[v(y - T(y) + (1 + r(b))b + (1 + x)z - t(\cdot))|z] - h(y, \theta) \quad (16)$$

Because of separability, the amount invested in the risky free asset  $b$  and the risky asset  $z$  for a given labor income  $y$  is the solution of the following subproblem:

$$V^0(y) \stackrel{\text{def}}{\equiv} \max_{b,z} u(A - b - z) + \mathbb{E}[v(y - T^0(y) + (1 + r(b))b - t^0(b, rb, z, xz))|z] \quad (17)$$

Following exactly the same route as in Section 2.3, I want to establish the validity of Lemma 1 in this two-assets context. As in Section 2.3, I impose the following regularity assumption on the capital tax function:

### Assumption 2.

- (i) The capital tax function  $t : (b, rb, z, xz) \mapsto t(b, rb, z, xz)$  is twice continuously differentiable.
- (ii) For all  $y$ , the objective function of problem (17) is strictly concave.

Then I prove in Appendix F.1 that under Assumption 2, it is always possible to offset the impact a reform of the capital tax  $t$  has on utility and labor supply by reforming the labor income tax  $T$ :

<sup>20</sup>Figure 3 of Fagereng et al. (2020) documents the increasing relationship between returns and wealth for safe assets. For instance the average rate of return is negative ( $-0.05\%$ ) for the bottom 10% of the wealth distribution while it rises to more than 1% for the top 10%. The authors suggest that part of this correlation can be attributed to scale effects in deposits management.

**Lemma 2.** *When a government moves from a given capital tax function  $t^0(b, rb, z, xz)$  to a candidate  $\hat{t}(b, rb, z, xz)$ , there always exists a labor income tax  $\hat{T}(y)$  that offsets the impact of the reform on:*

1.  $V(y)$ , that is, subutility from consumption for a given labor income  $y$ .
2.  $y(\theta)$ , that is, labor income for a given type  $\theta$ .
3.  $U(\theta)$ , that is, (expected) total utility for a given type  $\theta$ .

## 5.1 | Optimal capital tax function with two assets

Building on the reasoning made in Section 3, I guess that capital taxation would be used to fully insure agents against the volatility in returns to the risky asset. I therefore guess that the optimal capital tax schedule is given by the following tax function:

$$\hat{t} : (b, rb, z, xz) \mapsto xz - \bar{x}(z)z \quad (18)$$

Such a tax function would in particular imply that the risk-free asset  $b$  and its capital income  $r(b)b$  is not taxed at the optimum, even in presence of scale dependence on this specific asset.

**Proposition 6.** *In the case where the government can tax the risk-free and risky asset along with the capital income generated by these investments, the optimal capital tax function is given by:*

$$t^*(b, rb, z, xz) = xz - \bar{x}(z)z$$

The proof is given in Appendix F.2. Scale dependence in investment in the risky asset does not call for redistributive capital taxes but only for transfers between lucky and unlucky taxpayers, at a given level of  $z$ . Moreover, introducing scale dependence in the risk-free asset, that is, scale dependence in a deterministic environment, does not justify capital taxation, as discussed in Section 3. Proposition 6 can therefore justify a tax on “above normal” or “excess return” while leaving the risk-free return tax free. In this sense, it extends to a context of scale dependence the reasoning of Mirrlees et al. (2011) which states that taxes on the risk-free asset are not recommended as part of the optimal capital tax policy. However, although scale dependence does not per se justify capital taxation, the dependence between the rate of return and the amount invested has to be taken into account when defining what are “excess” or “above normal” return. Indeed, in a context of scale dependence on the risky asset, what is above normal depends on the amount invested so that “excess return” is defined as returns above the expected rate of return, conditional on the size of the investment. As documented by Fagereng et al. (2020), there are evidence of an increasing relationship between the average rate of return on risky assets and agents’ wealth.<sup>21</sup> Hence, implementing  $t^*$  instead of the naive tax on excess return  $t^{ER} = xz - \bar{x}z$  would actually imply a lower rate of return for people in the bottom of the wealth distribution and a higher rate of return for people in the top of the wealth

<sup>21</sup>See for instance figure 3 of Fagereng et al. (2020).

distribution. To give a quantitative illustration, assume that the empirical counterpart of what I call “risky asset” would be “direct stockholding” as measured by Fagereng et al. (2020). Then, implementing  $t^*$  would grant a 6% rate of return on direct stockholding for people in the 10th percentile of the (financial) wealth distribution while people in the 90th percentile would get a rate of return of 8%.<sup>22</sup> The intuition for this a priori counter-intuitive implication of Proposition 6 lies again in the assumption of Atkinson and Stiglitz (1976): under separability and unidimensional ex ante heterogeneity, people earning the same labor income will make the same investment decision. Hence if the government is concerned about the distribution of returns due to scale dependence, it only needs to change the distribution of labor income, using the labor income tax function.

## 5.2 | Capital taxation with two assets and imperfect information

Ruling out by hypothesis the optimal capital tax function described in Proposition 6, I constrain the information set of the government by assuming that some components of capital are only privately observed. I first study the case where only risky capital income  $xz$  is not observed by the government so that the capital tax function can only take the form  $t : (b, rb, z) \mapsto t(b, rb, z)$ . In other words, I suppose that both the initial investment and the capital income of the risk free asset can be taxed while only the initial investment in the risky asset can be taxed. This implies that only deterministic components of capital can be taxed. I prove in Appendix F.3 that in this case there is no need for capital taxation.

**Proposition 7.** *When the capital income generated by the risky asset is only privately observed then there is no capital taxation at the optimum:*

$$t^*(b, rb, z) = 0$$

Again, this formalizes in a context of scale dependence one implication of the *Rate of Return Allowance* (RRA) proposal of Mirrlees et al. (2011), namely that “for assets where only the risk-free (“normal”) rate of return is likely to be earned, this approach (i.e., *the RRA approach*) can be simplified, and returns on such assets can just be tax free.”<sup>23</sup> Not only this logic applies to the return of the risk-free asset, but also to the initial investment in both the risk-free and the risky asset. To sum up, every deterministic component of capital is left untaxed in this framework. The question is now to derive the optimal capital tax policy when the risky component is observed by the government, and can therefore be taxed, but the initial investment remains private information, so that the perfect insurance program described in Proposition 6 cannot be achieved.

**Proposition 8.** *Implementing a strictly positive linear tax on the capital income generated by the risky asset creates a first-order welfare improvement:  $t^*(xz) > 0$ .*

<sup>22</sup>Here I use Figure OA.12 of the Online Appendix to Fagereng et al. (2020).

<sup>23</sup>Mirrlees et al. (2011, p. 336).

The proof can be found in Appendix F.4. Although imperfect, the insurance provided by a tax on the capital income generated only by the risky asset justifies strictly positive capital taxation. In terms of optimal linear tax, the logic of formula (15) can be extended to the two asset case:

**Proposition 9.** *The optimal linear tax rate on the capital income associated to the risky asset verifies:*

$$\frac{\tau^*}{1 - \tau^*} = - \left( \frac{\text{Cov}[xv'(\cdot)|z]}{\bar{x}(z)E[v'(\cdot)|z]} \right) \cdot \frac{1}{\epsilon_{1-\tau}^z} \cdot \frac{1}{1 + \epsilon_z^{\bar{x}}} \quad (19)$$

where  $\epsilon_{1-\tau}^z$  is the elasticity of investment in the risky asset with respect to the net of tax rate and  $\epsilon_z^{\bar{x}}$  denotes the elasticity of the average rate of return on the risky asset with respect to the amount invested in the risky asset.

The proof is given in Appendix F.5. Formula (19) shows that the key trade-off for capital income taxation in this two version of the model is similar to the one captured by formula (15): taxation of risky capital income is welfare improving because of the insurance it creates, although this positive impact is tempered by the distortions on investment in the risky asset created. This negative effect on investment is amplified when the rate of return increases with the amount invested, as captured by the  $1 + \epsilon_z^{\bar{x}}$  term in formula (19).

## 6 | CONCLUSION

Capital taxation is widely used in developed economies although there still exists a debate regarding the theoretical justification for this specific tax instrument. In this paper, I discuss how two empirical properties of returns to savings, namely that they are uncertain and scale-dependent, affect the standard way of considering capital taxation. More precisely, starting from the Atkinson and Stiglitz (1976) benchmark where I know that absent these two effects, capital taxation would be zero at the optimum, I show how optimal policy is affected by uncertainty and scale dependence.

In every specification of the model, uncertainty in returns to savings unambiguously calls for nonzero capital taxation at the optimum because of the insurance it provides. The way capital taxation provides insurance against uncertain returns is however affected by the relationship between the rate of return and the amount invested. Such scale dependence first prevents the government from granting access to the same rate of return to everyone at the optimum: only agents investing the same amount should share the same rate of return. Besides, when the rate of return increases with the amount invested, it creates additional distortions to the tax base and therefore lower capital income tax rates.

One of the main takeaway of my analysis is that scale dependence on its own does not provide a strong rationale for deviating from the zero capital tax benchmark. However, when capital taxation is justified for other reasons, scale dependence has to be taken into account when designing the optimal policy. In my framework, I have shown how the shape of capital taxes are affected by scale dependence, although the reason why capital has to be taxed in the first place is the presence of uncertain returns. It would therefore be interesting to see how

scale dependence plays a role in optimal capital taxes in alternative frameworks considering other deviations from Atkinson and Stiglitz (1976) to justify non-zero capital taxation.

## ACKNOWLEDGMENT

I thank Pierre-Edouard Collignon, Antoine Ferey, Etienne Lehmann, Jonas Loebbing and Dylan Moore as well as participants of the CESifo Area Conference on Public Economics 2021 and of the IIPF 2020 and 2021 for their precious comments.

## ORCID

Eddy Zanoutene  <http://orcid.org/0000-0001-9093-6678>

## REFERENCES

- Atkinson, A. B., & Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, 6(1–2), 55–75. [https://doi.org/10.1016/0047-2727\(76\)90041-4](https://doi.org/10.1016/0047-2727(76)90041-4)
- Bach, L., Calvet, L. E., & Sodini, P. (2020). Rich pickings? Risk, return, and skill in household wealth. *American Economic Review*, 110(9), 2703–2747. <https://doi.org/10.1257/aer.20170666>
- Benhabib, J., & Bisin, A. (2018). Skewed wealth distributions: Theory and empirics. *Journal of Economic Literature*, 56(4), 1261–1291. <https://doi.org/10.1257/jel.20161390>
- Benhabib, J., Bisin, A., & Zhu, S. (2011). The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1), 123–157. <https://doi.org/10.3982/ECTA8416>
- Boadway, R., & Spiritus, K. (2021). *Optimal taxation of normal and excess returns to risky assets*. Tinbergen Institute Discussion Paper 2021-025/VI.
- Buchholz, W., & Konrad, K. A. (2014). *Taxes on risky returns—an update*. Technical report, Working Paper of the Max Planck Institute for Tax Law and Public Finance.
- Bulow, J. I., & Summers, L. H. (1984). The taxation of risky assets. *Journal of Political Economy*, 92(1), 20–39. <https://doi.org/10.1086/261206>
- Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica: Journal of the Econometric Society*, 54(3), 607–622. <https://doi.org/10.2307/1911310>
- Christiansen, V. (1984). Which commodity taxes should supplement the income tax? *Journal of Public Economics*, 24(2), 195–220. [https://doi.org/10.1016/0047-2727\(84\)90024-0](https://doi.org/10.1016/0047-2727(84)90024-0)
- Cremer, H., Pestieau, P., & Rochet, J.-C. (2003). Capital income taxation when inherited wealth is not observable. *Journal of Public Economics*, 87(11), 2475–2490. [https://doi.org/10.1016/S0047-2727\(02\)00050-6](https://doi.org/10.1016/S0047-2727(02)00050-6)
- Diamond, P., & Spinnewijn, J. (2011). Capital income taxes with heterogeneous discount rates. *American Economic Journal: Economic Policy*, 3(4), 52–76. <https://doi.org/10.1257/pol.3.4.52>
- Diamond, P. A. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment. *The Journal of Political Economy*, 75(5), 765. <https://doi.org/10.1086/259353>
- Domar, E. D., & Musgrave, R. A. (1944). Proportional income taxation and risk-taking. *The Quarterly Journal of Economics*, 58(3), 388–422. <https://doi.org/10.2307/1882847>
- Fagereng, A., Guiso, L., Malacrino, D., & Pistaferri, L. (2020). Heterogeneity and persistence in returns to wealth. *Econometrica*, 88(1), 115–170. <https://doi.org/10.3982/ECTA14835>
- Gabaix, X., Lasry, J.-M., Lions, P.-L., & Moll, B. (2016). The dynamics of inequality. *Econometrica*, 84(6), 2071–2111. <https://doi.org/10.3982/ECTA13569>
- Gahvari, F., & Micheletto, L. (2016). Capital income taxation and the Atkinson-Stiglitz theorem. *Economics Letters*, 147, 86–89. <https://doi.org/10.1016/j.econlet.2016.08.021>
- Gerritsen, A., Jacobs, B., Rusu, A. V., & Spiritus, K. (2020). *Optimal taxation of capital income with heterogeneous rates of return*. Technical report, CESifo Working Paper.
- Gordon, R. H. (1985). Taxation of corporate capital income: Tax revenues versus tax distortions. *The Quarterly Journal of Economics*, 100(1), 1–27. <https://doi.org/10.2307/1885733>
- Güvenen, F., Kambourov, G., Kuruscu, B., Ocampo-Diaz, S., & Chen, D. (2019). *Use it or lose it: Efficiency gains from wealth taxation*. Technical report, National Bureau of Economic Research.
- Harding, M. (2013). *Taxation of dividend, interest, and capital gain income*. Technical report, OECD.

- Harsanyi, J. C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy*, 61(5), 434–435. <https://doi.org/10.1086/257416>
- Jacquet, L., & Lehmann, E. (2021). Optimal income taxation with composition effects. *Journal of the European Economic Association*, 19(2), 1299–1341. <https://doi.org/10.1093/jeea/jvaa022>
- Judd, K. L. (1985). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics*, 28(1), 59–83. [https://doi.org/10.1016/0047-2727\(85\)90020-9](https://doi.org/10.1016/0047-2727(85)90020-9)
- Kaplow, L. (2006). On the undesirability of commodity taxation even when income taxation is not optimal. *Journal of Public Economics*, 90(6–7), 1235–1250. <https://doi.org/10.1016/j.jpubeco.2005.07.001>
- Kaplow, L. (2010). *The theory of taxation and public economics*. Princeton University Press.
- Konishi, H. (1995). A pareto-improving commodity tax reform under a smooth nonlinear income tax. *Journal of Public Economics*, 56(3), 413–446. [https://doi.org/10.1016/0047-2727\(94\)01444-S](https://doi.org/10.1016/0047-2727(94)01444-S)
- Kristjánsson, A. S. (2016). *Optimal taxation with endogenous return to capital*. Technical report, Memorandum.
- Krueger, D., & Perri, F. (2011). Public versus private risk sharing. *Journal of Economic Theory*, 146(3), 920–956. <https://doi.org/10.1016/j.jet.2010.08.013>
- Laroque, G. R. (2005). Indirect taxation is superfluous under separability and taste homogeneity: A simple proof. *Economics Letters*, 87(1), 141–144. <https://doi.org/10.1016/j.econlet.2004.10.010>
- Mirrlees, J., Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., Gammie, M., Johnson, P., Myles, G., & Poterba, J. (2011). The mirrlees review: Conclusions and recommendations for reform. *Fiscal Studies*, 32(3), 331–359. <https://doi.org/10.1111/fisc.2011.32.issue-3>
- Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2), 175–208. <https://doi.org/10.2307/2296779>
- Mirrlees, J. A., & Adam, S. (2010). *Dimensions of tax design: The Mirrlees review*. Oxford University Press.
- Piketty, T., & Saez, E. (2013). A theory of optimal inheritance taxation. *Econometrica*, 81(5), 1851–1886. <https://doi.org/10.3982/ECTA10712>
- Ramsey, F. P. (1927). A contribution to the theory of taxation. *The Economic Journal*, 37(145), 47–61. <https://doi.org/10.2307/2222721>
- Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics*, 83(2), 217–230. [https://doi.org/10.1016/S0047-2727\(00\)00159-6](https://doi.org/10.1016/S0047-2727(00)00159-6)
- Saez, E., & Zucman, G. (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 131(2), 519–578. <https://doi.org/10.1093/qje/qjw004>
- Scheuer, F., & Werning, I. (2017). The taxation of superstars. *The Quarterly Journal of Economics*, 132(1), 211–270. <https://doi.org/10.1093/qje/qjw036>
- Stiglitz, J. E. (1975). The effects of income, wealth, and capital gains taxation on risk-taking. In *Stochastic optimization models in finance* (pp. 291–311). Elsevier.
- Stiglitz, J. E. (2018). Pareto efficient taxation and expenditures: Pre-and re-distribution. *Journal of Public Economics*, 162, 101–119. <https://doi.org/10.1016/j.jpubeco.2018.01.006>
- Varian, H. R. (1980). Redistributive taxation as social insurance. *Journal of Public Economics*, 14(1), 49–68. [https://doi.org/10.1016/0047-2727\(80\)90004-3](https://doi.org/10.1016/0047-2727(80)90004-3)



## APPENDIX A: PROOF OF EQUATION (6)

To prove the existence of the utility-neutralizing labor income tax  $\widehat{T}(\cdot)$ , define the auxiliary function  $W : \mathbb{R}^2 \mapsto \mathbb{R}$  as:

$$W(y, a) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[v(y - a + (1 + r)s - \widehat{t}(s, rs))|s] \quad (\text{A1})$$

If  $\widehat{t}(s, rs)$  verifies Assumption 1, the first-order condition associated to problem (A1) is differentiable. Denoting by  $\widehat{s}(y, a)$  the solution of (A1), I know from the implicit function theorem that  $\widehat{s}(y, a)$  is continuously differentiable in  $a$  and  $y$ . Hence the value function  $W(y, a)$  is continuously differentiable in  $a$ . We can therefore apply the envelope theorem to (A1) to establish:

$$\frac{\partial W(y, a)}{\partial a} = -\mathbb{E}[v'(y - a + (1 + r)\widehat{s}(y, a) - \widehat{t}(\widehat{s}(y, a), r\widehat{s}(y, a)))|\widehat{s}(y, a)] < 0$$

Hence for every  $y$ , the equation  $W(y, a) = V^0(y)$  admits a single solution that I denote  $a(y)$ . Defining the utility-neutralizing labor income tax function as  $\widehat{T} : y \mapsto a(y)$  therefore implies for all  $y$ :

$$V^0(y) = \max_s u(A - s) + \mathbb{E}[v(y - \widehat{T}(y) + (1 + r)s - \widehat{t}(s, rs))|s]$$

## APPENDIX B: PROOF OF EQUATION (11)

The Lagrangian associated to (10) is:

$$L \equiv (1 + \bar{r}(s))s - \int_{r \in \mathbb{R}} c_2(r)f(r|s)dr + \lambda \left( u(A - s) + \int_{r \in \mathbb{R}} v(c_2(r))f(r|s)dr - V^0(y(\theta)) \right) \quad (\text{B1})$$

The first-order condition with respect to  $c_2(r)$  yields:

$$\lambda v'(c_2(r)) = 1 \quad (\text{B2})$$

$v'$  is monotonic so one can rewrite (B2) as:

$$c_2(r) = (v')^{-1}\left(\frac{1}{\lambda}\right) \quad (\text{B3})$$

which implies that  $c_2(r) = c_2$  is no longer a function of  $r$  at the optimum.

The F.O.C with respect to  $s$  is:

$$1 + \bar{r}(s) + \bar{r}'(s)s = \frac{\partial}{\partial s} \left( \int_{r \in \mathbb{R}} c_2(r) f(rls) dr \right) - \lambda \left[ \frac{\partial}{\partial s} \left( \int_{r \in \mathbb{R}} v(c_2(r)) f(rls) dr \right) - u'(A-s) \right] \quad (\text{B4})$$

Plugging (B2) in (B4) and using  $c_2(r) = c_2$  at the optimum yields:

$$1 + \bar{r}(s) + \bar{r}'(s)s = \frac{\partial}{\partial s} \left( \int_{r \in \mathbb{R}} f(rls) dr \right) c_2 - \frac{1}{v'(c_2)} \left[ \frac{\partial}{\partial s} \left( \int_{r \in \mathbb{R}} f(rls) dr \right) v(c_2) - u'(A-s) \right] \quad (\text{B5})$$

By definition of  $f(\cdot)$  as a density function,  $\int_{r \in \mathbb{R}} f(rls) dr = 1$ . So eventually the integral terms drop hence the Euler Equation:

$$u'(A-s) = (1 + \bar{r}(s) + \bar{r}'(s)s) v'(c_2)$$

## APPENDIX C: PROOF OF PROPOSITION 2

From Lemma 1, for any initial tax on savings  $t^0(s)$  and any candidate  $\hat{t}(s)$ , there exists a reformed labor income tax  $\hat{T}(y)$  such that utility and labor income is the same under  $\{\hat{t}, \hat{T}\}$  than under  $\{t^0, T^0\}$ . Then a necessary and sufficient condition for the optimality of  $\hat{t}$  is given by:

$$\hat{T}(y^0(\theta)) + \hat{t}(\hat{s}(\theta)) \geq T^0(y^0(\theta)) + t^0(s^0(\theta)) \quad (\text{C1})$$

Let  $\alpha = y - T(y) + s - t(s)$  denote the deterministic part of second-period consumption. Then one can rewrite (C1) as:

$$\hat{s}(\theta) - \hat{\alpha}(\theta) \geq s^0(\theta) - \alpha^0(\theta) \quad (\text{C2})$$

with  $\hat{\alpha}(\theta)$  and  $\alpha^0(\theta)$  denoting the deterministic part of second-period consumption under, respectively  $\{\hat{t}, \hat{T}\}$  and  $\{t^0, T^0\}$ . Then (C2) is verified when  $\{\hat{s}(\theta), \hat{\alpha}(\theta)\}$  solves:

$$\begin{aligned} & \max_{s, \alpha} s - \alpha \\ & \text{subject to: } u(A-s) + \int_{r \in \mathbb{R}} v(\alpha + rs) f(rls) dr = V^0(y(\theta)) \end{aligned} \quad (\text{C3})$$

A necessary and sufficient condition for solving (C3) is given by:

$$u'(A-s) = \int_{r \in \mathbb{R}} (1+r) v'(\alpha + rs) f(rls) dr + \int_{r \in \mathbb{R}} v(\alpha + rs) f'_s(rls) dr \quad (\text{C4})$$

Under  $\{\hat{t}, \hat{T}\}$ , an agent with labor income  $y$  solves:

$$V^0(y) = \max_s u(A-s) + \mathbb{E}[v(y + (1+r)s - \hat{T}(y) - \hat{t}(s)) | s] \quad (\text{C5})$$

The F.O.C of (C5) yields:

$$u'(A - s) = \int_{r \in \mathbb{R}(1+r-\hat{t}'(s))} v'(\alpha + rs)f(r|s)dr + \int_{r \in \mathbb{R}} v(\alpha + rs)f'_s(r|s)dr \quad (C6)$$

Hence for  $\{\hat{s}(\theta), \hat{\alpha}(\theta)\}$  to solve (C3),  $\hat{t}$  should be such that (C6) matches with (C4). This is only the case for  $\hat{t}(s) = \text{constant}$ , hence this is in particular true for  $\hat{t}(s) = 0$ .

#### APPENDIX D: PROOF OF PROPOSITION 3

From Lemma 1, for any initial tax on savings  $t^0(r(s)s)$  and any candidate  $\hat{t}(r(s)s)$ , there exists a reformed labor income tax  $\hat{T}(y)$  such that utility and labor income is the same under  $\{\hat{t}, \hat{T}\}$  than under  $\{t^0, T^0\}$ . Then a sufficient condition for the optimality of  $\hat{t}$  is given by:

$$\hat{T}(y^0(\theta)) + \hat{t}(r(\hat{s}(\theta))\hat{s}(\theta)) \geq T^0(y^0(\theta)) + t^0(r(s^0(\theta))s^0(\theta)) \quad (D1)$$

Let  $\beta = y - T(y) + s - t(r(s)s)$  Then one can rewrite (D1) as:

$$\hat{s}(\theta) - \hat{\beta}(\theta) \geq s^0(\theta) - \beta^0(\theta) \quad (D2)$$

Then (D2) is verified when  $\{\hat{s}, \hat{\beta}\}$  solves:

$$\begin{aligned} & \max_{s, \beta} s - \beta \\ & \text{subject to: } u(A - s) + v(\beta + r(s)s) = V^0(y(\theta)) \end{aligned} \quad (D3)$$

A necessary and sufficient condition for solving (D3) is given by:

$$u'(A - s) = (1 + r(s) + r'(s)s)v'(\beta + r(s)s) \quad (D4)$$

From Lemma 1, for any  $t^0(r(s)s)$  and for any candidate  $\hat{t}(r(s)s)$ , there exists a labor income tax function  $\hat{T}(y)$  such that:

$$V^0(y) = \max_s u(A - s) + v(y - \hat{T}(y) + (1 + r(s))s - \hat{t}(r(s)s)) \quad (D5)$$

The F.O.C of (D5) yields:

$$u'(A - s) = [1 + r(s) + r'(s)s - (r(s) + r'(s)s)\hat{t}'(\cdot)]v'(y - \hat{T}(y) + (1 + r(s))s - \hat{t}(r(s)s)) \quad (D6)$$

Hence for  $\{\hat{s}(\theta), \hat{\beta}(\theta)\}$  to solve (D3),  $\hat{t}$  should be such that (D6) matches with (D4). This is only the case for  $\hat{t}(r(s)s) = \text{constant}$ , hence this is in particular true for  $\hat{t}(r(s)s) = 0$ .

## APPENDIX E: PROOF OF THE RESULTS IN SECTION 4.2

### E.1. Proof of Equation (14)

In the economy without capital taxation, an agent with labor income  $y$  solves:

$$V^0(y) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[v(y - T^0(y) + (1 + r)s)|s] \quad (\text{E1})$$

Let  $s^0$  be the solution of (E1). Now introduce a linear tax  $\tau$  on capital income  $rs$ , so that  $\hat{i}(rs) = \tau \cdot rs$ . From Lemma 1, I know that for every tax rate  $\tau$ , there exists a corresponding labor income tax schedule  $\hat{T}(y, \tau)$  so that the reform does not affect the utility of taxpayers.

Under  $\{\hat{T}(\cdot); \tau\}$  an agent with income  $y$  must solve:

$$\hat{V}(y, \tau) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[v(y - \hat{T}(y, \tau) + (1 + (1 - \tau)r)s)|s] \quad (\text{E2})$$

Denote by  $\hat{s}(\tau)$  the solution of (E2). Under the candidate  $\{\hat{T}(\cdot); \tau\}$  the expected fiscal contribution of a taxpayer with labor income  $y$  is given by:

$$\mathcal{B}(y, \tau) = \hat{T}(y, \tau) + \tau \cdot \bar{r}(\hat{s}(\tau)) \cdot \hat{s}(\tau)$$

Note that  $\hat{T}(y, 0) = T^0(y)$  and  $\hat{s}(0) = s^0$ . Hence the impact on government revenue, evaluated at  $\tau = 0$ , of introducing linear capital income taxation is equal to:

$$\left. \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} \right|_{\tau=0} = \left. \frac{\partial \hat{T}(y, \tau)}{\partial \tau} \right|_{\tau=0} + \bar{r}(s^0)s^0 \quad (\text{E3})$$

Taxing capital income at rate  $\tau$  has an a priori negative effect on individual utility since it reduces the expected return on savings. As established in Lemma 1, this negative impact can always be compensated by adjusting labor income taxation. Hence the government should decrease labor income taxes to mute the impact of  $\tau$  on utility. To compute by how much should labor income taxation be reduced to keep utility unaffected by  $\tau$ , one can use the definition of  $\hat{T}(y, \tau)$  that guarantees for all  $\tau$ :

$$\begin{aligned} \hat{V}(y, \tau) &= V^0(y) \\ \Rightarrow \frac{\partial \hat{V}(y, \tau)}{\partial \tau} &= \frac{\partial V^0(y)}{\partial \tau} = 0 \end{aligned} \quad (\text{E4})$$

Applying the envelope theorem to (E2) and using (E4) yields:

$$\left. \frac{\partial \hat{T}(y, \tau)}{\partial \tau} \right|_{\tau=0} = -s^0 \frac{\mathbb{E}[rv'(y^0 - T^0(y) + (1 + r)s^0)|s^0]}{\mathbb{E}[v'(y^0 - T^0(y) + (1 + r)s^0)|s^0]} \quad (\text{E5})$$

Since  $v(\cdot)$  is strictly increasing, the right-hand side of equation (E5) is strictly negative: labor income taxation has to decrease to keep utility unchanged after capital income taxation is introduced. The  $\left. \frac{\partial \hat{T}(y, \tau)}{\partial \tau} \right|_{\tau=0}$  term in equation (E3) is therefore negative. The question is now to

compare this loss in labor income tax revenue with the gains associated to capital income taxation.

Plugging (E5) in (E3) gives:

$$\begin{aligned}
\left. \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} \right|_{\tau=0} &= s^0 \left( \bar{r}(s^0) - \frac{\mathbb{E}[rv'(\cdot) | s^0]}{\mathbb{E}[v'(\cdot) | s^0]} \right) \\
&= s^0 \left( \frac{\bar{r}(s^0) \mathbb{E}[v'(\cdot) | s^0] - \mathbb{E}[rv'(\cdot) | s^0]}{\mathbb{E}[v'(\cdot) | s^0]} \right) \\
&= -s^0 \left( \frac{\text{Cov}[r; v'(y^0 - T^0(y) + (1+r)s^0) | s^0]}{\mathbb{E}[v'(y^0 - T^0(y) + (1+r)s^0) | s^0]} \right)
\end{aligned} \tag{E6}$$

## E.2. The case of a tax on ex post wealth $(1+r)s$

Start from an economy where there is no tax on  $(1+r)s$  but only a labor income tax  $T^0(y)$ . Then an agent with labor income  $y$  solves:

$$V^0(y) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[v(y - T^0(y) + (1+r)s) | s] \tag{E7}$$

Denote by  $s^0$  the solution of (E7).

Now introduce a linear tax  $\tau$  on  $(1+r)s$ . Again Lemma 1 guarantees the existence of a corresponding labor income tax schedule  $\hat{T}(y, \tau)$  so that the reform does not affect the utility of taxpayers.

Under  $\{\hat{T}(\cdot); \tau\}$  an agent with income  $y$  must solve:

$$\hat{V}(y, \tau) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[v(y - \hat{T}(y, \tau) + (1 - \tau)(1+r)s -) | s] \tag{E8}$$

Denote by  $\hat{s}(\tau)$  the solution of (E8). Under the candidate  $\{\hat{T}(\cdot); \tau\}$  the expected fiscal contribution of a taxpayer with labor income  $y$  is given by:

$$\mathcal{B}(y, \tau) = \hat{T}(y, \tau) + \tau \cdot (1 + \bar{r}(\hat{s}(\tau))) \cdot \hat{s}(\tau)$$

so that the impact of implementing a tax with rate  $\tau$  on  $(1+r)s$  verifies:

$$\left. \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} \right|_{\tau=0} = \left. \frac{\partial \hat{T}(y, \tau)}{\partial \tau} \right|_{\tau=0} + \bar{r}(s^0) s^0 \tag{E9}$$

It follows from the definition of  $\hat{T}(y, \tau)$  that:

$$u(A - \hat{s}(\tau)) + \mathbb{E}[v(y - \hat{T}(y, \tau) + (1 - \tau)(1+r)\hat{s}(\tau)) | \hat{s}(\tau)] = V^0(y) \tag{E10}$$

Applying the envelope theorem to (E8) and using (E10) yields:

$$\frac{\partial \hat{T}(y, \tau)}{\partial \tau} = -\hat{s}(\tau) \frac{\mathbb{E}[(1+r)v'(\cdot) | \hat{s}(\tau)]}{\mathbb{E}[v'(\cdot) | \hat{s}(\tau)]} \tag{E11}$$

Evaluating (E11) at  $\tau = 0$  before plugging it in (E9) gives:

$$\begin{aligned}
\left. \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} \right|_{\tau=0} &= s^0 \left( 1 + \bar{r}(s^0) - \frac{\mathbb{E}[(1+r)v'(\cdot) | s^0]}{\mathbb{E}[v'(\cdot) | s^0]} \right) \\
&= s^0 \left( \frac{(1 + \bar{r}(s^0))\mathbb{E}[v'(\cdot) | s^0] - \mathbb{E}[(1+r)v'(\cdot) | s^0]}{\mathbb{E}[v'(\cdot) | s^0]} \right) \\
&= s^0 \left( \frac{\bar{r}(s^0)\mathbb{E}[v'(\cdot) | s^0] - \mathbb{E}[rv'(\cdot) | s^0]}{\mathbb{E}[v'(\cdot) | s^0]} \right) \\
&= -s^0 \left( \frac{\text{Cov}[r; v'(y^0 - T^0(y) + (1+r)s^0) | s^0]}{\mathbb{E}[v'(y^0 - T^0(y) + (1+r)s^0) | s^0]} \right)
\end{aligned} \tag{E12}$$

The concavity of  $v(\cdot)$  guarantees that the covariance term in (E12) is negative so that  $\left. \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} \right|_{\tau=0} > 0$ .

### E.3. Proof of Equation (15)

In the economy without capital taxation, an agent with labor income  $y$  solves:

$$V^0(y) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[y - T^0(y) + v((1+r)s)|s] \tag{E13}$$

Now introduce a linear tax  $\tau$  on capital income  $rs$ , so that  $\hat{t}(rs) = \tau \cdot rs$ . From Lemma 1, I know that for every tax rate  $\tau$ , there exists a corresponding labor income tax schedule  $\hat{T}(y, \tau)$  so that the reform does not affect the utility of taxpayers.

Under  $\{\hat{T}(\cdot); \tau\}$  an agent with labor income  $y$  must solve:

$$\hat{V}(y, \tau) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[v(y - \hat{T}(y, \tau) + (1 + (1 - \tau)r)s)|s] \tag{E14}$$

Denote by  $\hat{s}(\tau)$  the solution of (E14). Under the candidate  $\{\hat{T}(\cdot); \tau\}$  the expected fiscal contribution of a taxpayer with labor income  $y$  is given by:

$$\mathcal{B}(y, \tau) = \hat{T}(y, \tau) + \tau \cdot \bar{r}(\hat{s}(\tau)) \cdot \hat{s}(\tau)$$

The impact on government revenue of introducing linear capital income taxation is equal to:

$$\frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} = \frac{\partial \hat{T}(y, \tau)}{\partial \tau} + \bar{r}(\hat{s}(\tau)) \cdot \hat{s}(\tau) + \tau[\hat{s}'(\tau) \cdot \bar{r}'(\hat{s}(\tau)) \cdot \hat{s}(\tau) + \hat{s}'(\tau) \cdot \bar{r}(\hat{s}(\tau))] \tag{E15}$$

The definition of  $\hat{T}(y, \tau)$  implies:

$$\begin{aligned}
\hat{V}(y, \tau) &= V^0(y) \\
\Rightarrow \frac{\partial \hat{V}(y, \tau)}{\partial \tau} &= \frac{\partial V^0(y)}{\partial \tau} = 0
\end{aligned} \tag{E16}$$

Applying the envelope theorem to (E14) and using (E16) yields:

$$\frac{\partial \widehat{T}(y, \tau)}{\partial \tau} = -\widehat{s}(\tau) \frac{\mathbb{E}[rv'(\cdot) | \widehat{s}(\tau)]}{\mathbb{E}[v'(\cdot) | \widehat{s}(\tau)]} \quad (\text{E17})$$

Plugging (E17) in (E15) gives:

$$\begin{aligned} \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} &= -\widehat{s}(\tau) \frac{\mathbb{E}[rv'(\cdot) | \widehat{s}(\tau)]}{\mathbb{E}[v'(\cdot) | \widehat{s}(\tau)]} + \bar{r}(\widehat{s}(\tau)) \cdot \widehat{s}(\tau) + \tau[\widehat{s}'(\tau) \cdot \bar{r}'(\widehat{s}(\tau)) \cdot \widehat{s}(\tau) \\ &\quad + \widehat{s}'(\tau) \cdot \bar{r}(\widehat{s}(\tau))] \\ &= -\widehat{s}(\tau) \left( \frac{\text{Cov}[rv'(\cdot) | \widehat{s}(\tau)]}{\mathbb{E}[v'(\cdot) | \widehat{s}(\tau)]} \right) - \tau \frac{\partial \widehat{s}(\tau)}{\partial (1-\tau)} (\bar{r}'(\widehat{s}) \cdot \widehat{s}(\tau) + \bar{r}(\widehat{s})) \\ &= -\widehat{s}(\tau) \left( \frac{\text{Cov}[rv'(\cdot) | \widehat{s}(\tau)]}{\mathbb{E}[v'(\cdot) | \widehat{s}(\tau)]} \right) - \frac{\tau}{1-\tau} \cdot \widehat{s}(\tau) \cdot \bar{r}(s) \cdot \epsilon_{1-\tau}^s \cdot (1 + \epsilon_s^{\bar{r}}) \end{aligned} \quad (\text{E18})$$

At the optimum we should have  $\frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} = 0$ . It therefore follows from (E18) that:

$$\begin{aligned} \frac{\tau}{1-\tau} \cdot \widehat{s}(\tau) \cdot \bar{r}(s) \cdot \epsilon_{1-\tau}^s \cdot (1 + \epsilon_s^{\bar{r}}) &= -\widehat{s}(\tau) \left( \frac{\text{Cov}[rv'(\cdot) | \widehat{s}(\tau)]}{\mathbb{E}[v'(\cdot) | \widehat{s}(\tau)]} \right) \\ \frac{\tau}{1-\tau} &= - \left( \frac{\text{Cov}[rv'(\cdot) | \widehat{s}(\tau)]}{\bar{r}(s) \mathbb{E}[v'(\cdot) | \widehat{s}(\tau)]} \right) \cdot \frac{1}{\epsilon_{1-\tau}^s} \cdot \frac{1}{1 + \epsilon_s^{\bar{r}}} \end{aligned}$$

## APPENDIX F: PROOFS OF THE RESULTS OF SECTION 5

### F.1. Proof of Lemma 2

Consider an initial tax schedule  $\{T^0, t^0\}$  with  $t^0$  verifying Assumption 2. Under  $\{T^0, t^0\}$ , an agent with labor income  $y$  enjoys subutility from consumption:

$$\begin{aligned} V^0(y) &\stackrel{\text{def}}{=} \max_{b, z} u(A - b - z) + \mathbb{E}[v(y - T^0 + (1 + r(b))b \\ &\quad + (1 + x)z - t^0(b, rb, z, xz)) | z] \end{aligned} \quad (\text{F1})$$

Now suppose that the government moves from the initial capital tax  $t^0$  to a candidate  $\hat{t}$ , verifying Assumption 2. Consider the function  $W : \mathbb{R}^2 \mapsto \mathbb{R}$  defined as:

$$\begin{aligned} W(y, a) &\stackrel{\text{def}}{=} \max_{b, z} u(A - b - z) + \mathbb{E}[v(y - a + (1 + r(b))b \\ &\quad + (1 + x)z - \hat{t}(b, rb, z, xz)) | z] \end{aligned} \quad (\text{F2})$$

Because  $\hat{t}$  verifies Assumption 2, the first-order condition associated to problem (F2) is differentiable. Denoting by  $\{\hat{b}(y, a), \hat{z}(y, a)\}$  the solution of (F2), I know from the implicit function theorem that  $\{\hat{b}(y, a), \hat{z}(y, a)\}$  is continuously differentiable in  $a$ . Hence the value function  $W(y, a)$  is continuously differentiable in  $a$ . We can therefore apply the envelope theorem to (F2) to establish:

$$\frac{\partial W(y, a)}{\partial a} = -\mathbb{E}[v'(y - a + (1 + r(b)))b + (1 + x)z - \hat{t}(b, rb, z, xz)|z] < 0$$

Hence for every  $y$ , the equation  $W(y, a) = V^0(y)$  admits a single solution that I denote  $a(y)$ . Defining the utility-neutralizing labor income tax function as  $\hat{T} : y \mapsto a(y)$  therefore implies for all  $y$ :

$$V^0(y) = \max_{b,z} u(A - b - z) + \mathbb{E}[v(y - \hat{T} + (1 + r(b))b + (1 + x)z - \hat{t}(b, rb, z, xz)|z)] \quad (\text{F3})$$

Hence for every labor income  $y$ , subutility from consumption is the same under  $\{T^0, t^0\}$  than under  $\{\hat{T}, \hat{t}\}$ . Denoting by  $\hat{U}(\theta)$  the total utility of an agent with type  $\theta$  under the tax schedule  $\{\hat{T}, \hat{t}\}$ , it follows from (F3) that:

$$\begin{aligned} \hat{U}(\theta) &= \max_y V^0(y) - h(y, \theta) \\ &= U^0(\theta) \end{aligned} \quad (\text{F4})$$

It follows from (F4) that total utility is the same under  $\{\hat{T}(\cdot), \hat{t}(\cdot)\}$  than under  $\{T^0, t^0\}$  and that  $y^0(\theta)$  solves the labor supply problem under both tax regimes, for every type  $\theta$ .

## F.2. Proof of Proposition 6

For any initial tax schedule  $\{T^0, t^0\}$ ,  $\hat{t}$  is optimal if and only if:

$$\begin{aligned} &\hat{T}(y^0(\theta)) + \mathbb{E}[\hat{t}(\hat{b}(\theta), r\hat{b}(\theta), \hat{z}(\theta), x\hat{z}(\theta))|\hat{z}(\theta)] \\ &\geq T^0(y^0(\theta)) + \mathbb{E}[t^0(b^0(\theta), rb^0(\theta), z^0(\theta), xz^0(\theta))|z^0(\theta)] \end{aligned} \quad (\text{F5})$$

with  $\{b^0(\theta), z^0(\theta)\}$  and  $\{\hat{b}(\theta), \hat{z}(\theta)\}$  the optimal portfolio choices respectively under  $\{t^0, T^0\}$  and under  $\{\hat{t}, \hat{T}\}$ . Using the second-period budget constraint of an agent with type  $\theta$  one can rewrite (F5) as:

$$\begin{aligned} &(1 + r(\hat{b}(\theta)))\hat{b}(\theta) + (1 + \bar{x}(\hat{z}(\theta)))\hat{z}(\theta) - \mathbb{E}[\hat{c}_2(\theta, x)|\hat{z}(\theta)] \\ &\geq (1 + r(b^0(\theta)))b^0(\theta)(1 + \bar{x}(z^0(\theta)))z^0(\theta) - \mathbb{E}[c_2^0(\theta, x)|z^0(\theta)] \end{aligned} \quad (\text{F6})$$

with  $c_2^0(\theta, x)$  and  $\hat{c}_2(\theta, x)$  denoting second-period consumption for a given  $\theta$  and a given  $x$  respectively under  $\{t^0, T^0\}$  and  $\{\hat{t}, \hat{T}\}$ .

Hence the condition for optimality (F6) is verified if  $\{\hat{z}(\theta); \hat{b}(\theta); \hat{c}_2(\theta, x)\}$  is the solution of:

$$\begin{aligned} &\max_{b,z,c_2(x)} (1 + r(b))b + (1 + \bar{x}(z))z - \int_{x \in \mathbb{R}} c_2(x)f(x|z)dx \\ &\text{subject to: } u(A - b - z) + \int_{x \in \mathbb{R}} v(c_2(x))f(x|z)dx = V^0(y(\theta)) \end{aligned} \quad (\text{F7})$$



The Lagrangian associated to (F7) is:

$$L \equiv (1 + r(b))b + (1 + \bar{x}(z))z - \int_{x \in \mathbb{R}} c_2(x)f(x|z)dx + \lambda \left( u(A - b - z) + \int_{x \in \mathbb{R}} v(c_2(x))f(x|z)dx - V^0(y) \right) \quad (\text{F8})$$

The FOC with respect to  $c_2(x)$  yields:

$$\lambda v'(c_2(x)) = 1 \quad (\text{F9})$$

$v'(\cdot)$  is monotonic so one can rewrite (F9) as:

$$c_2(x) = (v')^{-1}\left(\frac{1}{\lambda}\right) \quad (\text{F10})$$

which implies that  $c_2(x) = c_2$  is no longer a function of  $x$  at the optimum.

The FOC with respect to  $b$  yields:

$$\lambda u'(A - b - z) = 1 + r(b) + r'(b)b$$

Using (F9) this yields:

$$u'(A - b - z) = (1 + r(b) + r'(b)b)v'(c_2)$$

Eventually the FOC with respect to  $z$  yields:

$$\lambda \left( u'(A - b - z) - \frac{\partial}{\partial z} \left( \int_{x \in \mathbb{R}} v(c_2(x))f(x|z)dr \right) \right) = 1 + \bar{x}(z) + \bar{x}'(z)z - \frac{\partial}{\partial z} \left( \int_{x \in \mathbb{R}} c_2(x)f(x|z)dr \right) \quad (\text{F11})$$

Using  $c_2(x) = c_2$  at the optimum, one can rewrite (F11) as:

$$\lambda u'(A - b - z) = 1 + \bar{x}(z) + \bar{x}'(z)z \quad (\text{F12})$$

And using (F9) this directly yields:

$$u'(A - b - z) = v'(c_2)(1 + \bar{x}(z) + \bar{x}'(z)z) \quad (\text{F13})$$

Hence a triplet  $\{b, z, c_2(x)\}$  solves problem (F7) iff:

$$\begin{cases} c_2(x) = c_2 \\ u'(A - b - z) = (1 + r(b) + r'(b)b)v'(c_2) \\ u'(A - b - z) = (1 + \bar{x}(z) + \bar{x}'(z)z)v'(c_2) \end{cases} \quad (\text{F14})$$

Now consider the tax candidate  $\hat{t}$  defined in (18). Under this specific  $\hat{t}$ , second-period consumption is given by  $c_2(x) = y - \hat{T}(y) + (1 + r(b))b + (1 + \bar{x}(z))z = c_2$  so that the first optimality condition of (F14) is verified.

The problem of an agent with labor  $y$  under  $\{\hat{T}(\cdot), \hat{t}\}$  is given by:

$$V^0(y) = \max_{b,z} u(A - b - z) + v(y - \hat{T}(y) + (1 + r(b))b + (1 + \bar{x}(z))z) \quad (\text{F15})$$

The FOC of (F15) implies that  $\{\hat{z}(\theta), \hat{b}(\theta), \hat{c}_2(\theta)\}$  verifies

$$u'(A - \hat{b}(\theta) - \hat{z}(\theta)) = (1 + r(\hat{b}(\theta)) + r'(\hat{b}(\theta))\hat{b}(\theta))v'(\hat{c}_2(\theta)) \quad (\text{F16})$$

$$u'(A - \hat{b}(\theta) - \hat{z}(\theta)) = (1 + \bar{x}(\hat{z}(\theta)) + \bar{x}'(\hat{z}(\theta))\hat{z}(\theta))v'(\hat{c}_2(\theta)) \quad (\text{F17})$$

Hence the three optimality conditions defined in (F14) are verified so  $\{\hat{z}(\theta), \hat{b}(\theta), \hat{c}_2(\theta)\}$  solves problem (F7). This proves Proposition 6.

### F.3. Proof of proposition 7

From Lemma 2, for any initial tax  $t^0(b, rb, z)$  and any candidate  $\hat{t}(b, rb, z)$ , there exists a reformed labor income tax  $\hat{T}(y)$  such that utility and labor income is the same under  $\{\hat{t}, \hat{T}\}$  than under  $\{t^0, T^0\}$ . Then  $\hat{t}$  is optimal iff:

$$\hat{T}(y^0(\theta)) + \hat{t}(\hat{b}(\theta), r(\hat{b}(\theta)), \hat{z}(\theta)) \geq T^0(y^0(\theta)) + t^0(b^0(\theta), r(b^0(\theta)), z^0(\theta)) \quad (\text{F18})$$

Then one can rewrite (F18) as:

$$(1 + r(\hat{b}(\theta))\hat{b}(\theta) + \hat{z}(\theta) - \hat{\alpha}(\theta)) \geq (1 + r(b^0(\theta))b^0(\theta) + z^0(\theta) - \alpha^0(\theta)) \quad (\text{F19})$$

where  $\alpha = y - T(\cdot) + (1 + r(b))b + z - t(\cdot)$  denote the deterministic component of second period consumption. Then (F19) is verified when  $\{\hat{b}, \hat{z}, \hat{\alpha}\}$  solves:

$$\begin{aligned} & \max_{b,z,\alpha} (1 + r(b))b + z - \alpha \\ & \text{subject to: } u(A - b - z) + \int_{x \in \mathbb{R}} v(\alpha + xz)f(x|z)dx = V^0(y^0(\theta)) \end{aligned} \quad (\text{F20})$$

The Lagrangian associated to (F20) is:

$$L \equiv (1 + r(b))b + z - \alpha + \lambda \left( u(A - b - z) + \int_{x \in \mathbb{R}} v(\alpha + xz)f(x|z)dx - V^0(y^0(\theta)) \right) \quad (\text{F21})$$

The FOC with respect to  $\alpha$  yields:

$$1 = \lambda \int_{x \in \mathbb{R}} v'(\alpha + xz)f(x|z)dx \quad (\text{F22})$$

The FOC with respect to  $z$  yields:

$$1 = \lambda \left[ u'(A - b - z) - \int_{x \in \mathbb{R}} xv'(\alpha + xz)f(x|z)dx - \int_{x \in \mathbb{R}} v(\alpha + xz)f'_z(x|z)dx \right]$$

Combining these two conditions yields:

$$u'(A - b - z) = \mathbb{E}[(1 + x)v'(\alpha + xz)|z] + \int_{x \in \mathbb{R}} v(\alpha + xz)f'_z(x|z)dx$$

Eventually, the FOC associated to  $b$  gives:

$$u'(A - b - z) = \frac{1}{\lambda}(1 + r(b) + r'(b)b)$$

so that combining the FOCs associated to  $\alpha$  and  $b$  yields:

$$u'(A - b - z) = (1 + r(b) + r'(b)b)\mathbb{E}[v'(\alpha + xz)|z]$$

A necessary and sufficient condition for solving (F20) is given by:

$$\begin{cases} u'(A - b - z) = \mathbb{E}[(1 + x)v'(\alpha + xz)|z] + \int_{x \in \mathbb{R}} v(\alpha + xz)f'_z(x|z)dx \\ u'(A - b - z) = (1 + r(b) + r'(b)b)\mathbb{E}[v'(\alpha + xz)|z] \end{cases} \quad (\text{F23})$$

Under  $\{\hat{T}, \hat{t}\}$  a taxpayer with labor income  $y$  solves:

$$V^0(y) = \max_{b,z} u(A - b - z) + \mathbb{E}[v(y - \hat{T}(y) + (1 + r(b))b + (1 + x)z - \hat{t}(b, rb, z))|z] \quad (\text{F24})$$

The F.O.C with respect to  $b$  is:

$$u'(A - b - z) = \left( 1 + r(b) + r'(b)b - \frac{d\hat{t}}{db} \right) \mathbb{E}[v'(\cdot)|z]$$

while the F.O.C with respect to  $z$  is:

$$u'(A - b - z) = \mathbb{E} \left[ \left( 1 + x - \frac{d\hat{t}}{dz} \right) v'(\alpha + xz) | z \right] + \int_{x \in \mathbb{R}} v(\alpha + xz)f'_z(x|z)dx$$

Hence the FOCs of the taxpayer's problem (F24) coincide with the FOCs of the government revenue maximization problem (F20) whenever  $\frac{d\hat{t}}{db} = \frac{d\hat{t}}{dz} = 0$ . This is the case for any  $\hat{t}(b, rb, z) = \text{constant}$ , and this is in particular the case for  $\hat{t}(b, rb, z) = 0$ .

#### F.4. Proof of Proposition 8

In the economy without capital taxation, an agent with labor income  $y$  solves:

$$V^0(y) \stackrel{\text{def}}{=} \max_{b,z} u(A - b - z) + \mathbb{E}[v(y - T^0(y) + (1 + r(b))b + (1 + x)z)|z] \quad (\text{F25})$$

Let  $z^0$  be the risky asset level solving (F25). Now introduce a linear tax  $\tau$  on  $xz$ . From Lemma 2, I know that for every tax rate  $\tau$ , there exists a corresponding labor income tax schedule  $\hat{T}(y, \tau)$  so that the reform does not affect the utility of taxpayers.

Under  $\{\hat{T}(\cdot); \tau\}$  an agent with income  $y$  must solve:

$$\hat{V}(y, \tau) \stackrel{\text{def}}{=} \max_{b,z} u(A - b - z) + \mathbb{E}[v(y - \hat{T}(y, \tau) + (1 + r(b))b + (1 + (1 - \tau)x)z)|z] \quad (\text{F26})$$

Denote by  $\hat{z}(\tau)$  the risky asset investment solving (F26). Under the candidate  $\{\hat{T}(\cdot); \tau\}$  the expected fiscal contribution of a taxpayer with labor income  $y$  is given by:

$$\mathcal{B}(y, \tau) = \hat{T}(y, \tau) + \tau \cdot \bar{x}(\hat{z}(\tau)) \cdot \hat{z}(\tau)$$

Note that  $\hat{T}(y, 0) = T^0(y)$  and  $\hat{z}(0) = z^0$ . Hence the impact on government revenue, evaluated at  $\tau = 0$ , of introducing linear capital income taxation is equal to:

$$\left. \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} \right|_{\tau=0} = \left. \frac{\partial \hat{T}(y, \tau)}{\partial \tau} \right|_{\tau=0} + \bar{x}(z^0)z^0 \quad (\text{F27})$$

To compute by how much should labor income taxation be reduced to keep utility unaffected by  $\tau$ , one can use the definition of  $\hat{T}(y, \tau)$  that guarantees for all  $\tau$ :

$$\begin{aligned} \hat{V}(y, \tau) &= V^0(y) \\ \Rightarrow \frac{\partial \hat{V}(y, \tau)}{\partial \tau} &= \frac{\partial V^0(y)}{\partial \tau} = 0 \end{aligned} \quad (\text{F28})$$

Applying the envelope theorem to (F26) and using (F28) yields:

$$\left. \frac{\partial \hat{T}(y, \tau)}{\partial \tau} \right|_{\tau=0} = -z^0 \frac{\mathbb{E}[xv'(\cdot)|z^0]}{\mathbb{E}[v'(\cdot)|z^0]} \quad (\text{F29})$$

Plugging (F29) in (F27) gives:

$$\begin{aligned} \left. \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} \right|_{\tau=0} &= z^0 \left( \bar{x}(z^0) - \frac{\mathbb{E}[xv'(\cdot)|z^0]}{\mathbb{E}[v'(\cdot)|z^0]} \right) \\ &= z^0 \left( \frac{\bar{x}(z^0)\mathbb{E}[v'(\cdot)|z^0] - \mathbb{E}[xv'(\cdot)|z^0]}{\mathbb{E}[v'(\cdot)|z^0]} \right) \\ &= -z^0 \left( \frac{\text{Cov}[x; v'(y - T^0(y) + (1 + r(b^0))b^0 + (1 + x)z^0) | z^0]}{\mathbb{E}[v'(\cdot)|z^0]} \right) > 0 \end{aligned} \quad (\text{F30})$$

### F.5. Proof of Equation (19)

In the economy without capital taxation, an agent with labor income  $y$  solves:

$$V^0(y) \stackrel{\text{def}}{=} \max_{b,z} u(A - b - z) + \mathbb{E}[y - T^0(y) + v((1 + r(b))b + (1 + x)z)|z] \quad (\text{F31})$$

Now introduce a linear tax  $\tau$  on capital income  $xz$ , so that  $\hat{t}(xz) = \tau \cdot xz$ . From Lemma 2, I know that for every tax rate  $\tau$ , there exists a corresponding labor income tax schedule  $\hat{T}(y, \tau)$  so that the reform does not affect the utility of taxpayers.

Under  $\{\hat{T}(\cdot); \tau\}$  an agent with income  $y$  must solve:

$$\hat{V}(y, \tau) \stackrel{\text{def}}{=} \max_s u(A - s) + \mathbb{E}[v(y - \hat{T}(y, \tau) + (1 + r(b))b + (1 + (1 - \tau)x)z)|z] \quad (\text{F32})$$

Denote by  $\hat{z}(\tau)$  the solution of (F32). Under the candidate  $\{\hat{T}(\cdot); \tau\}$  the expected fiscal contribution of a taxpayer with labor income  $y$  is given by:

$$\mathcal{B}(y, \tau) = \hat{T}(y, \tau) + \tau \cdot \bar{x}(\hat{z}(\tau)) \cdot \hat{z}(\tau)$$

The impact on government revenue of introducing linear risky capital income taxation is equal to:

$$\frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} = \frac{\partial \hat{T}(y, \tau)}{\partial \tau} + \bar{x}(\hat{z}(\tau)) \cdot \hat{z}(\tau) + \tau[\hat{z}'(\tau) \cdot \bar{x}'(\hat{z}(\tau)) \cdot \hat{z}(\tau) + \hat{z}'(\tau) \cdot \bar{x}(\hat{z}(\tau))] \quad (\text{F33})$$

The definition of  $\hat{T}(y, \tau)$  implies:

$$\begin{aligned} \hat{V}(y, \tau) &= V^0(y) \\ \Rightarrow \frac{\partial \hat{V}(y, \tau)}{\partial \tau} &= \frac{\partial V^0(y)}{\partial \tau} = 0 \end{aligned} \quad (\text{F34})$$

Applying the envelope theorem to (F32) and using (F34) yields:

$$\frac{\partial \hat{T}(y, \tau)}{\partial \tau} = -z(\tau) \frac{\mathbb{E}[xv'(\cdot)|\hat{z}(\tau)]}{\mathbb{E}[v'(\cdot)|\hat{z}(\tau)]} \quad (\text{F35})$$

Plugging (F35) in (F33) gives:

$$\begin{aligned} \frac{\partial \mathcal{B}(y, \tau)}{\partial \tau} &= -\hat{z}(\tau) \frac{\mathbb{E}[xv'(\cdot)|\hat{z}(\tau)]}{\mathbb{E}[v'(\cdot)|\hat{z}(\tau)]} + \bar{x}(\hat{z}(\tau)) \cdot \hat{z}(\tau) + \tau[\hat{z}'(\tau) \cdot \bar{x}'(\hat{z}(\tau)) \cdot \hat{z}(\tau) + \hat{z}'(\tau) \cdot \bar{x}(\hat{z}(\tau))] \\ &= -\hat{z}(\tau) \left( \frac{\text{Cov}[xv'(\cdot)|\hat{z}(\tau)]}{\mathbb{E}[v'(\cdot)|\hat{z}(\tau)]} \right) - \tau \frac{\partial \hat{z}(\tau)}{\partial (1 - \tau)} (\bar{x}'(\hat{z}(\tau)) \cdot \hat{z}(\tau) + \bar{x}(\hat{z}(\tau))) \\ &= -\hat{z}(\tau) \left( \frac{\text{Cov}[xv'(\cdot)|\hat{z}(\tau)]}{\mathbb{E}[v'(\cdot)|\hat{z}(\tau)]} \right) - \frac{\tau}{1 - \tau} \cdot \hat{z}(\tau) \cdot \bar{x}(\hat{z}(\tau)) \cdot \epsilon_{1-\tau}^z \cdot \left( 1 + \epsilon_{\bar{x}}^{\bar{x}} \right) \end{aligned} \quad (\text{F36})$$

At the optimum we should have  $\frac{\partial B(y, \tau)}{\partial \tau} = 0$ . It therefore follows from (F36) that:

$$\begin{aligned} \frac{\tau}{1-\tau} \cdot \hat{z}(\tau) \cdot \bar{x}(z) \cdot \epsilon_{1-\tau}^z \cdot \left(1 + \epsilon_z^{\bar{x}}\right) &= -\hat{z}(\tau) \left( \frac{\text{Cov}[xv'(\cdot) | \hat{z}(\tau)]}{\mathbb{E}[v'(\cdot) | \hat{z}(\tau)]} \right) \\ \frac{\tau}{1-\tau} &= - \left( \frac{\text{Cov}[xv'(\cdot) | \hat{z}(\tau)]}{\bar{r}(s)\mathbb{E}[v'(\cdot) | \hat{z}(\tau)]} \right) \cdot \frac{1}{\epsilon_{1-\tau}^z} \cdot \frac{1}{1 + \epsilon_z^{\bar{x}}} \end{aligned}$$

## APPENDIX G: PROOF OF PROPOSITION 1 WHEN LABOR INCOME IS EARNED AT TIME 1

To test whether the two-type results of Gabaix et al. (2016) and Kristjánsson (2016) extend to my continuous type framework with both scale dependence and risky returns, I consider an alternative to problem (1) where labor income is no longer earned at time 2 but at time 1:

$$U(\theta) \stackrel{\text{def}}{\equiv} \max_{y,s} u(A + y - s) + \int_{r \in \mathbb{R}} v((1+r)s - T(y) - t(s, rs)) f(r|s) dr - h(y, \theta) \quad (\text{G1})$$

An agent with labor income  $y$  solves:

$$V^0(y) \stackrel{\text{def}}{\equiv} \max_s u(A + y - s) + \mathbb{E}[v((1+r)s - T^0(y) - t^0(s, rs)) | s] \quad (\text{G2})$$

Note by  $s^0(y)$  the solution of (G2). The full problem an agent with type  $\theta$  must solve can be written as:

$$U^0(\theta) = \max_y V^0(y) - h(y, \theta) \quad (\text{G3})$$

Denote by  $y^0(\theta)$  the solution of (G3) and define the function  $W : \mathbb{R}^2 \mapsto \mathbb{R}$  as:

$$W(y, a) \stackrel{\text{def}}{\equiv} \max_s u(A + y - s) + \mathbb{E}[v((1+r)s - a - \hat{t}(s, rs)) | s] \quad (\text{G4})$$

Applying the envelope theorem to (G4), I know that  $\frac{\partial W(y, a)}{\partial a} < 0$ . Hence the equation:

$$W(y, a) = V^0(y) \quad (\text{G5})$$

Under the candidate  $\{\hat{T}(\cdot), \hat{t}(\cdot)\}$ , an agent with labor income  $y$  still enjoys utility  $V^0(y)$  by solving:

$$V^0(y) = \max_s u(A + y - s) + \mathbb{E}[v((1+r)s - \hat{T}(y) - \hat{t}(s, rs)) | s] \quad (\text{G6})$$

Denote by  $\hat{s}(y)$  the solution of (G6). therefore total utility under the candidate schedule  $\{\hat{T}(\cdot), \hat{t}(\cdot)\}$  verifies:

$$\begin{aligned}\hat{U}(\theta) &= \max_y V^0(y) - h(y, \theta) \\ &= U^0(\theta)\end{aligned}\tag{G7}$$

To show that the candidate  $\{\hat{T}(\cdot), \hat{t}(\cdot)\}$  generates more government revenue than the tax system  $\{T^0(\cdot), t^0(\cdot)\}$ , one can solve the following problem:

$$\begin{aligned}\max_{s, c_2(r)} & (1 + \bar{r}(s))s - \int_{r \in \mathbb{R}} c_2(r) f(r|s) dr \\ \text{subject to: } & u(A + y - s) + \int_{r \in \mathbb{R}} v(c_2(r)) f(r|s) dr = V^0(y)\end{aligned}\tag{G8}$$

The solution of (G8) is characterized by the Euler Equation:

$$u'(A + y - s) = (1 + \bar{r}(s) + \bar{r}'(s)s)v'(c_2)\tag{G9}$$

Let  $\hat{c}_2 = (1 + \bar{r}(\hat{s}))\hat{s} - \hat{T}(y)$  be second-period consumption under  $\{\hat{T}(\cdot), \hat{t}(\cdot)\}$ . It thus follows from the first-order condition of problem (G2) that the bundle  $\{\hat{s}, \hat{c}_2\}$  satisfies (G9). Therefore  $\{\hat{s}, \hat{c}_2\}$  solves problem (G8).

This implies:

$$\begin{aligned}(1 + \bar{r}(\hat{s}))\hat{s} - \hat{c}_2 &\geq (1 + \bar{r}(s^0))s^0 - \mathbb{E}\left[c_2^0(r) \middle| s^0\right] \\ \Rightarrow \hat{T}(y) &\geq T^0(y) + \mathbb{E}[t^0(s^0, rs^0) | s^0]\end{aligned}$$

The candidate  $\{\hat{T}(\cdot), \hat{t}(\cdot)\}$  generates more tax revenue than the initial tax system  $\{T^0(\cdot), t^0(\cdot)\}$ , without changing individual utility, which proves that Proposition 1 extends to the case where labor income is earned at time 1.