OPTIMAL CAPITAL TAXATION UNDER STOCHASTIC RETURNS TO SAVINGS

Eddy Zanoutene*

July 16, 2021

Abstract

I present a model of optimal capital taxation with heterogeneous labor productivity and stochastic, potentially scale dependent, returns to savings. I examine the relevance of both capital income and wealth taxation. First, the optimal policy combines confiscatory capital income taxes with wealth transfers to perfectly insure agents against risky returns. However as soon as returns exhibit scale dependence, an access to the expected rate of return conditional on initial savings must be guaranteed at the optimum. Second, in constrained settings where the government does not observe capital income but only initial savings, there is no capital taxation at the optimum as the logic of Atkinson and Stiglitz (1976) applies. However, in environments where the government does not observe initial savings but has only information on capital income, the optimum does feature positive capital taxation as it provides some form of insurance against risky returns. This positive capital taxation at the optimum extends to the case where the government only observes ex-post wealth, *i.e* the sum of savings and capital income. These results are valid under any social welfare function.

1 Introduction

In standard optimal capital taxation models, financial markets are assumed efficient, yielding a unique, deterministic, rate of return to savings. This homogeneity in returns usu-

*Université Paris II Panthéon Assas - CRED, 75005, 21 Rue Valette, Paris, France.

Email : eddy.zanoutene@u-paris2.fr

ally limit the relevance of capital taxation when labor income taxes are already implemented.[1] Yet, in an attempt to replicate the dynamics of wealth concentration documented for instance in Saez and Zucman (2016), a wave of theoretical papers 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.[2] On top of this theoretical argument, recent contributions by Fagereng et al. (2020) and Bach et al. (2020) present empirical evidences of an important dispersion in returns to savings.[3] In both papers, the distribution of returns appears highly correlated with wealth, with household in higher wealth brackets experiencing higher expected returns. This correlation between the rate of return and wealth, often described as scale dependence, could generate fast transition in wealth inequalities, potentially explaining the dynamic of wealth concentration at the top observed in the US.[4] Besides, scale dependence might not only influence the average returns to savings but also its riskiness, with the observed dispersion in returns varying across the wealth distribution.[5] The distributive and efficiency consequences of such stochastic, scale dependent, returns to savings can therefore have implications for the optimal design of capital taxes.

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 is uncertain and can depend on the amount saved. 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 in the first period. However, departing from the perfect capital market assumption of traditional optimal tax model, I suppose that the rate of return is randomly drawn from a wealth-correlated distribution, at the beginning of the second period. 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

[1] In the framework of Atkinson and Stiglitz (1976), capital taxation is zero at the optimum in presence of an optimal labor income tax schedule. Laroque (2005) and Kaplow (2006) extend this zero tax result to non-optimal labor income tax schedule. A review of this standard approach to capital taxation can be found in Chapter 9 of Kaplow (2010).

[2] See for instance Benhabib et al. (2011), Gabaix et al. (2016) or Benhabib and Bisin (2018).

[3] Fagereng et al. (2020) documents a 22% standard deviation in returns to net wealth, using Norwegian data.

[4] Scale dependence has been introduced by Gabaix et al. (2016) to actually produce such fast transitions in wealth inequality dynamics.

[5] Bach et al. (2020) shows on Swedish data that expected returns on gross wealth rises monotonically with wealth. The pattern in risk is more complex : volatility of the net wealth return is for instance higher in the ninth decile than in the third. Still, wealthier households experiments higher return heterogeneity with a 27% standard deviation of the excess logarithmic return.

know that the amount they save determines their expected rate of return on savings. In the second period, they use all their savings and the returns on their savings for consumption.

My first objective is to pin down the optimal capital tax induced by stochastic returns when both initial savings and capital income are observed, and can be taxed, by the government. In this context, the optimal policy consists in combining confiscatory capital income taxes with wealth transfers to perfectly insure agents against the risk associated to stochastic returns. However, to avoid distortions of the saving decision in a context where returns can exhibit scale dependence, wealth transfers need to be conditioned on the amount originally saved. Doing so, the government grant access to the rate of return taxpayers can expect, given their initial saving decision. This result first implies that introducing stochastic returns in a framework otherwise identical to Atkinson and Stiglitz (1976) is sufficient to break the zero-capital tax result. Second, conditioning the wealth subsidy on the saving decision originally made by taxpayers departs from the standard excess return tax, as described for instance in Mirrlees et al. (2011). Indeed, in a context of scale dependence, the standard distinction between normal and excess returns becomes blurry as it depends on the amount saved in the beginning : a return considered as excessive at low saving level could be considered as normal at higher saving level.

The other ambition of this paper is to assess the relevance of capital income and wealth taxation in constrained environments where the government does not observe all components of capital. First, there is no capital taxation at the optimum when the government only observe initial savings and not capital income. In this specific context, stochastic returns with scale dependence does not provide a rationale for taxing capital and the logic of Atkinson and Stiglitz (1976) applies : non-linear labor income taxation is sufficient to maximize any social welfare function. This is important as scale dependence, potentially yielding "rich get richer" phenomena, could intuitively justify capital taxation as soon as the government is sufficiently inequality averse. Second, it appears that the optimum does feature positive capital taxes when the government only observes capital income and not initial savings. This is fully explained by the form of insurance a capital income tax can provide as it reduces the variance of stochastic returns, balancing its negative impact on expected returns. Third, this insurance rationale for taxing capital extends to the case where the government only observes ex-post wealth, that is initial savings augmented with capital income. This specific form of wealth taxation is a priori ambiguous as it mitigates the riskiness of stochastic returns at the expense of inefficiently taxing savings. However I rigorously show that the insurance gains prevail and that the optimum does features positive capital taxation when only ex-post wealth is observed by the government.

Related literature.

First, this paper relates to previous works analyzing the consequences of uncertainty on capital taxation. Uncertain capital income has long been regarded as a poor rationale for capital taxation as illustrated in Domar and Musgrave (1944) where the implicit insurance provided by capital taxation is completely offset by an increase in the riskiness of portfolios. Extended and balanced by Stiglitz (1975), Bulow and Summers (1984) or Gordon (1985), this seminal paper has drawn major attention to the effect of capital taxation on risk-taking and portfolio choices. An optimal tax approach to such problematic has been recently provided by Boadway and Spiritus (2021) where capital income taxation is discussed in various risk and saving technology environments. In particular, their work also provide an insurance rationale for taxing above-normal returns arising from idiosyncratic risk. Departing from this portfolio choice analysis, a strand of the literature insists on the trade-off between social insurance and reduced incentives implied by capital taxation. In this sense, my main result echoes Varian (1980) which also studies the trade-off between social insurance and reduced incentives in a two-period model where the dispersion in second-period observed income only comes from exogenous differences in luck. However, in Varian (1980), agents do not make labor supply decision and start with an exogenous income in the first-period. The only endogenous outcome is the amount of savings, with agents deciding how to split this exogenous initial income between first and second-period consumption. Therefore, the impact of taxation on labor supply is ignored. On the contrary, in the present paper, labor income is endogenous and comes from the labor supply decision of agents with heterogeneous productivity, in the pure tradition of Mirrlees (1971).

Second, this paper allows the rate of return to be correlated with wealth and therefore falls within the literature studying the consequences for capital taxation of such return heterogeneity. Returns mechanically increasing with wealth, what it often referred as scale dependence in Gabaix et al. (2016)-terminology, could provide a rationale for taxing capital. Gerritsen et al. (2020) indeed show that scale effects in wealth management justify a positive capital income tax at the optimum. Their analysis also allows the rate of return to depend on ability, echoing two-types models such as Kristjánsson (2016) or Gahvari and Micheletto (2016), which is beyond the scope of this paper. The present work is however the first, to our knowledge, to include scale dependence in a risky environment. Therefore, two agents with the same labor income can expect different rates of return, either because they save different amounts (due to scale dependence) or because one got luckier than the other (due to the volatility of returns). Moreover, these two dimensions of return heterogeneity are intrinsically linked in our framework as our model exhibit both scale dependence in expected returns but also in the dispersion of returns. This feature of the model therefore matches the empirical evidences of Bach et al. (2020) and Fagereng et al. (2020) where both expectations and standards of deviation of returns vary dramatically across the wealth distribution.

Eventually, constraining the available capital tax instruments of the government allows me to contribute to the wealth against capital income taxation debate, in an Atkinson and Stiglitz (1976)-like framework where usually both form of capital taxation are ruled out at the optimum. While Kaplow (1991) argues that return uncertainty is unlikely to deliver key insights regarding this debate between a priori equivalent instruments, I show that stochastic returns rather advocate for capital income taxation because of the insurance it provides. Note however that a wealth tax understood as a tax on ex-post wealth does share some insurance properties with a pure capital income tax and can therefore be part of the optimal policy. The insurance argument put forward in this paper can counterbalance the rationale of Guvenen et al. (2019) where a wealth tax is preferred to capital income taxation as it allows to reallocate capital towards credit-constrained entrepreneurs. Furthermore, note that in our model, agents are ex-ante homogeneous with respect to their initial wealth so that part of the rationale for wealth taxation provided for instance by Piketty and Saez (2013) is not examined in the paper.[6]

The remainder of this paper is organized as follows : Section 2 sets up the model. Section 3 derives the optimal capital tax schedule in a world where the government observes both savings, returns to savings and labor income. Section 4 highlights that there is no capital taxation at the optimum when the government only observes labor income and initial savings. In Section 5, I show that, in a framework where the government does not observe initial savings but observe capital income, positive capital income taxation is required at the optimum. Besides, I show in Section 6 that this result extends to the case where the government only observes labor income and ex-post wealth, that is the sum of savings and capital income. Section 7 concludes.

[6] One could argue that such form of ex-ante heterogeneity resulting from differences in inherited wealth, if considered problematic, could be dealt with by using bequest taxation, which is not the topic of the present paper.

2 A Two-Period Economy

I model a two-period economy where a continuum of individuals work only in the first period and consume in both periods. 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 θ, individuals choose pre-tax labor income $y \in \mathbb{R}_{+}$ (for short labor income hereafter) and savings $s \in \mathbb{R}_{+}$.

At the beginning of the second period, agents draw a rate of return $r \in \mathbb{R}$ from some probability distribution. I call capital income the product of savings s and the rate of return r. This rate r may depend on savings itself so the density depends on savings s and is denoted $f(r|s)$ while the c.d.f is denoted $F(r|s)$. Such modeling strategy captures what Gabaix et al. (2016) identifies as scale dependence, *i.e* the influence the amount saved can have on the distribution of the rate of return. There exists various mechanisms, ranging from fixed costs in wealth management to decreasing relative risk aversion, that could lead agents with higher savings to earn higher rate of return on average. More generally, as empirically documented by Bach et al. (2020) or Fagereng et al. (2020), the amount of initial savings can influence both the expectation and the variance of the rate of return. Studying the conditional distribution of the rate of return therefore allows me to embed such empirical regularities in the optimal tax exercise.

2.1 Taxpayers

I denote by $u(.)$ the function measuring utility from first-period consumption, by $v(.)$ the function measuring utility from second-period consumption and by $h(.)$ the function measuring disutility from work effort. I assume that $u(.)$, $v(.)$ and $h(.)$ are both strictly increasing, twice continuously differentiable with $u(.)$ and $v(.)$ concave while $h(.)$ is convex.[7] Besides, both $u(.)$ and $v(.)$ satisfy Inada's conditions with $\lim_{x \to 0} u'(x) = \lim_{x \to 0} v'(x) = +\infty$ and $\lim_{x \to +\infty} u'(x) = \lim_{x \to +\infty} v'(x) = 0$.

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. Assuming additively separable preferences for consumption and leisure, an agent with type θ solves the following problem :

[7] One these three functions $u(.)$, $v(.)$ and $h(.)$ can be linear to include the quasilinear setting.

$$U(\theta) \overset{\text{def}}{\equiv} \max_{y,s} \quad u\left(y-s\right) + \int\limits_{r\in\mathbb{R}} v\left(\left(1+r\right)s - T(y) - t\left(s,rs\right)\right) f(r|s)dr - h\left(y,\theta\right) \tag{1}$$

The solution of (1) is denoted by $\{y(\theta), s(\theta)\}$. Note that the problem described in (1) departs from the common assumption that labor is taxed in the first-period while capital is taxed in the second period. To detail the reason for this modeling strategy, I first need to study the problem the government is facing.

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 $T(.)$ or capital taxation through $t(.)$. Both $T(.)$ and $t(.)$ can be non-linear. Among the many candidates for capital taxation, it is worth mentioning some specific cases such as :

- *Capital income taxation* : $t : rs \mapsto t(rs)$, which is a form of capital taxation that is solely based on the income generated by savings.
- *Ex-ante wealth taxation* : $t : s \mapsto t(s)$. Note that in this economy without bequest, the wealth of taxpayers at the beginning of the first period is zero.[8] Therefore wealth at the end of the first-period is just equal to savings s so that a tax on ex-ante wealth is just a tax on savings. This form of taxation can also be seen as a tax on the book value of wealth.
- *Ex-post wealth taxation* : $t : (1+r)s \mapsto t\left(\left(1+r\right)s\right)$. This is somehow analogous to a tax on the market value of wealth.

Assuming that both labor and capital taxes are levied in the second period, the intertemporal budget constraint of the government is given by :

$$\int\limits_{\theta\in\Theta} T\left(y\left(\theta\right)\right) dG\left(\theta\right) + \iint\limits_{\theta\in\Theta;\, r\in\mathbb{R}} t\left(s\left(\theta\right), rs\left(\theta\right)\right) f\left(r|s\left(\theta\right)\right) dr\, dG\left(\theta\right) \geq E \tag{2}$$

[8]In economies with bequest as studied for instance in Piketty and Saez (2013), wealth at time 0 is equal to some exogenous inheritance.

If instead of assuming that tax revenue are collected in the same period one makes the standard assumption that labor is taxed in the first period and capital in the second period, then writing the intertemporal budget constraint requires to choose the adequate discount factor. In the case where the draw of the rate of return does not depend on savings, one can rigorously prove, as done in Appendix A.1, that the right discount factor for government revenue is $\frac{1}{1+\bar{r}}$, with $\bar{r} = \mathbb{E}(r)$. This proof is achieved without specifying the saving technology the government has access to as the right discount factor naturally arises from the aggregate resource constraint in the economy.[9] However, discounting government revenue is not as trivial when the rate of return can depend on savings. Indeed, 1 dollar taken away from a taxpayer with high savings would not on average yield the same outcome in the second period than 1 dollar taken away from a taxpayer with low savings. To keep the optimal taxation exercise independent from the choice of the saving technology the government has access to, I circumvent the problem of the discount factor by assuming that all taxes are levied at the same time.

2.3 A Two-Step Procedure for Saving and Labor Supply Decisions

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 the utility of taxpayers. I therefore need a method that cancels the impact on utility of changing capital taxation. Building on the proof of Atkinson and Stiglitz (1976) given by Konishi (1995), Laroque (2005) and Kaplow (2006), I adjust labor income taxation in order to neutralize the impact a reform of capital taxation has on utility. I show in this section that, in my framework, there always exists such utility-neutralizing reform of the labor income tax schedule.

First, note that separability allows me to split the problem defined in (1) between a subproblem over the saving decision for a given level of labor income y and a subproblem over the labor supply choice. Hence for a given y, taxpayers solve :

[9] The intuition is that 1 dollar taken away from taxpayers in the first-period could have been saved by the government yielding on average $1 + \bar{r}$ dollar in the second-period. Such reasoning would not be valid if macroeconomic shocks could affect the whole distribution of r, thus creating aggregate uncertainty as studied for instance in Kocherlakota (2005) or more recently in Boadway and Spiritus (2021).

$$V(y) \overset{\text{def}}{\equiv} \max_s \quad u\,(y-s) + \int_{r\in\mathbb{R}} v\,((1+r)\,s - T(y) - t\,(s,rs))\, f(r|s)dr \tag{3}$$

To compare government revenue under various capital tax schedule, I rely on the first-order condition of problem (3) to characterize the saving decision made by taxpayers. First-order conditions might however not be sufficient as the concavity of the objective of problem (3) depends on the second derivatives with respect to s of both the capital tax function $t(s,rs)$ and the density function $f(r|s)$. I therefore restrict my analysis to cases where problem (3) is well-behaved by assuming the following:

Assumption 1. *The density function $f(r|s)$ and the capital tax function $t(s,rs)$ are such that the objective of problem 3 remains concave.*

To conduct the optimal capital tax analysis, start from a given tax environment $\{T^0(.), t^0(.)\}$ with $t^0(.)$ verifying assumption (1). An agent with labor income y solves :

$$V^0(y) \overset{\text{def}}{\equiv} \max_s \quad u\,(y-s) + \mathbb{E}\left[v\left((1+r)\,s - T^0(y) - t^0\,(s,rs)\right) \mid s\right] \tag{4}$$

Note by $s^0(y)$ the solution of (4). The full problem an agent with type θ must solve can be written as :

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

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

Now suppose that I want to study the welfare consequences of moving from the initial capital tax schedule $t^0(.)$ to a candidate schedule that verifies assumption 1 and that I denote by $\widehat{t}(.)$. To show that there exists a labor income tax that neutralizes the impact on utility of implementing the candidate $\widehat{t}(.)$, define the function $W : \mathbb{R}^2 \mapsto \mathbb{R}$ as :

$$W(y,a) \overset{\text{def}}{\equiv} \max_s \quad u\,(y-s) + \mathbb{E}\left[v\left((1+r)\,s - a - \widehat{t}\,(s,rs)\right) \mid s\right] \tag{6}$$

Under assumption 1, the first-order-condition of (6) admits a single solution. From the implicit function theorem, I know that this unique solution is continuous and differentiable in a.[10] Therefore $W(y,a)$ is continuous and differentiable in a. Applying the envelope theorem to (6), I know that $\frac{\partial W(y,a)}{\partial a} < 0$. Hence the equation :

$$W(y,a) = V^0(y) \tag{7}$$

admits a single solution denoted by $a(y)$. It follows from (7) that the labor income tax function $\hat{T} : y \mapsto a(y)$ neutralizes the impact the candidate capital tax function $\hat{t}(.)$ has on utility from consumption. Indeed under the candidate $\{\hat{T}(.), \hat{t}(.)\}$, an agent with labor income y still enjoys utility $V^0(y)$ by solving :

$$V^0(y) = \max_s \quad u\,(y-s) + \mathbb{E}\left[v\left((1+r)\,s - \widehat{T}(y) - \widehat{t}\,(s, rs)\right) \mid s\right] \tag{8}$$

Denote by $\hat{s}(y)$ the solution of (8). Total utility under the candidate schedule $\{\widehat{T}(.), \widehat{t}(.)\}$ therefore verifies :

$$\begin{aligned} \hat{U}(\theta) &= \max_y V^0(y) - h(y,\theta) \\ &= U^0(\theta) \end{aligned} \tag{9}$$

Note that $y^0(\theta)$ still solves (9) so that both total utility and labor supply are the same under the candidate $\{\widehat{T}(.), \widehat{t}(.)\}$ than under the original schedule $\{T^0(.), t^0(.)\}$:

Lemma 1. *When a government implements a candidate tax schedule on capital $\widehat{t}(.)$, there always exists an alternative labor income tax schedule $\widehat{T}(.)$ that mutes the impact of the reform on :*

1. *sub-utility from consumption*
2. *total utility*
3. *labor supply.*

3 Optimal Capital Taxation when both Savings and Capital Income are Observed

In this section, I suppose that the government observes both savings s and capital income rs, such that only skills θ are private information. Starting from a given tax schedule

[10] Let $w(a,y,s)$ be the derivative w.r.t s of the objective of (6). Thus $w(a,y,\tilde{s}) = 0$, with $\tilde{s}$ the solution of (6). Therefore from the implicit function theorem there exists a continuously differentiable function $\phi(.)$ such that $\tilde{s} = \phi(a,y)$

on capital $t^0(s, rs)$, an individual with labor income y solves the problem defined in (4) and reaches utility level $V^0(y)$.

Denote by $\bar{r}(s) = \mathbb{E}[r|s]$ the average rate of return for a given level of savings s. Now suppose that the government implements a new policy combining a confiscatory tax on capital income rs with a wealth subsidy equal to $\bar{r}(s)s$. Formally this reform replaces the original schedule $t^0(.)$ by an alternative tax function $\widehat{t}(.)$ defined as :

$$\widehat{t} : (s, rs) \mapsto \widehat{t}(s, rs) = rs - \bar{r}s \tag{10}$$

The candidate $\widehat{t}(.)$ provides perfect insurance at each saving level s since second-period consumption is now given by $c_2 = (1 + \bar{r}(s))s - T(y)$, which does not depend on the draw of r. If the draw of r is independent of savings, this full insurance policy would grant the same rate of return $\bar{r}$ to every taxpayers as they would face exactly the same risk. However, if the draw of r can depend on s, agents making different saving decisions can expect different rate of return and therefore do not face the same risk. A natural candidate for the optimal capital tax function therefore takes the form of $\hat{t}(.)$. As explained in Section 2.3, to rigorously establish the optimality of the candidate $\widehat{t}(.)$, I will show that it is possible to generate more government revenue by implementing $\widehat{t}(.)$ instead of $t^0(.)$, without affecting utility. From Lemma 1, I know that there exists a labor income tax schedule $\widehat{T}(.)$ such that the utility of taxpayers is not changed by the introduction of $\widehat{t}(.)$. So suppose that the government moves from the tax system $\{T^0(.), t^0(.)\}$ to the full insurance candidate $\left\{\widehat{T}(.), \widehat{t}(.)\right\}$.

Under the candidate $\left\{\widehat{T}(.), \widehat{t}(.)\right\}$, an agent with labor income y must solve :

$$\hat{V}(y) = \max_{s} \quad u(y - s) + v\left((1 + \bar{r}(s))s - \widehat{T}(y)\right) \tag{11}$$

Under assumption 1 the saver's problem admits a single solution, denoted by $\hat{s}$, characterized by the Euler Equation :

$$u'(y - \hat{s}) = (1 + \bar{r}(\hat{s}) + \bar{r}'(\hat{s})\,\hat{s})\, v'\left((1 + \bar{r}(\hat{s}))\hat{s} - \widehat{T}(y)\right) \tag{12}$$

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

$$\begin{aligned} &\max_{s,c_2(r)} (1+\overline{r}(s))\, s - \int\limits_{r\in\mathbb{R}} c_2(r) f(r|s) dr \\ &\text{subject to}: u(y-s) + \int\limits_{r\in\mathbb{R}} v\left(c_2(r)\right) f(r|s) dr \geq V^0(y) \end{aligned} \tag{13}$$

In the simple case where the draw of r is independent of savings s, problem (13) is convex. Since it is not necessarily the case when the draw of r can depend on savings s, I assume that the density function $f(.)$ is such that the objective of (13) is at least quasiconcave.

I show in Appendix A.2 that the solution of (13) is characterized by the Euler Equation :

$$u'\left(y-s\right) = \left(1+\overline{r}\left(s\right)+\overline{r}'(s)\; s\right) v'(c_2) \tag{14}$$

Let $\hat{c}_2 = (1+\overline{r}(\hat{s}))\,\hat{s} - \widehat{T}(y)$ be second-period consumption under the candidate $\{\widehat{T}\,(.)\,,\hat{t}\,(.)\}$. It thus follows from the individual's first-order condition (12) that the bundle $\{\hat{s},\hat{c}_2\}$ satisfies (14). Therefore $\{\hat{s},\hat{c}_2\}$ solves problem (13).

Let $c_2^0(r) = (1+r)\,s^0 - T^0(y) - t^0\left(s^0, rs^0\right)$ be second-period consumption under the initial tax system $\{T^0(.), t^0(.)\}$. The bundle $\{s^0, c_2^0\}$ satisfies the inequality constraint of problem (13) but not the Euler Equation (14). Therefore it is not a solution of (13), which implies :

$$\begin{aligned} (1+\overline{r}(\hat{s}))\,\hat{s} - \hat{c}_2 &> \left(1+\overline{r}(s^0)\right) s^0 - \mathbb{E}[c_2^0(r) \mid s^0] \\ \Rightarrow \widehat{T}(y) &> T^0(y) + \mathbb{E}\left[t^0(s^0, rs^0) \mid s^0\right] \end{aligned} \tag{15}$$

The candidate $\left\{\widehat{T}\,(.)\,,\hat{t}\,(.)\right\}$ generates more tax revenue than the initial tax system $\{T^0(.), t^0(.)\}$, without changing individual utility.

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 :*

$$t^*\left(s, rs\right) = s\left(r - \overline{r}\left(s\right)\right)$$

Recall that under $t^*(.)$ second-period consumption always amounts to $(1+\overline{r})s - \widehat{T}(y)$. Thus proposition 1 implies that the optimal policy fully insure agents against the uncertainty arising from the draw of r. By delivering the average rate of return an agent can expect given the saving decision she made, perfect insurance is achieved while incentives to save are preserved. Note first that if returns are no longer stochastic, $\overline{r} = r$ so that Atkinson and Stiglitz (1976) applies with $t^*(.) = 0$. Second, remark that if the rate of return is independent of savings, then the optimal capital tax boils down to a tax on excess return : $t^{ER} = s(r - \bar{r})$. This would result in redistribution between group of savers. However when the rate of return actually depends on savings, implementing t^{ER} would certainly provide insurance but at the expense of distorting the saving decision made by taxpayers. On the contrary the optimal policy defined in proposition 1 rewards savings accordingly and does not inefficiently alter the individual's intertemporal choice.

Such a policy can sound counterintuitive from an equity perspective. Indeed, suppose that rates of return rise monotonically with savings, creating a "rich get richer" effect. The automatic rise in wealth inequality induced by such a pattern could provide a rationale for progressive capital taxation, to curb wealth accumulation at the top, or transfers between different level of savings to allocate returns to capital more equally. Yet, the optimal capital tax schedule described in proposition 1 does not advocate for transfers between groups of savers but only within a group of savers. Indeed, insurance is here provided by only transferring money between people who have saved the same amount. The intuition behind this result is that in our framework, two components of wealth can be seen as unfair. The first one is the luck of drawing a high rate of return r, given the saving decision. Capital taxation deals with this form of unfairness. The second one is the luck of drawing a high skill parameter θ that yields different labor income for the same effort. Progressive labor income taxation deals with this second form of unfairness in the wealth accumulation process.[11] Therefore, in such an Atkinson and Stiglitz (1976) framework, there is no reason to transfer money from high savers to low savers, as long as there exists a progressive labor income tax schedule. Indeed, at a given level of labor income, the saving decision is just the result of an arbitrage between first and second-period consumption. Nevertheless, note that the optimal capital tax schedule $t^*(.)$ can have some interesting consequences for wealth inequality. At the top, it can prevent high savers from earning unusually high level of capital income. At the bottom, it could inhibit some poverty trap patterns, where agents with low level of savings experiment unusually low rates of returns. According to Fagereng et al. (2020) and Bach et al. (2020), the dispersion in rates of return appears particularly strong at both the top and the bottom of the distribution so that this mitigating impact of $t^*(.)$ on inequality could be significant.

The optimal policy $t^*(.)$ requires that the government perfectly observes both savings and capital income, which could prove costly in practice. Besides, given the Atkinson and

[11]Obviously there exists other patterns in wealth creation that can be seen as unfair, starting from the capital one inherits at birth, but those are beyond our Atkinson and Stiglitz (1976)-like framework.

Stiglitz (1976) structure of the problem analyzed here, it is not obvious that capital taxation would even be needed if we depart from the perfect information benchmark. Studying the optimal taxation problem in a constrained environment where either savings or capital income are not observed could therefore be insightful from both a practical and a theoretical point of view.

4 No Capital Taxation when only Savings are Observed

In this section I suppose that the government observes savings but not capital income, so that the rate of return r is private information. An ex-ante wealth tax $s \mapsto t(s)$ therefore becomes the only available form of capital taxation in this framework. I will show that moving from an economy where there exists such a wealth tax $t(.)$ to an economy with only labor income taxation increases welfare.

In an economy where labor is taxed through a function $T^0(y)$ and wealth through a function $t^0(s)$, the saving problem of a taxpayer with labor income y is given by :

$$V^0(y) \overset{\text{def}}{\equiv} \max_s \quad u\,(y-s) + \int\limits_{r\in\mathbb{R}} v\left((1+r)\,s - T^0(y) - t^0(s)\right) f(r|s)dr \tag{16}$$

Denote by s^0 the solution of (16). Let $c_2^0 = s^0 - T^0(y) - t^0\left(s^0\right)$ be the associated deterministic component of second-period consumption. Indeed, this part of second-period consumption is independent of the draw of r and is known by taxpayers before the realization of the shock.

Now suppose wealth taxation is erased. Again from Lemma 1, $T^0(.)$ can be replaced by a labor tax function $\widehat{T}(.)$ to keep individual utility unchanged. A taxpayer with labor income y now solves :

$$V^0(y) = \max_s \quad u\,(y-s) + \int\limits_{r\in\mathbb{R}} v\left((1+r)\,s - \hat{T}(y)\right) f(r|s)dr \tag{17}$$

Denote by $\hat{s}$ the solution of (17) with $\hat{c}_2 = \hat{s} - \hat{T}(y)$. Then $\{\hat{s}, \hat{c}_2\}$ verifies the Euler Equation :

$$u'(y-\hat{s}) = \int\limits_{r\in\mathbb{R}} (1+r)\,v'\left(\hat{c}_2 + r\hat{s}\right) f\left(r|\hat{s}\right) dr + \int\limits_{r\in\mathbb{R}} v\left(\hat{c}_2 + r\hat{s}\right) f_s\left(r|\hat{s}\right) dr \tag{18}$$

To show that $\widehat{T}(.)$ generates more government revenue than $\{T^0(.), t^0()\}$, one can solve the following cost minimization problem :

$$\begin{aligned}
&\min_{s,c_2} c_2 - s \\
&\text{subject to} : u(y-s) + \int\limits_{r\in\mathbb{R}} v\left(c_2 + rs\right) f(r|s)dr \geq V^0(y)
\end{aligned} \tag{19}$$

Note that both $\{s^0, c_2^0\}$ and $\{\hat{s}, \hat{c}_2\}$ satisfy the inequality constraint of problem (19). Suppose that the constraint set is convex, then a necessary and sufficient condition to solve (19) is given by:

$$u'(y-s) = \int\limits_{r\in\mathbb{R}} (1+r)\, v'\left(c_2 + rs\right) f\left(r|s\right) dr + \int\limits_{r\in\mathbb{R}} v\left(c_2 + rs\right) f_s(r|s)dr \tag{20}$$

It directly follows from (18) that the bundle $\{\hat{s}, \hat{c}_2\}$ satisfies the condition (20) and is therefore a solution of problem (19). This implies that for all y :

$$\begin{aligned}
&\hat{c}_2 - \hat{s} \leq c_2^0 - s^0 \\
&\Rightarrow \widehat{T}(y) \geq t^0(s^0) + T^0(y)
\end{aligned} \tag{21}$$

where the inequality is strict whenever $\frac{\partial t^0(s)}{\partial s} \neq 0$. Thus erasing capital taxation is Pareto-improving as it increases government revenue without affecting individual utility.

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

In this constrained environment, the government does not observe the rate of return drawn by the agent. Therefore capital taxation, which boils down here to a tax on savings, can no longer serve as an insurance device against the volatility of returns. Because of the assumptions made on preferences and the dimensions of unobserved heterogeneity, capital taxation is useless at redistributing resources in presence of a non-linear labor income tax. Since it does not provide insurance either, the logic of Atkinson and Stiglitz (1976) applies and there is no need for capital taxation. Again, it is worth-mentioning that the impact savings have on the draw of the rate of return, which departs from the standard Atkinson and Stiglitz (1976) framework, does not necessarily provide a rationale for taxing capital. Indeed, even in settings where $f(r|s)$ is increasing with s, yielding a "rich get richer" effect, capital taxation in the form of an ex-ante wealth tax is zero at the optimum. The intuition for this result is that savings remain a residual of labor income. Hence if one is concerned by the "rich get richer" effect of wealth-correlated returns, then adjustment in the non-linear labor income tax schedule are sufficient to tackle such inequalities. Again, this reasoning would not be valid under larger dimensions of heterogeneity in endowments or preferences.[12] The constrained environment used in this section mutes the insurance properties of capital taxation. The question is now to understand the relevance of such taxation in constrained settings where capital taxes could still grant some form of insurance against risky returns.

5 The Need for Capital Taxation when only Capital Income is Observed

In this section, I suppose that capital income rs is perfectly observed by the government while savings s are only private information. Therefore capital taxation boils down to capital income taxation $t : rs \mapsto t\,(rs)$. The aim of this section is to assess the relevance of taxing capital in such a constrained framework. In that respect, the objective here is less ambitious than in Section 3 and 4 as I do not try to pin down the optimal capital tax. To see if a capital income tax is needed at the optimum, I gauge the impact on welfare of introducing a linear tax on capital income. Therefore, instead of canceling capital taxation in an economy where both labor and capital are taxed, as done in Section 4, I introduce capital taxation in an economy that initially only taxes labor, *i.e* where $t^0(rs) = 0$. Denote by $T^0(.)$ the initial labor income tax function. Then, in the initial economy, an agent with labor income y solves :

$$V^0(y) \stackrel{\text{def}}{\equiv} \max_s \quad u\,(y-s) + \mathbb{E}\left[v\left((1+r)\,s - T^0(y)\right) \mid s\right] \tag{22}$$

Let s^0 be the solution of (22). Now introduce a linear tax τ on capital income rs, so that $\hat{t}(rs) = \tau \cdot rs$. From Lemma 1, I know that for every tax rate τ, there exists a corresponding labor income tax schedule $\widehat{T}(y,\tau)$ so that the reform does not affect the utility of taxpayers.

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

$$\hat{V}(y,\tau) \stackrel{\text{def}}{\equiv} \max_s \quad u\,(y-s) + \mathbb{E}\left[v\left((1+(1-\tau)\,r)\,s - \widehat{T}\,(y,\tau)\right) \mid s\right] \tag{23}$$

Denote by $\hat{s}(\tau)$ the solution of (23). Under the candidate $\{\widehat{T}(.)\ ;\ \tau\}$ the expected fiscal contribution of a taxpayer with labor income y is given by :

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

[12] The heterogeneity in inheritance in Piketty and Saez (2013) or the heterogeneity in preferences for the future in Farhi and Werning (2013) typically advocate for more tools then just labor income taxation to achieve redistribution.

Note that $\widehat{T}(y,0) = T^0(y)$ 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:

$$\left.\frac{\partial \mathscr{B}(y,\tau)}{\partial \tau}\right|_{\tau=0} = \left.\frac{\partial \widehat{T}(y,\tau)}{\partial \tau}\right|_{\tau=0} + \bar{r}(s^0)s^0 \tag{24}$$

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

$$\begin{aligned} \hat{V}(y,\tau) &= V^0(y) \\ \Rightarrow \frac{\partial \hat{V}(y,\tau)}{\partial \tau} &= \frac{\partial V^0(y)}{\partial \tau} = 0 \end{aligned} \tag{25}$$

Applying the envelope theorem to (23) and using (25) yields :

$$\left.\frac{\partial \widehat{T}(y,\tau)}{\partial \tau}\right|_{\tau=0} = -s^0 \frac{\mathbb{E}\left[rv'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}{\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]} \tag{26}$$

Since $v(.)$ is strictly increasing, the right-hand side of equation (26) is strictly negative : labor income taxation has to decrease to keep utility unchanged after capital income taxation is introduced. The $\left.\frac{\partial \widehat{T}(y,\tau)}{\partial \tau}\right|_{\tau=0}$ term in equation (24) is therefore negative. The question is now to compare this loss in labor income tax revenue with the gains associated to capital income taxation. To do so I show in Appendix A.3 that the impact on government revenue evaluated at $\tau = 0$ can be rewritten as :

$$\left.\frac{\partial \mathscr{B}(y,\tau)}{\partial \tau}\right|_{\tau=0} = -s^0 \left(\frac{Cov\left[r\,;\ v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}{\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]} \right) \tag{27}$$

The strict concavity of $v(.)$ guarantees that the covariance term in (42) is strictly negative. So $\left.\frac{\partial \mathscr{B}(y,\tau)}{\partial \tau}\right|_{\tau=0} > 0$: the introduction of τ yields the same utility but increases government revenue. Hence the following proposition :

Proposition 3. *Introducing a linear tax on capital income in an economy without capital taxation generates a first-order welfare improvement.*

Proposition 3 implies that capital taxation, even in its less sophisticated form which is linear taxation, is required at the optimum. In an environment where capital income is observed but not savings, the government ignores if a high capital income is due to high savings or to a lucky draw of r. Therefore, it is impossible to fully insure agents against the volatility of returns. Nevertheless the impossibility to reach the second-best optimum described in proposition 1 does not mean that the government should only rely on labor income taxation: here a linear tax on capital partially insure taxpayers against the risk on returns without discouraging savings too much. The key intuition is that capital income taxation, as established since Domar and Musgrave (1944), 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 τ has on utility. This is why the difference between what the government gets, *i.e* $\overline{r}(s)s$, and what the government loses, through the decrease in labor income taxes $\frac{\partial \hat{T}(y,\tau)}{\tau}$, is positive.

6 The Case of an ex-post Wealth Tax

So far, I have shown that taxing only savings through an ex-ante wealth tax is not optimal while taxing only capital income is actually welfare-improving. An intermediary case arises when the government taxes ex-post wealth $(1+r)s$. Indeed, ex-post wealth is the sum of savings s and capital income rs. However, if the government only observes this sum $(1+r)s$, it is not possible to distinguish between what results from an effort made by taxpayers and what is random. Although a tax on $(1+r)s$ *a priori* shares the social insurance properties of capital income taxation, it adds some undesired effects on incentives by also taxing savings. The question is therefore to understand if these additional adverse effects of taxing $(1+r)s$ instead of rs makes a tax on ex-post wealth irrelevant at the optimum. To do so, I again constrain the information set of the government by assuming that the only form of capital observed by the government is ex-post wealth $(1+r)s$. Capital taxation can therefore only occur through a function $t : (1+r)s \mapsto t\left((1+r)\,s\right)$.

I follow the same route as in Section 5 since I do not try to pin down the optimal tax but I am only interested in knowing if an ex-post wealth tax is even needed at the optimum. To do so, 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 :

$$V^0(y) \overset{\text{def}}{\equiv} \max_s \quad u\left(y-s\right) + \mathbb{E}\left[v\left(\left(1+r\right)s - T^0(y)\right) \mid s\right] \tag{28}$$

Denote by s^0 the solution of (28).

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

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

$$\hat{V}(y,\tau) \overset{\text{def}}{\equiv} \max_{s} \quad u\left(y-s\right)+\mathbb{E}\left[v\left(\left(1-\tau\right)\left(1+r\right)s-\widehat{T}\left(y,\tau\right)\right)\mid s\right] \tag{29}$$

Denote by $\hat{s}(\tau)$ the solution of (29). Under the candidate $\{\widehat{T}(.)\ ;\ \tau\}$ the expected fiscal contribution of a taxpayer with labor income y is given by :

$$\mathscr{B}\left(y,\tau\right)=\widehat{T}\left(y,\tau\right)+\tau.\left(1+\bar{r}\left(\hat{s}\left(\tau\right)\right)\right).\hat{s}(\tau)$$

Note that $\widehat{T}(y,0)=T^0(y)$ and $\hat{s}(0)=s^0$. Using the same method as in Section 5, I show in Appendix A.4 that the introduction of a linear tax on ex-post wealth has the following impact on government revenue:

$$\left.\frac{\partial\mathscr{B}(y,\tau)}{\partial\tau}\right|_{\tau=0}=-s^0\left(\frac{Cov\left[r\ ;\ v'\left(\left(1+r\right)s^0-T^0(y)\right)\mid s^0\right]}{\mathbb{E}\left[v'\left(\left(1+r\right)s^0-T^0(y)\right)\mid s^0\right]}\right) \tag{30}$$

Again, the concavity of $v(.)$ guarantees that the covariance term in (30) is negative so $\left.\frac{\partial\mathscr{B}(y,\tau)}{\partial\tau}\right|_{\tau=0}>0$. Hence the following proposition :

Proposition 4. *Introducing a linear ex-post wealth tax in an economy without capital taxation generates a first- order welfare improvement*

Proposition 4 is analogous to proposition 3 since both advocates for positive capital taxation at the optimum. Combining equation (26) with equation (44) in A.4, capital income and ex-post wealth taxation can be linked through :

$$\left.\frac{\partial\widehat{T}\left(y,\tau\right)^{\text{ExPostWealth}}}{\partial\tau}\right|_{\tau=0}=-s^0+\left.\frac{\partial\widehat{T}\left(y,\tau\right)^{\text{CapIncome}}}{\partial\tau}\right|_{\tau=0} \tag{31}$$

This relation is not surprising : a tax on ex-post wealth $(1+r)s$ is just a combination of a tax on capital income rs and a tax on savings s. Therefore, to keep utility unaffected by the ex-post wealth tax, the government can use labor income taxation to :

- cancel the impact of the "saving" tax induced by ex-post wealth taxation : this is the $-s^0$ term in equation (31). In a sense, this adjustment removes the superfluous component of ex-post wealth taxation. Indeed, as documented in Section 4, taxing savings is of no use in our framework.

- cancel the impact of the "capital income" tax induced by ex-post wealth taxation : this is the $\left.\frac{\partial \widehat{T}(y,\tau)^{\text{CapIncome}}}{\partial \tau}\right|_{\tau=0}$ term in (31).

Doing this adjustment through labor income taxation, the government can expect the same gains from taxing ex-post wealth than from taxing capital income. This equivalence is obvious when comparing (30) with (27) : ex-post wealth and capital income taxation generates exactly the same first-order improvement in government revenue. This can prove important in practice if for political, legal or other institutional reasons the government is biased towards either capital income or ex-post wealth taxation as our results suggest that these two tools produce similar first-order welfare gains.

7 Conclusion

This paper provides an efficiency rationale for taxing capital to insure agents against risky returns without distorting their saving decision. Such insurance could be either provided by a tax on capital income or on ex-post of wealth. However, taxing ex-ante wealth is inefficient. The possibility for wealthier individual to access higher returns does not *per se* provide a rationale for capital taxation. This is not surprising given the Atkinson and Stiglitz (1976) framework studied in this paper : savings are just a residual of labor income and redistribution should therefore occur through labor income taxation. Nevertheless, given the observed strong dispersion in returns at both the top and the bottom of the wealth distribution, a capital tax in the spirit of the perfect insurance optimum described in proposition 1 could mitigate wealth inequality. Although randomness could explain a large share of the observed dispersion in returns at the bottom of the wealth distribution, this narrative seems incomplete at the top. Indeed, Bach et al. (2020) attributes a large share of the observed dispersion in returns at the top of the wealth distribution to entrepreneurial activity. In this context, different returns for a same amount of initial savings could be attributed not only to luck but also to entrepreneurial effort or entrepreneurial skills. These features of the wealth accumulation process are not taken into account in this paper. Conceptually, the frontier between labor and capital income for entrepreneurs is blurry and would require a specific modeling strategy to assess the respective role of not only labor and capital taxation but also corporate taxation.

References

A. B. Atkinson and J. E. Stiglitz. The design of tax structure: direct versus indirect taxation. *Journal of public Economics*, 6(1-2):55–75, 1976.

L. Bach, L. E. Calvet, and P. Sodini. Rich pickings? risk, return, and skill in household wealth. *American Economic Review*, 110(9):2703–47, 2020.

J. Benhabib and A. Bisin. Skewed wealth distributions: Theory and empirics. *Journal of Economic Literature*, 56(4):1261–91, 2018.

J. Benhabib, A. Bisin, and S. Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1):123–157, 2011.

R. Boadway and K. Spiritus. Optimal taxation of normal and excess returns to risky assets. *Tinbergen Institute Discussion Paper 2021-025/VI*, 2021.

J. I. Bulow and L. H. Summers. The taxation of risky assets. *Journal of Political Economy*, 92(1): 20–39, 1984.

E. D. Domar and R. A. Musgrave. Proportional income taxation and risk-taking. *The Quarterly Journal of Economics*, 58(3):388–422, 1944.

A. Fagereng, L. Guiso, D. Malacrino, and L. Pistaferri. Heterogeneity and persistence in returns to wealth. *Econometrica*, 88(1):115–170, 2020.

E. Farhi and I. Werning. Estate taxation with altruism heterogeneity. *American Economic Review*, 103(3):489–95, 2013.

X. Gabaix, J.-M. Lasry, P.-L. Lions, and B. Moll. The dynamics of inequality. *Econometrica*, 84(6): 2071–2111, 2016.

F. Gahvari and L. Micheletto. Capital income taxation and the atkinson–stiglitz theorem. *Economics Letters*, 147:86–89, 2016.

A. Gerritsen, B. Jacobs, A. V. Rusu, and K. Spiritus. Optimal taxation of capital income with heterogeneous rates of return. Technical report, CESifo Working Paper, 2020.

R. H. Gordon. Taxation of corporate capital income: tax revenues versus tax distortions. *The Quarterly Journal of Economics*, 100(1):1–27, 1985.

F. Guvenen, G. Kambourov, B. Kuruscu, S. Ocampo-Diaz, and D. Chen. Use it or lose it: Efficiency gains from wealth taxation. Technical report, National Bureau of Economic Research, 2019.

L. Kaplow. Taxation and risk taking: a general equilibrium perspective. Technical report, National Bureau of Economic Research, 1991.

L. Kaplow. On the undesirability of commodity taxation even when income taxation is not optimal. *Journal of Public Economics*, 90(6-7):1235–1250, 2006.

L. Kaplow. *The theory of taxation and public economics*. Princeton University Press, 2010.

N. R. Kocherlakota. Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. *Econometrica*, 73(5):1587–1621, 2005.

H. Konishi. A pareto-improving commodity tax reform under a smooth nonlinear income tax. *Journal of Public Economics*, 56(3):413–446, 1995.

A. S. Kristjánsson. Optimal taxation with endogenous return to capital. Technical report, Memorandum, 2016.

G. R. Laroque. Indirect taxation is superfluous under separability and taste homogeneity: a simple proof. *Economics Letters*, 87(1):141–144, 2005.

J. Mirrlees, S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles, and J. Poterba. The mirrlees review: conclusions and recommendations for reform. *Fiscal Studies*, 32(3):331–359, 2011.

J. A. Mirrlees. An exploration in the theory of optimum income taxation. *The review of economic studies*, 38(2):175–208, 1971.

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

E. Saez and G. Zucman. Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics*, 131(2):519–578, 2016.

J. E. Stiglitz. The effects of income, wealth, and capital gains taxation on risk-taking. In *Stochastic Optimization Models in Finance*, pages 291–311. Elsevier, 1975.

H. R. Varian. Redistributive taxation as social insurance. *Journal of public Economics*, 14(1):49–68, 1980.

A Appendix

A.1 The Discount Factor For Government Revenue

An agent with type θ faces the following budget constraint :

$$\begin{aligned} c_1(\theta) &= y(\theta) - T\left(y\left(\theta\right)\right) - s\left(\theta\right) \\ c_2(\theta, r) &= (1+r)s(\theta) - t\left(s\left(\theta\right), rs\left(\theta\right)\right) \end{aligned} \tag{32}$$

Assuming that public goods E are financed at time 1, the first-period aggregate resource constraint is given by :

$$\int_{\theta\in\Theta} c_1\left(\theta\right) + s\left(\theta\right) dG\left(\theta\right) + E \leq \int_{\theta\in\Theta} y\left(\theta\right) dG\left(\theta\right) \tag{33}$$

while the second-period aggregate resource constraint is :

$$\begin{aligned} &\iint_{\theta\in\Theta; r\in\mathbb{R}} c_2\left(\theta, r\right) f(r)\, drdG\left(\theta\right) \leq \iint_{\theta\in\Theta; r\in\mathbb{R}} \left(1+r\right) s\left(\theta\right) f(r) dr\, dG\left(\theta\right) \\ \Leftrightarrow &\iint_{\theta\in\Theta; r\in\mathbb{R}} c_2\left(\theta, r\right) f\left(r\right) drdG\left(\theta\right) \leq \int_{r\in\mathbb{R}} \left(1+r\right) f(r) dr \int_{\theta\in\Theta} s\left(\theta\right) dG\left(\theta\right) \end{aligned} \tag{34}$$

where the last line follows from the independence between the distribution of skills θ and the distribution of rates of return r. Therefore the intertemporal aggregate resource constraint is given by :

$$\frac{1}{1+\overline{r}} \iint_{\theta\in\Theta; r\in\mathbb{R}} c_2\left(\theta, r\right) f(r) dr\, dG\left(\theta\right) \leq \int_{\theta\in\Theta} y\left(\theta\right) - c_1\left(\theta\right) dG\left(\theta\right) - E \tag{35}$$

with $\overline{r} = \int_{r\in\mathbb{R}} rdF\left(r\right)$ the average rate of return.

Plugging the budget constraint of taxpayers in (35) yields the government intertemporal budget constraint :

$$B = \int_{\theta\in\Theta} T\left(y\left(\theta\right)\right) dG\left(\theta\right) + \frac{1}{1+\overline{r}} \iint_{\theta\in\Theta; r\in\mathbb{R}} t\left(s\left(\theta\right), rs\left(\theta\right)\right) f(r) dr\, dG\left(\theta\right) \geq E \tag{36}$$

where B stands for government revenue. Hence, because of the independence between skills and rates of return, the government should discount future revenue at rate $\frac{1}{1+\overline{r}}$.

A.2 Proof of Equation 14

The Lagrangian associated to (13) is:

$$L \equiv (1+\bar{r}(s))\, s - \int\limits_{r\in\mathbb{R}} c_2(r) f(r|s) dr + \lambda \left(u(y-s) + \int\limits_{r\in\mathbb{R}} v\left(c_2(r)\right) f(r|s) dr - V^0(y) \right) \tag{37}$$

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

$$\lambda v'\left(c_2\left(r\right)\right) = 1 \tag{38}$$

$v'(.)$ is monotonic so one can rewrite (38) as :

$$c_2(r) = (v')^{-1}(\frac{1}{\lambda}) \tag{39}$$

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 :

$$1+\bar{r}(s)+\bar{r}'(s)s = \frac{\partial}{\partial s}\left(\int\limits_{r\in\mathbb{R}} c_2(r) f(r|s) dr \right) - \lambda_1 \left[\frac{\partial}{\partial s} \left(\int\limits_{r\in\mathbb{R}} v\left(c_2(r)\right) f(r|s) dr \right) - u'\left(y-s\right) \right] \tag{40}$$

Plugging (38) in (40) and using $c_2(r) = c_2$ at the optimum yields :

$$1+\bar{r}(s)+\bar{r}'(s)s = \frac{\partial}{\partial s}\left(\int\limits_{r\in\mathbb{R}} f(r|s) dr \right) c_2 - \frac{1}{v'(c_2)} \left[\frac{\partial}{\partial s} \left(\int\limits_{r\in\mathbb{R}} f(r|s) dr \right) v(c_2) - u'\left(y-s\right) \right] \tag{41}$$

By definition of $f(.)$ as a (conditional) density function, $\int\limits_{r\in\mathbb{R}} f(r|s)dr = 1$. So eventually the integral terms drop and equation (41) boils down to (14) :

$$u'\left(y-s\right) = \left(1+\bar{r}(s)+\bar{r}'(s)\, s\right) v'(c_2)$$

A.3 Proof of equation 24

Plugging (26) in (24) gives :

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

$$= s^0\left(\frac{\bar{r}\left(s^0\right)\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right] - \mathbb{E}\left[rv'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}{\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}\right)$$

$$= -s^0\left(\frac{Cov\left[r\,;\; v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}{\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}\right) \tag{42}$$

A.4 Proof of equation 30

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

$$u\left(y - \hat{s}\left(\tau\right)\right) + \mathbb{E}\left[v\left((1-\tau)\left(1+r\right)\hat{s}\left(\tau\right) - \widehat{T}\left(y,\tau\right)\right) \mid \hat{s}\left(\tau\right)\right] = V^0(y) \tag{43}$$

Applying the envelope theorem to (23) yields :

$$\frac{\partial\widehat{T}\left(y,\tau\right)}{\partial\tau} = -\hat{s}(\tau)\frac{\mathbb{E}\left[(1+r)\,v'\left((1+r)\,(1-\tau)\,\hat{s}\left(\tau\right) - \widehat{T}\left(y,\tau\right)\right) \mid \hat{s}(\tau)\right]}{\mathbb{E}\left[v'\left((1+r)\,(1-\tau)\,\hat{s}\left(\tau\right) - \widehat{T}\left(y,\tau\right)\right) \mid \hat{s}(\tau)\right]} \tag{44}$$

Evaluating (44) at $\tau = 0$ before plugging it in (30) gives :

$$\begin{aligned}
\left.\frac{\partial\mathscr{B}(y,\tau)}{\partial\tau}\right|_{\tau=0} &= s^0\left(1 + \bar{r}\left(s^0\right) - \frac{\mathbb{E}\left[(1+r)\,v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}{\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}\right)\\
&= s^0\left(\frac{\left(1+\bar{r}\left(s^0\right)\right)\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right] - \mathbb{E}\left[(1+r)\,v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}{\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}\right)\\
&= s^0\left(\frac{\bar{r}\left(s^0\right)\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right] - \mathbb{E}\left[rv'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}{\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}\right)\\
&= -s^0\left(\frac{Cov\left[r\,;\; v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}{\mathbb{E}\left[v'\left((1+r)\,s^0 - T^0(y)\right) \mid s^0\right]}\right)
\end{aligned} \tag{45}$$