

# Optimal Taxation of Normal and Excess Returns to Risky Assets

Robin Boadway, Kevin Spiritus\*

March 15, 2021

## Abstract

We study the optimal taxation of risk-free and excess capital income with heterogeneous rates of return, alongside an optimal nonlinear earnings tax. Households can hold three assets: one risk-free, one risky but diversifiable, and one a private investment with idiosyncratic risk whose expected return differs among households. Contrary to expectations, the optimal tax on excess returns to risky assets is ineffective for redistribution, because its effects are annulled by a Domar-Musgrave effect. It assumes only an insurance role, and is positive. The optimal tax on risk-free returns does fulfill a redistributive role, insofar the risk-free returns reveal information about the investors' types beyond what is revealed by the earnings tax base. The optimal nonlinear earnings tax takes the standard Mirrleesian form amended to take account of the stochasticity of capital income tax revenue.

Keywords: optimal capital taxation, Rate-of-Return Allowance, risk, excess returns

JEL Classification: H21, H23, H24

---

\*Boadway: Department of Economics, Queen's University, K7L 3N6 Kingston, Ontario, Canada, e-mail: boadwayr@econ.queensu.ca. Spiritus: Department of Economics, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands, e-mail: spiritus@ese.eur.nl. We are grateful to Vidar Christiansen, André Decoster, Bas Jacobs, Laurence Jacquet, Jukka Pirttilä, Dirk Schindler, Erik Schokkaert and Jo Van Biesebroeck for their time and detailed insights. We also thank Bart Capéau, Annelies Casteleyn, Aart Gerritsen, Louis Kaplow, Etienne Lehmann and Erwin Ooghe for valuable comments, as well as participants to numerous seminars and conferences. Kevin Spiritus acknowledges the financial support of the Belgian Federal Science Policy Office (BELSPO) via the BRAIN.be project BR/121/A5/CRESUS, and of the Flemish Research Foundation (FWO) via travel grant V464515N.

# 1 Introduction

The consensus among public economists is that capital income should be taxed at positive rates, despite suggestions in the earlier literature of the contrary.<sup>1</sup> Recently this view has been strengthened by the observation that rates of return to capital differ among households, with some households obtaining above-normal returns on their investments.<sup>2</sup> Heterogeneous rates of return lead to excess returns, which may reflect rents, luck or differences in investment ability. This observation has led some observers to recommend a differential tax treatment of normal versus excess capital returns.

The Mirrlees Review (2011) proposes a Rate-of-Return Allowance (RRA) that would fully exempt all risk-free returns from taxation, while taxing all excess returns at the same marginal rate as labour income.<sup>3</sup> A recent report by the Institute for Fiscal Studies reinforces this reasoning, applying it to the taxation of business owner-managers (Adam and Miller (2021)). In the corporate tax literature, the differential taxation of rents versus normal capital income has long been proposed, beginning especially with the recommendation for an Allowance for Corporate Equity, which is equivalent to a tax on rents and risk. (Gammie (1991); Mirrlees Review, 2011). More recently, proposals for cash-flow taxation accomplish the same result. For example, Auerbach et al. (2017) recommend the use of Destination-Based Cash Flow Taxation for multinational corporations, fully exempting the normal returns to capital from taxation. Devereux et al. (2019) acknowledge the fairness problem involved in taxing only at destination. Their Residual Profit Allocation by Income (RPA-I) proposal allows for routine profits to generate tax revenue where actual economic activity takes place, while still taxing residual profits at destination. Some Scandinavian countries already partly implement the recommendation to focus on excess returns, adding returns that surpass a normal rate to the labour income tax base, while taxing normal returns at a lower rate.<sup>4</sup>

A common thread underlying the above policy recommendations appears to be that taxing excess returns with perfect loss offsets, would redistribute the heterogeneous rates of return and prevent tax avoidance, without discouraging risk. Taxing normal rates of return, on the other hand, would impose additional efficiency losses, without offering distributional benefits over other tax bases.

In this paper, we revisit the case for taxing excess returns and normal returns to capital at different rates, in a context in which excess returns are risky. We find that, contrary to expectations, the taxation of excess returns does not reduce the after-tax excess returns that investors obtain, so it cannot serve a redistribution role. The intuition is that as a tax on excess returns with perfect loss offsets reduces the after-tax risk for investors,

---

<sup>1</sup>The arguments are reviewed by Banks and Diamond (2010), Bastani and Waldenström (2020) and Scheuer and Slemrod (2021).

<sup>2</sup>Fagereng et al. (2020) provide direct evidence, showing a presence of persistent, individual-specific rates of return that cannot be explained by differences in the allocation of wealth between risk-free and risky assets. Piketty (2014) finds that universities with larger endowments obtain higher rates of return on their investments, and Saez and Zucman (2016) report similar findings for foundations in the United States. Gabaix et al. (2016), Lusardi et al. (2017), Kacperczyk et al. (2018) and Benhabib et al. (2019) provide indirect evidence, showing that the observed wealth inequality can only be explained using life-cycle models if one accounts for the presence of heterogeneous rates of return.

<sup>3</sup>The Mirrlees Review (2011) states that the empirical relevance of known violations of the Atkinson-Stiglitz (1976) theorem is not clear, with doubts arising even about the optimal sign of the tax on normal returns to capital. They then recommend to not tax normal returns at all, and to focus on the excess returns.

<sup>4</sup>Norway applies a Rate of Return Allowance in its treatment of owner-managers of businesses. Finland taxes profits distributed by closely-held companies at higher rates if they exceed 8% of the equity value.

they respond by expanding their investments in risky assets. Following Domar and Musgrave (1944), we show that a tax on excess returns then does not affect the expected utilities of the individuals. The government is thus unable to use the tax on excess returns to redistribute between individuals with different expected rates of return. Instead, the main reason to tax excess returns is to provide insurance against risky returns to capital. The government then turns to the tax on normal returns for redistributive purposes, insofar the normal returns include information about the investment productivities of the individuals that is not already revealed by the labour income tax base. In doing so, the government balances the distortions caused by such a tax, against the distributional benefits.

**Our Model** We adopt an optimal income tax approach to study the optimal taxation of risk-free normal returns and risky excess returns to capital alongside a nonlinear earnings tax. We allow for symmetric loss offsets in the tax on excess returns.<sup>5</sup> An important consideration is that the government cannot separate the different components of the excess returns. It is practically impossible to tax rents, which may be desirable on efficiency grounds, without at the same time taxing returns to risk. To confront this issue, we allow for three types of assets: a risk-free asset, risky assets that yield a competitive return and face aggregate risk, and those that yield an idiosyncratic and possibly above-normal return. We assume the latter assets are personal investments whose expected returns vary among individuals, but which cannot be insured or diversified on the market. Individuals then differ in two dimensions, labour skills as in the standard optimal income tax literature and investment productivities, which we assume are not perfectly correlated. In a first-best world, lump-sum taxes would differentiate according to these two productivity types, but informational limitations faced by the government preclude that. In our second-best setting, a nonlinear income tax is imposed on labour income, and separate linear taxes are applied to the risk-free component of all assets and to the excess return on the risky assets.

Since the empirical literature does not yet offer clear insights into the origins of the observed heterogeneous rates of return to capital, we remain agnostic on this front. We assume that the expected rates of return to the risky assets are linear and differ exogenously among individuals. The presence of risk suffices to limit the amounts that individuals are prepared to invest into the higher-yielding assets. To prevent the government from fully insuring all risk in the optimum, we follow the approach of Christiansen (1993) and Schindler (2008), assuming that the government returns the risk of its revenues to the tax payers using a stochastic public good. Since there is then a social cost associated with taxing the excess returns, the government is unwilling to fully tax them away. Different modelling assumptions could be used to prevent full insurance of the risky capital incomes, e.g. including decreasing returns to scale or incentive effects of taxing excess returns, in which case a tax on excess returns would yield additional distortions. Our setup however allows us to most cleanly illustrate the main mechanisms at work, because it allows us to apply standard portfolio theory. Alternatively, we follow Atkinson and Stiglitz (1980) and Gordon (1985), by assuming that stochastic revenues are returned

---

<sup>5</sup>Adam and Miller (2021) present equivalent methods to implement a tax on excess returns. An expenditure tax, for example, allows taxing the excess returns without identifying the risk-free returns. With an expenditure tax, normal returns to capital are exempted by design.

to individuals as state-contingent equal per capita lump sums. Qualitatively similar results apply in this case.

A problem in a framework with many periods and yearly taxation, would be that the normal returns to capital at any point in time, partly reflect excess returns to capital in the past. Since we wish to clearly distinguish between taxes on normal and excess returns to capital, we work in a two-period life-cycle setting. Individuals supply labour in the first period, and they save part of their labour incomes into their portfolio to finance consumption over the two periods. We adopt assumptions that would lead to zero taxation of capital income in a risk-free setting with no excess returns (Atkinson and Stiglitz, 1976). Taxpayers thus have identical preferences, which are weakly separable between goods and labour.

**Main Results** We first show that the Domar-Musgrave (1944) result remains valid, despite the inclusion of variable labour supply, heterogeneous rates of return and different rates of tax on normal and excess returns: individuals alter their portfolio compositions in such a way that their excess capital income moves in inverse proportion with the net of tax rate. The tax on excess returns does not affect the taxpayers' labour supply or the distribution of their consumption in either period, so it does not affect the expected utilities of the individuals. The optimal tax on the excess returns does not depend on the social welfare weights of the individuals: the presence of heterogeneous rates of return does not yield an additional distributional reason to tax the excess returns.

While the tax on excess returns cannot address redistributive concerns, the tax on risk-free returns will generally have redistributive consequences. If individuals differed only in labour skills, then under some common assumptions (e.g. when the Mean-Variance framework applies or when there is only one type of risk) the tax on risk-free capital income would be zero because of the separability assumptions we have adopted. All redistribution can be accomplished by the nonlinear earnings tax. That is no longer the case when individuals have different investment productivities. The amount of savings and therefore the risk-free capital income of individuals with the same earnings will vary with investment productivity, implying that the tax on risk-free capital income can achieve some redistribution over and above that achieved by the earnings tax.

**Relationship to Existing Literature** The evidence of a positive gradient in the rates of return to capital has sparked a line of research that investigates the implications for the optimal taxation of capital income. Gahvari and Micheletto (2016) and Kristjánsson (2016) study a two-type model where the type with a higher labour-earning ability also has a higher rate of return, finding that the optimal marginal tax on capital income is positive. Gerritsen et al. (2020) study the optimal mix of non-linear taxes on incomes from labour and capital with continuous types. They find that the marginal tax rates on capital income should differ from zero, both when rates of return depend on labour ability and when they depend on the scale of the portfolio. Guvenen et al. (2019) find that a linear wealth tax performs better than a linear capital income tax when investment productivities are heterogeneous. A crucial difference between the papers in this literature and ours, is that we assume that investors do not yet know the returns to their investments when they make their decisions. Furthermore, we allow for a separate tax on excess returns, which avoids distorting the intertemporal allocation of the taxpayers' consumption.

Regarding the taxation of risky capital, the seminal contributions are those of Domar and Musgrave (1944), Mossin (1968) and Stiglitz (1969). In a comparative statics exercise, they study a representative individual who maximizes expected utility by allocating savings over a safe and a risky asset. They show that when the government introduces a linear tax on excess returns to capital, the individual will increase investment in the risky asset. The total after-tax risk will remain unaltered for the investor, as will expected utility. Sandmo (1977) shows that the results of Domar and Musgrave (1944) remain valid when investors optimize a portfolio consisting of multiple risky assets. Gordon (1985) and Kaplow (1994) show how the neutrality of capital income taxes for risk taking remains true in general equilibrium models.

One should not conclude from the Domar-Musgrave (1944) result that the government can costlessly tax all excess returns at 100%, and then obtain a welfare gain by redistributing the revenue to the taxpayers. Buchholz and Konrad (2014) summarize the consequences of taxing idiosyncratic returns, noting the private sector may choose not to insure all risks for incentive reasons. If this is the case, then government insurance will introduce inefficiencies. Only in specific cases, for example when there are positive spillovers from the activities of inventors or when an asset cannot be traded because it is productive only in the hands of a specific individual, can there be good reasons to provide public insurance.

Another reason why taxing excess returns at very high rates may be undesirable is the presence of aggregate risk. In theory, the government could use the financial markets to smooth the aggregate risk of its revenues over time. In practice however, it is unlikely that the government can do so without limit or without distributional consequences. Christiansen (1993) studies the optimal taxation of capital income when the government returns the risky tax revenues to the households via changes in public goods. He does so in a representative individual model with fixed income and one risky asset with linear returns. In this case, as long as the public good is not a perfect substitute for consumption, stochastic public goods mitigate the consumption risk from holding risky assets. An increase in the tax on capital income reduces uncertain capital income and therefore consumption risk, but it increases risk associated with consuming the public good. Schindler (2008) alters this model to allow for separate taxes on risk-free and excess returns. He finds that risk-free assets should not be taxed in an efficient tax system, while excess returns should be taxed, balancing public risk against private risk. We follow this latter road, assuming that the government returns part of the aggregate risk using a stochastic public good, to prevent the government from taxing the excess returns at 100% in the optimum.

The literature that studies optimal redistributive taxation has neglected the possibility of risky capital income. For example, the recent dynamic optimal tax literature focuses instead on idiosyncratic wage risk. The natural conclusion appears to be that there are good reasons to tax capital income at positive rates for redistributive purposes, and that given the Domar-Musgrave (1944) result, we should not worry too much about discouraging risk. We show however that the shortcomings of this reasoning follow from the Domar-Musgrave (1944) result itself: given that a tax on excess returns does not affect expected utilities, there are no good grounds to expect that in absence of tax avoidance, a tax on excess returns should serve a redistributive purpose.

**Road Map** We proceed as follows. Section 2 outlines a general model with both risky and risk-free assets, in which the risky assets can take two forms. As mentioned, one is the purchase of a market portfolio where competitive capital markets ensure that all remaining risk is aggregate. The other is a personal investment, such as a private business or private equity investments where expected returns differ among individuals and risks are assumed to be idiosyncratic. Individuals differ in wage rates and investment productivities, and face a nonlinear tax on labour income and linear taxes on risk-free and excess capital income. Section 3 solves the individual optimization problem, and finds relevant properties. We derive optimal linear tax rates on risk-free and excess returns to capital, optimal nonlinear tax rates on labour income and public good decision rules in Section 4. We use a solution method suitable for our problem that involves deriving the optimal tax system using a perturbation method analogous to that of Saez (2001), Werquin et al. (2015) and Lehmann et al. (2019). Section 5 considers the case where the government returns the risk to individuals using a stochastic lump sum, and Section 6 concludes.

## 2 The model

### 2.1 Individuals

There is a continuum of individuals endowed with labour earning abilities, or skills,  $w \in [\underline{w}, \bar{w}]$ , which are distributed by the function  $G^w(w)$ . The density function is  $g^w(w) = G_w^w(w)$ , where  $G_w^w(w) \equiv dG^w(w)/dw$ , a convention we adopt throughout the paper. Skills determine the effective labour generated by a given effort. Besides their skill levels, individuals differ in the expected returns on their private investments as discussed below.

Individuals live for two periods. In the first period, they supply labour  $\ell$ , yielding gross labour income (effective labour supply)  $z \equiv w\ell$ , which is non-stochastic and declared to the tax authorities. Labour income is taxed according to the nonlinear tax function  $t^\ell(z)$ , so first-period disposable income is  $z - t^\ell(z)$ . Individuals consume  $c^1$  of their disposable income and save  $s$ , so:

$$c^1 = z - t^\ell(z) - s. \quad (1)$$

Savings  $s$  are invested in three assets: bonds  $b$ , market funds  $f$  and private investment opportunities  $p$  so that:

$$s = b + f + p. \quad (2)$$

Using three assets allows us to clearly separate the different sources of risk. Bonds are risk-free and yield a normal return  $r^b$ , so bond income in period two is  $br^b$ . Market funds yield a stochastic market rate of return  $\tilde{r}^m$ , impacted only by aggregate shocks. We denote stochastic variables by a tilde. Total returns from market funds are then  $f\tilde{r}^m$ . Investment in private investment opportunities  $p$  yields a return  $p(\alpha + \tilde{\varepsilon})$ , where  $\tilde{\varepsilon}$  is an idiosyncratic shock that is independent and identically distributed. The shock  $\tilde{\varepsilon}$  is thus uncorrelated with

the market returns. We follow the common assumption that the idiosyncratic shock  $\tilde{\varepsilon}$  is added linearly to the expected return  $\alpha$ . The assumption that market risks are aggregate while private investment risks are idiosyncratic captures the assumption that capital markets fully insure idiosyncratic risk on market portfolios but private investments are not fully insured.

Individuals differ in their expected rates of return to private investment  $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ , which are distributed by the function  $G^\alpha(\alpha)$ , with density function  $g^\alpha(\alpha)$ . Therefore, they differ ex ante in both their labour productivities and their expected rates of return. We denote the two-dimensional types as  $\boldsymbol{\omega} \equiv (w, \alpha)$ , with domain  $\Omega \equiv [\underline{w}, \bar{w}] \times [\underline{\alpha}, \bar{\alpha}]$ . Individuals of the same type  $\boldsymbol{\omega}$  make the same decisions. We denote the joint cumulative distribution of skills and expected returns to private investment by  $G^\omega(\boldsymbol{\omega})$ , with density function  $g^\omega(\boldsymbol{\omega})$ . We make no assumptions about the joint distributions of labour earning abilities  $w$  and returns to private investment  $\alpha$ : they could be strongly correlated, but they could also be entirely independent.

The rate  $\tilde{r}^m$  is drawn from a distribution function  $G^m(r^m)$  with density function  $g^m(r^m)$ . The domain for these functions is denoted  $\mathcal{R}$ . The realizations of the stochastic shocks  $\tilde{\varepsilon}$  are drawn from a distribution function  $G^p(\varepsilon)$ , with density function  $g^p(\varepsilon)$ . The domain for these distribution functions is denoted  $\mathcal{E}$ . When choosing labour supply, savings and portfolio composition, the individuals know their skills  $w$ , their expected returns to private investment  $\alpha$  and the distribution functions  $G^m$  and  $G^p$  of the capital income shocks, but not the realizations of the shocks  $\tilde{r}^m$  and  $\tilde{\varepsilon}$ . Individuals who are ex ante equal differ ex post in the realizations of the idiosyncratic shock  $\tilde{\varepsilon}$ .

Total capital income in the second period is denoted by  $\tilde{y} \equiv (s - f - p)r^b + f\tilde{r}^m + p(\alpha + \tilde{\varepsilon})$ . For tax purposes it is split in two separately declared components: a risk-free part  $y^n$  at interest rate  $r^b$ , and the remaining excess part  $\tilde{y}^e$  such that  $\tilde{y} \equiv y^n + \tilde{y}^e$ , with:

$$y^n \equiv sr^b, \quad \text{and} \quad \tilde{y}^e \equiv f\tilde{r}^m + p(\alpha + \tilde{\varepsilon}) - (f + p)r^b. \quad (3)$$

Recall that we define *excess returns* to refer to all capital income that deviates from the risk-free return, including risk premiums, rents and stochastic shocks. This corresponds with the definition used in the Mirrlees Report (2011) in their proposed RRA system.

Individuals pay taxes  $t^n y^n$  on the risk-free part of their capital income, and  $t^e \tilde{y}^e$  on excess returns. Second-period consumption equals assets saved plus second period after-tax capital income:

$$\tilde{c}^2 \equiv s + (1 - t^n)y^n + (1 - t^e)\tilde{y}^e. \quad (4)$$

The government chooses a labour income tax function and capital income tax rates in the first period. It obtains labour income tax revenues in the first period and capital income tax revenues in the second, and must satisfy an intertemporal budget constraint described below. If there is aggregate risk in the capital markets, tax revenues will be stochastic. The government returns this risk to the individuals in the second period, using a stochastic provision of a pure public good  $\tilde{P}$ .

Given effective labour supply  $z$ , first-period consumption  $c^1$ , realization of second-period consumption  $c^2$  and the realized level of the public good  $P$ , an individual with skill  $w$  obtains utility:

$$U\left(u(c^1, c^2), \frac{z}{w}, P\right). \quad (5)$$

This utility function displays weak separability between intertemporal consumption allocation and labour effort, so preferences over  $c^1$  and  $c^2$  are independent of  $\ell$ . As well, the public good is weakly separable from all individual choice variables, so does not affect them.<sup>6</sup> For ease of notation, we do not write the corresponding subutility function.

Individuals make all their decisions in the first period and are passive recipients of capital income and the public good in the second period. In particular, they choose their labour supply, first period consumption, total savings and portfolio composition to maximize expected utility, subject to their lifetime budget constraints. Second period consumption is determined as a residual.

## 2.2 The government

Each individual reports three variables to the government: labour income  $z$ , risk-free capital income  $y^n \equiv sr^b$  and excess capital income  $y^e \equiv y - sr^b$ . The government imposes a nonlinear labour income tax  $t^\ell(z)$  and linear taxes  $t^n$  and  $t^e$  on risk-free and excess capital income  $y^n$  and  $y^e$ . It commits to these tax rates, as is standard in the literature. We use linear tax rates on capital income for two reasons. The first is that any curvature of the tax function would affect the concavity of the net-of-tax returns to capital, which would influence portfolio choices by the individuals. The additional behavioural effects caused by the nonlinearity of the tax schedule would complicate our analysis, without adding much intuition. The second reason is that linear taxes on capital income might be levied through financial institutions, easing compliance.

We assume the government cannot observe the allocation of savings among the different types of assets. The reason is that the difference is not always clear. Financial institutions might repackage bundles of assets, for fiscal or other reasons; it is not obvious which types of bonds can really be regarded as risk-free investments; and it is difficult to distinguish between aggregate and idiosyncratic components of risk.

The government has access to the bond market between the periods and it balances its budget over time. The *law of large numbers* ensures that, with sufficiently large populations at each type, the government budget constraint is not affected by the idiosyncratic shocks to private investments. Aggregate shocks do however cause government revenue to be stochastic. The provision of the public good varies with the aggregate shock, according to the intertemporal budget constraint in second-period values:

$$P(\tilde{r}^m) = \iint_{\Omega} [(1 + r^b)t^\ell(z) + t^e E_{\mathcal{E}}[\tilde{y}^e | \tilde{r}^m] + t^n y^n] dG^\omega(\omega), \quad (6)$$

---

<sup>6</sup>The case where preferences over private and public consumption are not separable goes beyond the scope of our paper. We merely use the public good to return the aggregate risk to the individuals. See Schindler (2008) for a discussion of the more general cases. We verify the robustness of our results in Section 5, where the government returns the risk of its revenues through a stochastic lump sum, potentially affecting the decisions of the individuals.



where, recall,  $\Omega$  is the domain of the types  $(w, \alpha)$  and where  $E_{\mathcal{E}}$  denotes an average over the realizations of the private investment risk, conditional on the realization of the market rate of return.<sup>78</sup>

We treat the levels of  $P(\tilde{r}^m)$  to be provided for each potential realization of the market shock as policy instruments in the government's optimization problem. The government takes an ex ante perspective. It sets the tax instruments, together with the spending on the public good  $P(r^m)$ , to maximize the sum of the taxpayers' expected utilities:

$$\max_{t^\ell(\cdot), t^n, t^e, P(\cdot)} \iint_{\Omega} E[\tilde{V}(t^\ell(\cdot), t^n, t^e, P(\cdot))] dG^\omega(\omega), \quad (7)$$

subject to the intertemporal budget constraint (6). Here  $E[\tilde{V}(t^\ell(\cdot), t^n, t^e, P(\cdot))]$  is the expected maximized level of utility of an individual with type  $\omega$ , given government policies (where we have suppressed  $\omega$  for simplicity).<sup>9</sup>

Note that we define social welfare in the first period, while we define the government's budget constraint in the second period. After all, the government commits to its policies in the first period, maximizing *expected* social welfare before the shocks are realized, while public good provision is only set in the second period, to close the budget constraint.

### 3 Individual behaviour

#### 3.1 Parameterizations of the tax instruments

We parameterize the nonlinear policy instruments in a way that allows us to study the comparative statics of individual behaviour. This will allow us to formulate the optimal policies in terms of behavioural elasticities.

We parameterize the labour tax function as follows:<sup>10</sup>

$$\tau^\ell(z|\rho, \sigma) \equiv t^\ell(z) + \sigma z - \rho,$$

where the parameter  $\rho$  shifts the intercept and  $\sigma$  shifts the slope of the tax function. The corresponding marginal tax rate is  $\tau_z^\ell(z|\rho, \sigma) \equiv t_z^\ell(z) + \sigma$ .

To study the comparative statics for the linear tax rates on risk-free and excess capital income, we directly perturb the parameters  $t^n$  and  $t^e$ . We summarize the vector of tax perturbation parameters as  $\vec{\sigma} \equiv (\rho, \sigma, t^e, t^n)$ .

<sup>7</sup>Whenever we use the expectation operator  $E[\cdot]$ , we refer to the average of the operand over all potential realizations of the shocks  $(r^m, \varepsilon) \in \mathcal{R} \times \mathcal{E}$ . I.e. for any stochastic variable  $\tilde{x} \equiv x(\tilde{r}^m, \tilde{\varepsilon})$ , its expected value is defined as  $E[\tilde{x}] \equiv \int_{\mathcal{R}} \int_{\mathcal{E}} x(r^m, \varepsilon) dG^{\mathcal{P}}(\varepsilon) dG^m(r^m)$ . The conditional expectation operator  $E_{\mathcal{E}}[\cdot|r^m]$  denotes an average over all potential realizations  $\varepsilon \in \mathcal{E}$  of private investment risk, conditional on the realization  $r^m$  of market risk:  $E_{\mathcal{E}}[\tilde{x}|r^m] \equiv \int_{\mathcal{E}} x(r^m, \varepsilon) dG^{\mathcal{P}}(\varepsilon)$ . The quantity  $E_{\mathcal{E}}[\tilde{y}^e|\tilde{r}^m]$  still contains aggregate risk, but no idiosyncratic risk.

<sup>8</sup>Note that although the government can borrow freely to balance its budget over the two periods of our model, we assume that it cannot pool aggregate risk over time by borrowing. For example, the aggregate risk might manifest itself as a once-over shock in the second period of our model. To pool that shock would involve borrowing over the indefinite future which would be very demanding.

<sup>9</sup>Here and elsewhere in the paper, when a function argument is a function itself rather than its value, we indicate this by adding  $(\cdot)$  behind the function name. In (7) this convention means that indirect utility depends on the tax system and on the stochastic public good provision in their entirety, rather than on specific values of these functions.

<sup>10</sup>For previous papers that parameterize nonlinear tax schedules to study comparative statics, see e.g. Christiansen (1981) and Jacquet et al. (2013).

### 3.2 Parameterization of the public good

For the stochastic public good, we assume that the government commits to levels of  $P(r^m)$  for each potential realization of the market shock  $\tilde{r}^m$ , such that its intertemporal budget constraint (6) is satisfied. When we discuss a perturbation of the level of  $P(r^m)$  at a specific realization  $r^m$ , we implicitly mean a perturbation of the function  $P$  in a small interval around the realization  $r^m$ . We discuss the perturbations of the function  $P$  more rigorously in Appendix A.1.

The fact that we implement perturbations for different realizations of the market return  $r^m$  does not imply that we redistribute resources between the different states of the world. The government budget constraint needs to be satisfied for each state of the world separately. Our approach differs from that of Christiansen (1993) and Schindler (2008), who in a representative agent model directly substitute the government budget constraint for the public good parameter in the individual's utility function. An advantage of our approach, besides tractability in our more general setting, is that it leads to a stochastic government budget Lagrange multiplier, yielding optimal tax equations that are more straightforward to interpret. We discuss this in more detail in Section 4.

### 3.3 First-order conditions

Consider the individual optimization problem in a situation where the policy reform parameters  $\rho$  and  $\sigma$  are not necessarily zero. Any individual of type  $\omega$  chooses effective labour supply  $z$ , first-period consumption  $c^1$ , total savings  $s$  and investment in assets with excess returns  $f$  and  $p$  to maximize expected utility. Using the utility function (5), the Lagrangian for this maximization problem is:

$$\mathcal{L}(c^1, z, s, f, p | \vec{\sigma}, P(\cdot)) \equiv \mathbb{E} \left[ U \left( u(c^1, c^2(s, f, p | \vec{\sigma}, \tilde{r}^m, \tilde{\varepsilon})), \frac{z}{w}, \tilde{P} \right) \right] - \mu \cdot [c^1 - z + \tau^\ell(z | \rho, \sigma) + s], \quad (8)$$

with budget Lagrange multiplier  $\mu$ , and with second-period consumption obtained from (4) using (3):

$$c^2(s, f, p | \vec{\sigma}, \tilde{r}^m, \tilde{\varepsilon}) = s + (1 - t^n)sr^b + (1 - t^e)[f\tilde{r}^m + p(\alpha + \tilde{\varepsilon}) - (f + p)r^b]. \quad (9)$$

We treat second-period consumption as a residual rather than a choice variable, since its realization is not known at the time of the optimization.

We assume throughout that second-order conditions are met and focus solely on the first-order conditions. To simplify notation, we denote  $\tilde{U}_1 \equiv (\partial \tilde{U} / \partial u) \cdot (\partial u / \partial c_1)$  and  $\tilde{U}_2 \equiv (\partial \tilde{U} / \partial u) \cdot (\partial u / \partial c_2)$ . The first-order conditions on earnings  $z$  and savings  $s$  are standard and are as follows:

$$\frac{\mathbb{E}[\tilde{U}_\ell]}{\mathbb{E}[\tilde{U}_1]} = -(1 - \tau_z^\ell)w, \quad \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} = 1 + (1 - t^n)r^b. \quad (10)$$

From the first-order conditions on portfolio choices  $f$  and  $p$ , we obtain the marginal risk premiums required by

the individuals:

$$(1 - t^e)\mathbb{E}[\tilde{r}^m - r^b] = -(1 - t^e)\frac{\text{cov}(\tilde{U}_2, \tilde{r}^m)}{\mathbb{E}[\tilde{U}_2]}, \quad (11)$$

$$(1 - t^e)(\alpha - r^b) = -(1 - t^e)\frac{\text{cov}(\tilde{U}_2, \tilde{\varepsilon})}{\mathbb{E}[\tilde{U}_2]}. \quad (12)$$

We assume that there is sufficient risk to limit the amounts that individuals are willing to invest in the risky assets. We thus assume an interior solution for the choice between risk-free and risky assets. Combining conditions (11) and (12) yields:<sup>11</sup>

$$(1 - t^e)\left(\mathbb{E}[\tilde{y}^e] + \frac{\text{cov}(\tilde{U}_2, \tilde{y}^e)}{\mathbb{E}[\tilde{U}_2]}\right) = 0. \quad (13)$$

The term  $\mathbb{E}[\tilde{y}^e]$  reflects the expected excess return to capital, which is positive (although some realizations of  $y^e$  are negative). The normalized covariance  $\text{cov}(\tilde{U}_2, \tilde{y}^e)/\mathbb{E}[\tilde{U}_2]$  is negative. Its size reflects the required total risk premium from the risky investments. The left-hand side of (13) reflects the certainty equivalent of the after-tax excess capita incomes. Individuals thus invest in the risky assets up to the point where the certainty equivalent of their excess capital income equals zero.

### 3.4 Unconditional demand and supply functions

The solution to the individual optimization problem (8)–(9) yields uncompensated demand functions for goods and assets and an uncompensated supply function of labour:  $c^1(\vec{\sigma})$ ,  $p(\vec{\sigma})$ ,  $s(\vec{\sigma})$ ,  $f(\vec{\sigma})$  and  $\ell(\vec{\sigma})$ . (We omit type  $\omega$  as an argument for notational convenience.) These functions depend on the individual type and on the reform parameters for the taxes, but by our weak separability assumption not on the public good. We use these uncompensated functions to construct the following uncompensated functions for risk-free and excess capital income:

$$y^n(\vec{\sigma}) = s(\vec{\sigma})r^b, \quad (14)$$

$$\tilde{y}^e(\vec{\sigma}) = f(\vec{\sigma})(\tilde{r}^m - r^b) + p(\vec{\sigma})(\alpha + \tilde{\varepsilon} - r^b). \quad (15)$$

The demand function  $\tilde{c}^2(\vec{\sigma})$  follows by substituting the appropriate demand functions into budget constraint (9).

If for each individual the policy reforms are complemented by lump sum payments such that utility remains constant at level  $\bar{V}$ , then we obtain compensated supply and demand functions, which we denote using a superscript asterisk:  $c^{1*}(\vec{\sigma}, P(\cdot), \bar{V})$ ,  $\ell^*(\vec{\sigma}, P(\cdot), \bar{V})$ ,  $s^*(\vec{\sigma}, P(\cdot), \bar{V})$ ,  $f^*(\vec{\sigma}, P(\cdot), \bar{V})$  and  $p^*(\vec{\sigma}, P(\cdot), \bar{V})$ . The corresponding tax bases are  $z^*(\vec{\sigma}, P(\cdot), \bar{V})$ ,  $y^{n*}(\vec{\sigma}, P(\cdot), \bar{V})$  and  $\tilde{y}^{e*}(\vec{\sigma}, P(\cdot), \bar{V})$ .<sup>12</sup>

<sup>11</sup>To see this, rewrite (11) and (12) as follows:  $(1 - t^e)\mathbb{E}[\tilde{U}_2(\tilde{r}^m - r^b)f] = 0$  and  $(1 - t^e)\mathbb{E}[\tilde{U}_2(\alpha + \tilde{\varepsilon} - r^b)p] = 0$ . Add these two results together and substitute definition (3) to find  $(1 - t^e)\mathbb{E}[\tilde{U}_2\tilde{y}^e] = 0$ . Rearrange to find (13).

<sup>12</sup>We introduce compensated responses more formally in Appendix A.1.

We can now formulate some properties of individual demand. Domar and Musgrave (1944) and Sandmo (1977) derive comparative statics for the effect of capital income tax changes on the demand for risky assets. In the following lemma, we extend their findings to our setting with endogenous labour supply.

**Lemma 1.** *A reform to the tax on excess returns  $t^e$  has the following effects:*

1. *excess capital income moves in inverse proportion with the net of tax rate:*

$$\tilde{y}_{t^e}^e = \frac{\tilde{y}^e}{1 - t^e}. \quad (16)$$

2. *labour income and consumption in either period are not affected by a change in the tax on excess capital income:*

$$z_{t^e} = c_{t^e}^1 = \tilde{c}_{t^e}^2 = 0. \quad (17)$$

3. *individual welfare is not affected by the tax on excess capital income:*

$$E[\tilde{V}_{t^e}] = 0. \quad (18)$$

*Proof.* See Appendix A.2. □

Intuitively, if there is an increase in the tax on excess returns, then without behavioural responses, both the expected value and the standard deviation of net-of-tax excess returns  $(1 - t^e)\tilde{y}^e$  decrease proportionally. As a consequence, the net-of-tax marginal risk premiums required by the individuals (the right-hand sides of (11) and (12)) decline more than the effective decrease in the net-of-tax marginal risk premiums (the left-hand sides of (11) and (12)). The individuals will extend their investments in the risky assets up to the point where their required marginal risk premiums again equal the effective marginal risk premiums — when their after-tax returns  $(1 - t^e)\tilde{y}^e$  and thus second-period consumption  $\tilde{c}^2$  are back at their original levels. Since also labour income and first-period consumption are not affected by a change in the tax on excess returns, a reform to the tax on excess returns does not affect individual welfare.

### 3.5 Demand and supply functions conditional on labour income

Our aim is to characterize the optimal taxes on capital income in the presence of an optimal nonlinear tax on labour income. To do so, it is useful to express the behavioural responses of the individuals conditional on labour supply.<sup>13</sup> Imagine the individual problem as a two-stage optimization problem in which labour supply and therefore earnings are chosen first, and then earnings are allocated to consumption, savings and portfolio composition. We study the problem starting from the second stage. Taking labour supply and first-period

---

<sup>13</sup>Christiansen (1984) shows that the optimal differentiation of commodity taxes balances the additional distributional benefits and efficiency gains against the additional distortionary costs, compared to a tax system that fully relies on the tax on labour income. The relevant comparative statics to characterize the optimum are then conditional on labour income. See e.g. Jacobs and Boadway (2014) and Gerritsen et al. (2020) for other papers that use comparative statics conditional on labour income.

disposable income as given, individuals maximize utility by choosing first-period consumption, savings and portfolio composition. Denote the resulting uncompensated conditional demand functions using a superscript  $\cdot^c$ :  $c^{1c}(\vec{\sigma}|z, \alpha)$ ,  $s^c(\vec{\sigma}|z, \alpha)$ ,  $f^c(\vec{\sigma}|z, \alpha)$  and  $p^c(\vec{\sigma}|z, \alpha)$ , with corresponding capital income tax bases  $y^{nc}(\vec{\sigma}|z, \alpha)$  and  $y^{ec}(\vec{\sigma}|z, \alpha)$ . The conditional demand functions depend on the expected return  $\alpha$ , but they do not depend on labour earnings ability  $w$  due to the separability of preferences between leisure and consumption. In the first stage individuals optimize their overall expected utility by choosing their labour incomes, anticipating the outcome of the resulting second-stage choices.

## 4 Government optimization

We now have all the building blocks to study the government optimization problem (7) subject to budget constraint (6), taking into account the behavioural responses of the individuals. We characterize the optimal policies in Appendix A. To find the optimal linear tax rates  $t^n$  and  $t^e$ , we use the standard approach, demanding that small perturbations of the tax rates do not affect social welfare. To characterize the nonlinear instruments  $t^\ell(\cdot)$  and  $P(\cdot)$ , we use a perturbation approach similar to that introduced by Saez (2001) and formalized by Werquin et al. (2015) and Lehmann et al. (2019). We use the calculus of variations to study the effects of arbitrary perturbations to the nonlinear policy instruments, and we require that the effects of such perturbations on social welfare sum to zero.

A difficulty in constructing the Lagrangian is that the government budget depends on the state of the world. We assume for purposes of analysis that the levels of the public good for each shock are chosen before the realization of the market shock. Thus, there is no single budget multiplier for the government optimization problem. To each realization of the market shock  $r^m$  corresponds a budget multiplier  $\lambda^{r^m}$ . To reflect this, we introduce a *stochastic budget Lagrange multiplier*  $\tilde{\lambda} \equiv \lambda^{r^m}$ . Each realization  $\lambda^{r^m}$  of the stochastic multiplier can be interpreted as the social value of an additional unit of resources in the second period if the realization of the market shock equals  $r^m$ .<sup>14</sup> We obtain the following Lagrangian for the government optimization problem:

$$\begin{aligned} \Lambda(t^\ell(\cdot), t^n, t^e, P(\cdot), \lambda(\cdot)) &= \iint_{\Omega} \mathbb{E}[\tilde{V}(t^\ell(\cdot), t^n, t^e, P(\cdot))] dG^\omega(\omega) - \mathbb{E}[\tilde{\lambda}\tilde{P}] \\ &+ (1 + r^b)\mathbb{E}[\tilde{\lambda}] \iint_{\Omega} t^\ell(z) dG^\omega(\omega) + t^n \mathbb{E}[\tilde{\lambda}] \iint_{\Omega} y^n dG^\omega(\omega) + t^e \iint_{\Omega} \mathbb{E}[\tilde{\lambda}\tilde{y}^e] dG^\omega(\omega). \end{aligned} \quad (19)$$

The arguments of the demand and supply functions  $z$ ,  $y^n$  and  $\tilde{y}^e$  capture the behavioural responses to policy changes. For notational simplicity, we omit these function arguments in (19).

We introduce some notation before showing the optimality conditions. Suppose the government gives an additional unit of income in the first period to an individual of type  $\omega$ . This is equivalent to a change in the intercept of the labour income tax schedule. The effect on social welfare consists of two parts. First, there is a direct effect on the expected utility of this individual. Second, there are income effects on the different

<sup>14</sup>This interpretation of the budget multipliers shows why we choose to write the government's budget constraint (6) in second-period values. Writing (6) in first-period values would not alter our results, but it would complicate our notations.

tax bases, which affect the tax liability. The total effect of the additional unit of income on the government objective (19) equals  $E[\tilde{U}_1 + \tilde{\lambda} \cdot \{(1 + r^b)t_z^\ell z_\rho + t^n y_\rho^n + t^e \tilde{y}_\rho^e\}]$ . Divide by  $E[\tilde{\lambda}]$  to find the monetary value of the effect on social welfare:

$$\beta(\boldsymbol{\omega}) \equiv \frac{E[\tilde{U}_1]}{E[\tilde{\lambda}]} + (1 + r^b)t_z^\ell z_\rho + t^n y_\rho^n + t^e \frac{E[\tilde{\lambda} \tilde{y}_\rho^e]}{E[\tilde{\lambda}]}.$$
 (20)

The term  $\beta(\boldsymbol{\omega})$  indicates the net marginal social utility of income for an individual of type  $\boldsymbol{\omega}$ , following Diamond (1975). The variation of  $\beta(\boldsymbol{\omega})$  with skill captures the benefits of redistributing income between individuals with different earning abilities, whereas the variation of  $\beta(\boldsymbol{\omega})$  with the expected return to private investment captures the benefits of redistributing income from individuals with a high expected return to individuals with a low expected return.

We denote the marginal excess burden of a change in the marginal tax rate on labour income as:

$$\mathcal{W}(\boldsymbol{\omega}) \equiv -(1 + r^b)t_z^\ell z_\sigma^* - t^n y_\sigma^{n*} - t^e \frac{E[\tilde{\lambda} \tilde{y}_\sigma^{e*}]}{E[\tilde{\lambda}]}.$$
 (21)

This marginal excess burden quantifies, in monetary terms, the social welfare loss due to the compensated revenue effects of a small increase in the marginal tax rate on labour income at the income level chosen by a type- $\boldsymbol{\omega}$  individual. The first term equals the loss from compensated responses of labour income, the second and third terms equal the losses from compensated responses of risk-free and excess capital incomes.

Let  $G^z(z)$  denote the cumulative distribution function for labour incomes in the tax optimum, and let  $g^z(z)$  be the density function. Denote the cumulative distribution of the individual types  $\boldsymbol{\omega}$  conditional on labour income  $z$  as  $G^{\boldsymbol{\omega}|z}(\boldsymbol{\omega}|z)$ . For any function  $h(\boldsymbol{\omega})$ , let  $\bar{h}(z) \equiv \iint_{\Omega} h(\boldsymbol{\omega}) dG^{\boldsymbol{\omega}|z}(\boldsymbol{\omega}|z)$  denote the average of  $h(\boldsymbol{\omega})$  for all types  $\boldsymbol{\omega}$  who choose labour income  $z$ .

Following Saez (2002), we assume that among individuals who earn the same labour income, the compensated behavioural responses to a labour tax reform are not systematically correlated with the propensities to save out of labour income.<sup>15</sup> We thus assume that for any tax base  $k = z, y^n, \tilde{y}^e$ :<sup>16</sup>

$$\text{cov}_{\Omega}(k_\sigma^*, y_z^{nc}|z) = 0,$$
 (22)

where  $\text{cov}_{\Omega}(\cdot)$  denotes a population covariance (as opposed to  $\text{cov}(\cdot)$ , which denotes a covariance over the potential realizations of the investment risks).

Given the definitions of  $\beta(\boldsymbol{\omega})$  and  $\mathcal{W}(\boldsymbol{\omega})$ , we derive the first-order conditions for the government optimization problem in Appendix A. Rearranging them yields the following conditions, where the equation labels refer to

<sup>15</sup>Saez (2002) argues that there is no obvious reason why such correlations should exist. Moreover, such systematic relations would be empirically difficult to detect.

<sup>16</sup>A direct equivalent to Saez' (2002) Assumption 2 would be e.g.  $\text{cov}_{\Omega}(z_{\tilde{t}^n}^*, y_z^{nc}|z) = 0$ . Using the Slutsky symmetry between  $z$  and  $y^n$  (see Appendix A.2), this is equivalent to  $\text{cov}_{\Omega}(y_\sigma^{n*}, y_z^{nc}|z) = 0$ . We use the latter formulation in (22) because it leads to simpler equations for stochastic tax bases. Assumption (22) only affects the optimal tax on risk-free returns. Appendix A.5 contains characterizations for the optimal tax on risk-free returns that do not use Assumption (22).

the policy variable being perturbed:

$$t_z^\ell(\cdot) : \quad \bar{W}(z)g^z(z) = \int_z^{\bar{z}} (1 + r^b - \bar{\beta}(z))dG^z(z), \quad (23)$$

$$t^\ell(0) : \quad \iint_{\Omega} \beta(\boldsymbol{\omega})dG^\omega(\boldsymbol{\omega}) = 1 + r^b, \quad (24)$$

$$t^n : \quad t^n \int_z^{\bar{z}} \bar{y}_{t^n}^{nc*} dG^z(z) = \frac{E[\tilde{U}_2]}{E[\tilde{U}_1]} \int_z^{\bar{z}} \text{cov}_\Omega(\beta, y^n|z)dG^z(z) - t^e \int_z^{\bar{z}} \frac{E[\tilde{\lambda}\tilde{y}_{t^n}^{ec*}]}{E[\tilde{\lambda}]} dG^z(z), \quad (25)$$

$$t^e : \quad \iint_{\Omega} \frac{\text{cov}(\tilde{\lambda}, \tilde{y}^e)}{E[\tilde{\lambda}]} dG^\omega(\boldsymbol{\omega}) = \iint_{\Omega} \frac{\text{cov}(\tilde{U}_2, \tilde{y}^e)}{E[\tilde{U}_2]} dG^\omega(\boldsymbol{\omega}), \quad (26)$$

$$P(\cdot) : \quad \lambda^{r^m} = \iint_{\Omega} E_{\mathcal{E}}[\tilde{U}_P|r^m]dG^\omega(\boldsymbol{\omega}). \quad (27)$$

The effect of most policy reforms on tax revenue from excess capital income is uncertain. Definitions (20) and (21) and optimality condition (25) take into account the welfare effects of this uncertainty. Before turning to the interpretation of the optimality conditions, we reflect on how the uncertainty in the government's cost-benefit analysis affects optimal policy decisions.

#### 4.1 Marginal social risk premiums

The optimizing government compares the marginal costs and benefits of using different policy instruments. When there is a marginal reform to any of the policy instruments, the effects will generally include a change in tax revenue from excess capital income. Since the excess returns to capital are uncertain, the effects on government revenue will be uncertain. For this reason, a term  $E[\tilde{\lambda}t^e\tilde{y}_\nu^e]/E[\tilde{\lambda}]$  appears in definitions (20) and (21) and in optimality condition (25). This term is the government's certainty equivalent of the effect of the reform on the tax revenues from the excess returns. It can be decomposed into expectation and risk terms:

$$\frac{E[\tilde{\lambda}t^e\tilde{y}_\nu^e]}{E[\tilde{\lambda}]} = E[t^e\tilde{y}_\nu^e] + \frac{\text{cov}(\tilde{\lambda}, t^e\tilde{y}_\nu^e)}{E[\tilde{\lambda}]}. \quad (28)$$

The normalized covariance in (28) captures the effect of a policy reform on the uncertainty of the government's revenues. The size of this normalized covariance  $\text{cov}(\tilde{\lambda}, t^e\tilde{y}_\nu^e)/E[\tilde{\lambda}]$  is the *marginal social risk premium* associated with the instrument  $\nu$ . The marginal social risk premium is analogous to the private marginal risk premiums in individual first-order conditions (11) and (12). An advantage of this concept is that when different groups of the population are affected in different ways by a given realization of the state of the world, the distributional effects are taken into account in the value of the corresponding government budget multiplier,  $\lambda^{r^m}$ . The marginal social risk premium allows for distribution-sensitive cost-benefit analyses in stochastic environments.

The sign of the term  $\text{cov}(\tilde{\lambda}, t^e\tilde{y}_\nu^e) = t^e f_\nu \text{cov}(\tilde{\lambda}, \tilde{r}^m)$  is opposite to that of  $E[\tilde{y}_\nu^e]$ , since  $\lambda^{r^m}$  rises as  $r^m$  falls and vice versa. It is not clear a priori whether the total impact of the term  $E[\tilde{\lambda}t^e\tilde{y}_\nu^e]$  on social welfare is positive or negative. To further interpret this term, introduce the notation  $\mathcal{A} \equiv -E[\tilde{\lambda}\tilde{y}_\nu^e]/E[\tilde{\lambda}]$ . If there is a perturbation of  $\nu$ , then  $-t^e \iint_{\Omega} \mathcal{A}_\nu dG^\omega(\boldsymbol{\omega})$  captures the effect on social welfare caused by behavioural responses of excess

capital income  $y^e$ . We show in the following lemma how the term  $\iint_{\Omega} \mathcal{A}_{\nu} dG^{\omega}(\omega)$  reflects the change in the relative importance of aggregate risk due to a perturbation of  $\nu$ .

**Lemma 2.** *The effect of any policy reform  $\nu = \sigma, \rho, t^n$  on the expected social value of government revenue from excess returns is determined by the relative shift towards or away from private investment:*

$$t^e \iint_{\Omega} \mathcal{A}_{\nu} dG^{\omega}(\omega) = - \iint_{\Omega} t^e \mathbb{E}[\tilde{r}^m - r^b] f dG^{\omega}(\omega) \cdot \frac{d}{d\nu} \left( \frac{\iint_{\Omega} t^e (\alpha - r^b) p dG^{\omega}(\omega)}{\iint_{\Omega} t^e \mathbb{E}[\tilde{r}^m - r^b] f dG^{\omega}(\omega)} \right).$$

*This result extends to compensated effects and to effects conditional on labour income.*

*Proof.* See Appendix A.9. □

If idiosyncratic risk on average becomes relatively more important, e.g. because entrepreneurs undertake new investments that are not correlated with the market, then the total impact on social welfare is positive. The welfare effect of the increased expected government revenue outweighs the effect of the increased uncertainty, because the sources of government revenue become better diversified. If on the other hand market investment becomes relatively more important, the total impact on social welfare is negative, even if the expected tax revenue increases.

Lemma 2 illustrates the importance of correcting for uncertainty. If we were to ignore uncertainty, then we would simply include *the expected* effect on excess capital income as a sufficient statistic in the government's cost-benefit analysis. After correcting for uncertainty, it is not the expected effect on excess capital income, but the effect on the relative importance of aggregate risk that matters.<sup>17</sup> We show in Appendix B that seemingly innocuous simplifications of our model lead to the conclusion that  $\iint_{\Omega} \mathcal{A}_{\nu} dG^{\omega}(\omega)$  equals zero. This is most obviously the case when there is only idiosyncratic or only aggregate risk. Also, when individuals optimize their portfolios according to the Mean-Variance framework, we find that  $\iint_{\Omega} \mathcal{A}_{\nu} dG^{\omega}(\omega)$  equals zero when expected returns to private investment are equal for all investors, or when the semi-elasticities of  $p$  and  $f$  with respect to  $\nu$  are constant over the income distribution.

In the sections that follow we discuss optimality conditions (23)–(27) for the different policy instruments. We first discuss optimal linear taxes on excess returns in Subsection 4.2, and on risk-free capital income in Subsection 4.3. We treat the optimal nonlinear tax on labour income in Subsection 4.4. We discuss the optimal provisions of public goods in Subsection 4.5. Finally, in Subsection 4.6, we combine our insights to discuss the optimality of the Rate-of-Return Allowance proposed by the Mirrlees Review (2011).

## 4.2 Optimal linear tax on excess capital income

Condition (26) shows how the government should balance the risk of private consumption against the uncertainty of public revenues. To gain further intuition, substitute (27) into (26) to find the following theorem, which characterizes the optimal tax on excess capital income.

<sup>17</sup>Our findings illustrate the warning of Kleven (2020), who notes that the sufficient statistics approach implicitly relies on strong assumptions on preferences and on the decision environment.



**Theorem 3.** *The optimal linear tax on excess capital income balances public consumption risk against private consumption risk:*

$$\frac{\text{cov}(\iint_{\Omega} \tilde{U}_P dG^{\omega}(\omega), \iint_{\Omega} \tilde{y}^e dG^{\omega}(\omega))}{\iint_{\Omega} E[\tilde{U}_P] dG^{\omega}(\omega)} = \iint_{\Omega} \frac{\text{cov}(\tilde{U}_2, \tilde{y}^e)}{E[\tilde{U}_2]} dG^{\omega}(\omega). \quad (29)$$

*The optimal tax on excess returns is not affected by behavioural elasticities or by the social welfare weights.*

As the uncertainty of government revenue increases, also the uncertainty of public good provision increases. If individual preferences are concave in the public good, uncertainty in its provision will have a welfare cost. The optimal tax on excess capital income thus balances the uncertainty of public good provision against the uncertainty of private consumption.

A striking feature of optimality condition (29) is that it does not depend on the social welfare weights  $\beta$ . The optimal tax on excess returns does not serve to redistribute between individuals based on ex ante differences, such as ability or expected rates of return to capital. Lemma 1 shows why this is the case: individuals respond to a tax increase on excess returns by proportionally increasing their investments in the risky assets, such that their expected utility remains unaltered. A tax on excess returns is unable to redistribute welfare differences that stem from ex ante characteristics.

There is another way to understand the absence of the social welfare weights in (29). Recall from (7) that the government cares about the distribution of *expected* utilities. Thus, what matters from a distributional perspective are the certainty equivalents of the excess capital incomes. By (13) the certainty equivalents of excess capital income are equal to zero for all optimizing individuals. Hence, for the tax on excess returns, there is no scope for redistribution based on ex ante characteristics.

We now show that  $t^e$  should be strictly between zero and unity. Suppose first that  $t^e = 0$ . If there is a high realization of the market rate of return, excess returns  $y^e$  will be high, and thus second-period consumption  $c^2$  will be high. The marginal utility of second-period consumption  $U_2$  will be low. Conversely, when there is a low realization of the market rate of return, the marginal utility of second-period consumption will be high. It follows that the covariance on the right-hand side of (29) is strictly negative. The left-hand side is zero, since without a tax on excess returns government revenues are not stochastic. Therefore,  $t^e = 0$  cannot be optimal.

Similarly, we can exclude the case  $t^e < 0$ , because in this case the covariance on the left-hand side would be positive: a high realization of the market rate of return would lead to low government revenues and to a higher value of  $\lambda^m$ . Similar reasoning shows that the only remaining possibility, a positive tax on excess returns  $t^e > 0$ , yields the correct signs for both covariances, as long as  $t^e$  does not become larger than 100%, in which case the covariance on the right-hand side of (29) becomes positive.

To see that  $t^e$  should be strictly smaller than 100%, suppose that  $t^e$  approaches unity from below and that excess returns are all taxed away. The covariance on the right-hand side of (29) approaches zero. The size of the covariance on the left-hand side on the other hand increases: the government budget is more responsive to market shocks when taxes on excess returns are high. As long as individuals are averse towards risk in the provision of the public good, the optimal tax on excess returns is thus strictly smaller than 100%. We summarize

these findings in the following corollary.

**Corollary 4.** *If all risk is returned using uncertain provision of a public good, then the optimal tax rate on excess capital returns is strictly positive, and strictly smaller than 100%:*

$$0 < t^e < 1.$$

*If the importance of aggregate risk increases, then the optimal tax rate on excess returns decreases.*

### 4.3 Optimal linear tax on risk-free capital income

Substitute the definition of  $\mathcal{A}$  into condition (25) to find the optimal tax on risk-free capital income in the following Theorem.

**Theorem 5.** *The optimal tax on risk-free capital income balances its distortions against its distributional benefits:*

$$t^n \int_{\underline{z}}^{\bar{z}} \bar{y}_{t^n}^{nc*} dG^z(z) = \frac{E[\tilde{U}_2]}{E[\tilde{U}_1]} \int_{\underline{z}}^{\bar{z}} \phi^n(z) dG^z(z) + t^e \int_{\underline{z}}^{\bar{z}} \mathcal{A}_{t^n}^{c*} dG^z(z), \quad (30)$$

where  $\phi^n(z)$  denotes the distributional characteristic of risk-free capital income conditional on labour income:

$$\phi^n(z) \equiv \text{cov}_\Omega(\beta, y^n | z).$$

Condition (30) shows that the optimal tax on risk-free capital income depends on three effects: the compensated effects of a change in  $t^n$  on a) tax revenue from risk-free capital income, b) the relative importance of aggregate risk, and c) the covariance of the social welfare weights  $\beta$  with the risk-free returns to capital  $y^n$ , all conditional on labour income  $z$ . Given that we assume that preferences are separable between consumption and leisure, there are no Corlett and Hague (1953) motives to tax risk-free capital income. The absolute value of the tax on risk-free capital income should be larger if the compensated responses of risk-free capital income are smaller. The integral on the left-hand side of (30) is negative: a compensated increase of the tax on risk-free capital income decreases savings and thus risk-free returns. Furthermore, we showed in the previous subsection that the tax rate  $t^e$  on excess returns is never negative. The tax on risk-free capital income should thus be larger if the relative importance of aggregate risk decreases as a consequence, and vice versa.

The term involving the covariance between  $\beta$  and  $y^n$  reflects the potential to use the tax on risk-free returns to obtain distributional benefits that cannot be obtained through a tax on labour income. If individuals only differ in their ability to earn labour income, then the conditional covariance of the welfare weights with the risk-free capital incomes is zero:  $\text{cov}_\Omega(\beta, y^n | z) = 0$ . In this case there is no scope for redistribution through the tax on risk-free capital income: all ex ante redistribution takes place through the tax on labour income. If instead  $\text{cov}_\Omega(\beta, y^n | z) \neq 0$ , then the tax on risk-free returns should be used to redistribute from individuals with lower welfare weights to those with higher welfare weights, conditional on labour income.

Given the heterogeneity of the expected returns to private investment, the welfare weights do vary with the amount of risk-free capital income, conditional on labour income, so  $\text{cov}_\Omega(\beta, y^n|z) \neq 0$ . We show in Appendix A.2 that when the utility function is additively separable between consumption in both periods and between consumption and labour supply, and individual preferences exhibit constant absolute risk aversion, then changes in the expected rate of return  $\alpha$  only have income effects, both on savings  $s$  and on labour income  $z$ . Those individuals with a low return to private investment, and thus those with a higher welfare weight conditional on labour income, are then the ones who save the most to smooth their consumption over the life cycle. The covariance  $\text{cov}_\Omega(\beta, y^n|z)$  would thus be positive: conditional on labour income, the individuals with the highest risk-free capital income would also have the highest welfare weights. The presence of the distributional characteristic of risk-free capital income in (30) then puts a downward pressure on the optimal tax on risk-free capital income.

With more general preferences, one cannot unambiguously sign the covariance  $\text{cov}_\Omega(\beta, y^n|z)$  based on theoretical considerations alone. Even if there is growing empirical evidence of persistent heterogeneity in rates of return after correcting for risk (e.g. Fagereng et al. (2020)), it remains unclear to what extent such differences are independent of differences in the ability to earn labour income. More research is needed to find whether the conditional covariance  $\text{cov}_\Omega(\beta, y^n|z)$  differs substantially from zero, and in which direction.

Similarly, even if the term  $\int_{\underline{z}}^{\bar{z}} \mathcal{A}_{t^n}^{c*} dG^z(z)$  easily becomes zero under common theoretical assumptions, its sign is ultimately an empirical question. There is limited empirical evidence about the effects of capital taxes on investment decisions. Akcigit et al. (2018) find that higher income taxes decrease the activities of inventors. Higher income taxes thus discourage activities that are less correlated with the market. Jakobsen et al. (2020) find that wealth taxes have a stronger effect on wealth accumulation at the top. We know that private equity is more concentrated at the top of the income distribution. Taken together, these results suggest that a tax on wealth decreases the relative importance of investments in activities whose returns correlate less with the market. Lemma 2 then indicates that the term  $\int_{\underline{z}}^{\bar{z}} \mathcal{A}_{t^n}^{c*} dG^z(z)$  is positive.

Recall that the integral on the left-hand side of (30) is negative. If both terms on the right-hand side of (30) are indeed positive, then Theorem 5 indicates that the optimal tax on risk-free returns to capital income is negative.

#### 4.4 Optimal nonlinear tax on labour income

The optimality condition for the intercept (24) is standard. The optimal marginal labour income tax (23) resembles the standard Mirrlees result, amended to take into account induced expected capital income tax revenue effects of a labour income tax reform, and the uncertainty involved therein. The standard Mirrlees result can be recovered when all individuals earn equal returns to capital (i.e.  $\text{cov}(\beta, y^n|z) = 0$ ) and policy reforms do not alter the relative shares of aggregate and idiosyncratic risk (i.e.  $\mathcal{A}_\nu = 0$ ). According to Theorem

5, the optimal tax on capital income then equals zero, and the marginal social utility of income can be simplified:

$$\beta(w) \equiv (1 + r^b) \left( \frac{\mathbb{E}[\tilde{U}_2]}{\mathbb{E}[\tilde{\lambda}]} - t_z^\ell z_\rho \right). \quad (31)$$

The optimal tax on labour income then takes the traditional form, as shown in the following theorem.

**Theorem 6.** *When all individuals earn equal returns to capital and compensated policy reforms do not alter the relative shares of aggregate and idiosyncratic risk, then the marginal optimal tax rates on labour income are determined by the classic Mirrlees equation:*

$$\forall z : \frac{t_z^\ell}{1 - t_z^\ell} = \frac{1}{e^z} \cdot \frac{1 - G^z}{z g^z} \cdot \frac{\int_z^{\bar{z}} (1 + r - \bar{\beta}(z)) dG^z}{1 - G^z},$$

with present value compensated elasticity of labour income:

$$\forall z : e^z \equiv -\frac{1}{1 + r^b} \frac{1 - t_z^\ell}{z} z_\sigma^*.$$

## 4.5 Optimal stochastic public good provision

Condition (27) shows that for each market outcome  $r^m$ , the government should choose the level of the public good such that the average value of the expected marginal utility of its provision  $\tilde{U}_P$  equals the shadow price of government revenue  $\lambda^{r^m}$ .

This can be related to a Samuelson-type rule for public goods provision by adopting the approach of Christiansen (1981). Take the expected value of (27), divide by  $\mathbb{E}[\tilde{\lambda}]$  and rearrange to find:

$$1 = \iint_{\Omega} (1 + r^b) \frac{\mathbb{E}[\tilde{U}_P]}{\mathbb{E}[\tilde{U}_1]} dG^\omega(\omega) + \iint_{\Omega} \left( \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{\lambda}]} - (1 + r^b) \right) \frac{\mathbb{E}[\tilde{U}_P]}{\mathbb{E}[\tilde{U}_1]} dG^\omega(\omega). \quad (32)$$

The term  $\mathbb{E}[\tilde{U}_P]/\mathbb{E}[\tilde{U}_1]$  reflects an individual's expected marginal valuation of the public good in terms of first-period consumption. The first term on the right-hand side of (32) reflects the effect on social welfare of a marginal increase in public good provision when all individuals have equal social weights, so  $\mathbb{E}[\tilde{U}_1]/\mathbb{E}[\tilde{\lambda}] = 1 + r^b$  for each individual. This term reflects the traditional Samuelson rule (1954) for public good provision, where the marginal utility of private consumption should be equalized to the marginal utility of the public good. The second term on the right-hand side reflects the distributional effects of a marginal increase in public good provision, due to the fact that different individuals have different social weights. If the expected marginal valuation of the public good  $\mathbb{E}[\tilde{U}_P]/\mathbb{E}[\tilde{U}_1]$  increases sufficiently with the social weight  $\mathbb{E}[\tilde{U}_1]/\mathbb{E}[\tilde{\lambda}]$ , then the second term on right-hand side of (32) is positive. This implies that the first term on the right-hand side of (32) must be smaller than one, and thus that the provision of the public good must be larger than if distributional motives were ignored.

## 4.6 Optimality of the Rate-of-Return Allowance

An RRA is defined as a zero-tax on risk-free capital income combined with a positive tax on excess returns. With an RRA, the entire burden of insurance and redistribution of capital income falls on the tax on excess returns.

We have shown that the tax on excess returns serves the purpose of insurance, by pooling idiosyncratic risk and by balancing the uncertainty of private consumption against the uncertainty of public good provision. However, we demonstrated that with fairly weak assumptions, the tax on excess returns cannot be used for redistribution. If a tax on excess return is introduced, then individuals will reoptimize their portfolios such that their net-of-tax returns to capital and their utility remain unchanged, analogous to the well-known Domar-Musgrave (1944) result.

Whether the tax on risk-free capital income should be zero depends on two factors. The first is whether conditional on labour income, risk-free capital income is correlated with the welfare weights. In our model this is the case when returns to private investment are heterogeneous even after controlling for risk. In this case there is scope for redistribution through the tax on risk-free capital income, beyond what is possible using a tax on labour income alone.

The second factor determining whether the tax on risk-free capital income should be zero is whether a tax on risk-free returns alters the relative importance of aggregate and idiosyncratic risk in the individual investment portfolios. In the presence of a tax on excess returns, tax revenues become more uncertain if the importance of aggregate risk increases, which has adverse consequences for social welfare. A tax on risk-free capital income should be larger if this decreases the importance of aggregate risk, and vice versa. We found that commonly made theoretical assumptions can reduce this second factor to zero.

Summarizing, we find that a tax on risk-free capital income has a different function than a tax on excess capital income. A tax on excess returns serves to insure against potential realizations of the investment risk. Any redistribution based on ex ante characteristics conditional on labour income, however, should occur through the tax on risk-free capital income. Our results thus go against recommendations to introduce an RRA.<sup>18</sup>

## 5 Returning the risk using a stochastic lump sum

Up to now we have assumed that the government returns the risk of its revenues using a stochastic public good. The stochasticity of the government's revenues then did not affect private consumption risk. With the assumption that preferences were separable between public and private consumption, the provision of the public good did not affect individual behaviour.

We now consider the case where the government returns the risk of its revenues using a stochastic lump sum  $\tilde{K}$ . The model remains the same as in Section 2, however the individual budget constraint in the second period

---

<sup>18</sup>This does not necessarily mean that the RRA proposal should be discarded altogether. Suppose that the optimal taxes on risk-free and excess capital incomes equal  $t^{n*}$  and  $t^{e*}$  respectively. If the government would exempt a rate of return  $r^E = (1 - t^{n*}/t^{e*})r^b$  from taxation, and if it would tax any deviations from  $r^E$  at rate  $t^{e*}$ , then the government would still be implementing the optimum.

is altered as follows:

$$\tilde{c}^2 \equiv s + (1 - t^n)y^n + (1 - t^e)\tilde{y}^e + \tilde{K}, \quad (33)$$

and the government's intertemporal budget constraint becomes:

$$K(\tilde{r}^m) = \iint_{\Omega} [(1 + r^b)t^\ell(z) + t^e E_{\mathcal{E}}[\tilde{y}^e | \tilde{r}^m] + t^n y^n] dG^\omega(\omega). \quad (34)$$

The presence of the stochastic lump sum affects individual behaviour. The supply and demand functions now also depend on the function  $K(\cdot)$ , as does the indirect utility function.

We treat the levels of  $K(\tilde{r}^m)$  to be provided for each potential realization of the market shock as policy instruments in the government's optimization problem. The government sets the tax instruments together with the levels of the lump sum  $K(r^m)$  to maximize social welfare:

$$\max_{t^\ell(\cdot), t^n, t^e, K(\cdot)} \iint_{\Omega} E[\tilde{V}(t^\ell(\cdot), t^n, t^e, K(\cdot))] dG^\omega(\omega), \quad (35)$$

subject to intertemporal budget constraint (34).

We derive the first-order conditions for the government optimization problem in Appendix A. We find that optimality conditions (23)–(26) remain valid. In Appendix B we show that if the Mean-Variance framework applies in the presence of a stochastic lump sum that is positively correlated with the market risk, then with some simplifying assumptions the relative importance of the aggregate risk responds in the same direction as the size of the risky portfolio:<sup>19</sup>

$$\iint_{\Omega} \mathcal{A}_\nu dG^\omega(\omega) \gtrless 0 \iff p_\nu \gtrless 0 \text{ and } f_\nu \gtrless 0. \quad (36)$$

The intuition is that in the presence of the stochastic lump sum, individuals are less interested in investing in assets that are also correlated with the market risk. As the risky portfolio grows however, the relative importance of the stochastic lump sum decreases, and individuals will be more interested in investing in the market asset.

Result (36) reduces the probability that the optimal tax on risk-free capital income is negative. To see this, note that when there is a tax increase on risk-free returns, the compensated response of total savings conditional on labour income is unambiguously negative ( $s_{t^n}^{c*} < 0$ ), and the risky portfolio is likely to decline in size ( $p_\nu^{c*} < 0$  and  $f_\nu^{c*} < 0$ ). Result (36) then implies that  $\iint_{\Omega} \mathcal{A}_{t^n} dG^\omega(\omega) < 0$ : the relative importance of aggregate risk in the portfolios declines. From a theoretical point of view, the second term on the right-hand side of Theorem 5 is then more likely to be negative, and the sign of the optimal tax on risk-free capital income is more likely to be positive.

The optimality condition (27) for the public good is now replaced by the following optimality condition for

---

<sup>19</sup>The additional assumptions needed are that either the semi-elasticities of  $p$  and  $f$  with respect to  $\nu$  are equal for all individuals, or that all individuals earn the same expected rates of return  $\alpha$  to private investment. See Appendix B for proofs. If there are no assets with idiosyncratic risk, we again find that  $\iint_{\Omega} \mathcal{A}_\nu dG^\omega(\omega) = 0$ , even in the presence of the stochastic lump sum.

the stochastic lump sum:<sup>20</sup>

$$K(r^m) : \quad \frac{\lambda^{r^m}}{E[\lambda]} = \iint_{\Omega} \frac{E_{\mathcal{E}}[\tilde{U}_2|r^m]}{E[\tilde{U}_2]} dG^{\omega}(\omega) + \iint_{\Omega} \left( t^n \bar{y}_{K(r^m)}^{nc*} + t^e \bar{A}_{K(r^m)}^{c*} \right) dG^{\omega}(\omega) \quad (38)$$

$$+ \frac{E[\tilde{U}_2]}{E[\tilde{U}_1]} \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega} \left( \frac{E_{\mathcal{E}}[\tilde{U}_2|r^m]}{E[\tilde{U}_2]}, \beta \middle| z \right) dG^z(z).$$

The left-hand side of (38) reflects the direct effect on the government budget of an increase of the lump sum in a specific realization  $r^m$  of the state of the world. The first term on the right-hand side contains the direct welfare effects of the perturbation on the individuals. The remaining terms on the first line of (38) reflect compensated behavioural effects of the perturbation on government revenue from the capital incomes, conditional on labour income. A change of  $K(r^m)$  for a specific realization of the market risk  $r^m$  changes the risk properties of second-period consumption. This causes individuals to alter their investments, even if they are compensated for the perturbation. The term on the second line of (38) reflects the distributional consequences of increasing the stochastic lump sum in a specific realization of the state of the world. The terms on the right-hand side of (38) are difficult to characterize without further assumptions on individual behaviour.

Substituting condition (38) into the optimality condition (26) for the tax on excess returns, we obtain the following Theorem.

**Theorem 7.** *If the government returns the risk of its tax revenues using a stochastic lump sum, then the optimal tax on excess returns satisfies the following necessary condition:*

$$\iint_{\Omega} \frac{\text{cov}(\tilde{U}_2, \tilde{\varepsilon}^e p)}{E[\tilde{U}_2]} dG^{\omega}(\omega) = t^n \text{cov} \left( \iint_{\Omega} \bar{y}_{K}^{nc*} dG^{\omega}(\omega), \iint_{\Omega} \tilde{y}^e dG^{\omega}(\omega) \right) + t^e \text{cov} \left( \iint_{\Omega} \bar{A}_{K}^{c*} dG^{\omega}(\omega), \iint_{\Omega} \tilde{y}^e dG^{\omega}(\omega) \right). \quad (39)$$

*Proof.* See Appendix A. □

The left-hand side of (39) indicates how much of the idiosyncratic risk remains after the tax on excess returns is imposed. If the distortions on the right-hand side of (39) were to approach zero, then the optimal tax on excess returns would approach 100%, such that the idiosyncratic risk would be fully insured. If moreover there were no idiosyncratic risk, but only aggregate risk, then the left-hand side of (39) would always be zero, and the value of the tax on excess returns would become irrelevant. In this case it would optimally be set to zero, to minimize the compliance costs.

The irrelevance of the tax on excess returns in absence of idiosyncratic risk and compensated behavioural responses extends the findings of Atkinson and Stiglitz (1980), who assume the presence of a representative

<sup>20</sup>To arrive at (38), we extend assumption (22) to include the following statement for any  $k = z, y^n, \tilde{y}^e$ :

$$\text{cov} \left( k_{\sigma}^*, \frac{d}{dz} \left( \frac{E_{\mathcal{E}}[\tilde{U}_2|r^m]}{E[\tilde{U}_2]} \right) \middle| z \right) = 0, \quad (37)$$

where  $E_{\mathcal{E}}[\tilde{U}_2|r^m]/E[\tilde{U}_2]$  is the certainty equivalent of an increase in the stochastic lump sum for a specific realization  $r^m$  of the market risk. We thus assume, again following Saez (2002), that among individuals who earn the same labour income, the compensated behavioural responses to a labour tax reform are not systematically correlated with the propensity of this certainty equivalent to increase with labour income. Formulations of the optimum that do not use Assumption (37) can be found in Appendix A.8.

consumer for whom the tax on excess returns is exactly undone by the stochastic lump sum, and of Gordon (1985), who assumes that the lump sums are tailored to the individuals. In our case, the fact that the lump sums are not matched to the individuals does not undo the irrelevance of the tax on excess returns. Any redistributive effects of introducing a positive tax on excess returns are absorbed as the other policy instruments are re-optimized, and the final effect on social welfare is zero.

In presence of both aggregate and idiosyncratic risk, Theorem 7 shows that the optimal insurance is no longer mitigated by a balancing act between public risk and private risk. This extends the results of Gordon (1985), who finds in a comparative statics exercise that a tax on excess returns has no behavioural effects when the lump sums are tailored to the individuals in the non-corporate sector. He finds that it would be better, from a welfare perspective, to fully insure the idiosyncratic risk. Condition (39) states that in presence of a nonlinear tax on earnings and in absence of compensated effects of the stochastic lump sum, this full insurance result remains valid even when the lump sums are not tailored to the individuals.

The right-hand side of (39) indicates the distortions caused by not tailoring the stochastic lump sums to individual risks. Even if individuals are compensated for changes in their expected utility, an increase of the stochastic lump sum in high realizations of the market rate will generally have different behavioural effects than an increase of the stochastic lump sum in low realizations of the market rate. If these distortions indeed depend on the state of the world, then the optimal tax on excess capital income differs from 100%. These distortions however are difficult to characterize without further assumptions on individual behaviour. This question is beyond the scope of our paper.

In conclusion, our main results remain valid when the government returns the aggregate risk of its revenues using a stochastic lump sum. The optimal tax on excess capital income is still unambiguously positive for insurance reasons, though the government now only insures idiosyncratic risk, taking into account the distortions caused by the way the aggregate risk is returned to the individuals. The tax on risk-free capital income still serves to redistribute based on differences between individuals conditional on labour income, and to reduce the risk of tax revenues from excess capital income. The sign of the optimal on risk-free capital income is ambiguous.

## 6 Conclusion

We have studied optimal linear taxation of risk-free and excess returns to capital, alongside an optimal nonlinear tax on labour income, in an intertemporal model with risky assets. Our approach differs from the standard Mirrlees optimal income tax analysis in a number of dimensions. We incorporate portfolio choice into a two-period model with heterogeneous households who both supply labour and save in the first period. Portfolios include both safe and risky assets where the latter combine market assets with aggregate risk and private investment with idiosyncratic risk. Individuals differ both in their labour productivity and in the expected return on their private investments. Aggregate risk on market assets is reflected in the government's revenue stream, which in turn leads to uncertainty in government spending either on public goods or on identical



lump-sum transfers to all individuals.

In our base case, we assume the government returns uncertain tax revenues to the economy by varying the supply of public goods. A tax on excess returns taxes both rents of private investments and returns to risk. One might expect that since individuals differ in the productivity of their private investments, taxing excess returns might serve a redistribution role. However, a Domar-Musgrave effect nullifies such a role: individuals respond to a tax on excess returns by adjusting their portfolios such that their expected utility remains unchanged. The tax does fulfill an insurance role. Such a tax reduces the riskiness of private consumption while increasing that of public goods consumption. At an optimum, a positive tax on excess returns trades off these two risks. At the same time, a tax on risk-free returns can have redistributive consequences and can serve as a complement to the progressive earnings tax. Individuals of different investment productivities but the same earnings will generally differ in their savings and risk-free capital income. However, the relation between investment productivity and savings is ambiguous: under reasonable assumptions more productive persons might save less. If so, the optimal tax on risk-free capital income would be negative. This implies that, while it is optimal to tax risk-free capital income and excess returns to capital at different rates, the RRA system proposed by the Mirrlees Review (2011) will generally not be optimal. These results extend to the case where the government returns its uncertain revenues to households as lump sums.

Marginal labour income tax rates will take into account capital income tax revenue changes induced by changes in the marginal income tax rate. Public goods decision rules are standard and are equivalent to equity-adjusted Samuelson rules. For each realization of aggregate risk, the sum of the expected marginal social utilities of public good should equal the shadow value of government revenue.

Our stylized model shows that optimal tax results are sensitive to the inclusion of risky assets. As more refined models are developed to study the optimal taxation of capital in presence of heterogeneous rates of return, one should account for the risk incurred in obtaining those returns. The fact that a tax on excess returns does not discourage risk-taking due to the Domar-Musgrave (1944) effect, does not imply that one can simply extrapolate the results from a model with deterministic capital incomes to a context with risky capital incomes. One should be particularly careful when applying a sufficient-statistics framework. The change in the tax liability due to a tax reform on risky excess returns, is not an adequate measure for its effect on individual welfare. Moreover, elasticities of expected tax revenues are no longer sufficient if uncertain tax revenues decrease social welfare.

## References

- Stuart Adam and Helen Miller. Taxing work and investment across legal forms: pathways to well-designed taxes. *The Institute for Fiscal Studies*, 2021.
- Ufuk Akcigit, John Grigsby, Tom Nicholas, and Stefanie Stantcheva. Taxation and innovation in the 20th century. *National Bureau of Economic Research Working Paper Series w24982*, 2018.

- George B Arfken and Hans J Weber. *Mathematical methods for physicists international student edition*. Academic press, 2005.
- Anthony B. Atkinson and Joseph E. Stiglitz. The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, 6(1-2):55–75, 1976.
- Anthony B. Atkinson and Joseph E. Stiglitz. *Lectures on public economics*. McGraw-Hill College, 1980.
- Alan J Auerbach, Michael P Devereux, Michael Keen, and John Vella. Destination-based cash flow taxation. *Oxford Legal Studies Research Paper WP 17/01*, 2017.
- James Banks and Peter A Diamond. The base for direct taxation. *Institute for Fiscal Studies, ed., Dimensions of Tax Design, New York: Oxford University Press*, 2010.
- Spencer Bastani and Daniel Waldenström. How should capital be taxed? *Journal of Economic Surveys*, 34(4): 812–846, 2020.
- Jess Benhabib, Alberto Bisin, and Mi Luo. Wealth distribution and social mobility in the us: A quantitative approach. *American Economic Review*, 109(5):1623–47, 2019.
- Radim Boháček and Michal Kejak. Optimal government policies in models with heterogeneous agents. *Journal of Economic Theory*, 176:834–858, 2018.
- Wolfgang Buchholz and Kai A Konrad. Taxes on risky returns – an update. *Max Planck Institute for Tax Law and Public Finance Working Papers 2014-10*, 2014.
- Vidar Christiansen. Evaluation of public projects under optimal taxation. *The Review of Economic Studies*, 48(3):447–457, 1981.
- Vidar Christiansen. Which commodity taxes should supplement the income tax? *Journal of Public Economics*, 24(2):195–220, 1984.
- Vidar Christiansen. A normative analysis of capital income taxes in the presence of aggregate risk. *The Geneva Papers on Risk and Insurance-Theory*, 18(1):55–76, 1993.
- Wilfred J Corlett and Douglas C Hague. Complementarity and the excess burden of taxation. *The Review of Economic Studies*, 21(1):21–30, 1953.
- Michael P Devereux, Alan J Auerbach, Michael Keen, Paul Oosterhuis, Wolfgang Schön, and John Vella. Residual profit allocation by income. *Oxford University Centre for Business Taxation Working Paper WP19/01*, 2019.
- Peter A Diamond. A many-person ramsey tax rule. *Journal of Public Economics*, 4(4):335–342, 1975.

- Evsey D Domar and Richard A Musgrave. Proportional income taxation and risk-taking. *The Quarterly Journal of Economics*, pages 388–422, 1944.
- Andreas Fagereng, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. Heterogeneity and persistence in returns to wealth. *Econometrica*, 88(1):115–170, 2020.
- Xavier Gabaix, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. The dynamics of inequality. *Econometrica*, 84(6):2071–2111, 2016.
- Firouz Gahvari and Luca Micheletto. Capital income taxation and the atkinson–stiglitz theorem. *Economics Letters*, 147:86–89, 2016.
- Malcolm Gammie. *Equity for Companies: a Corporation Tax for the 1990s*. Institute for Fiscal Studies, 1991.
- Aart Gerritsen, Bas Jacobs, Alexandra V. Rusu, and Kevin Spiritus. Optimal taxation of capital income with heterogeneous rates of return. *CEPrifo Working Paper No. 8395*, 2020.
- Roger H Gordon. Taxation of corporate capital income: tax revenues versus tax distortions. *The Quarterly Journal of Economics*, 100(1):1–27, 1985.
- Fatih Guvenen, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen. Use it or lose it: Efficiency gains from wealth taxation. Technical report, National Bureau of Economic Research Working Paper Series w26284, Cambridge, MA., 2019.
- Bas Jacobs and Robin Boadway. Optimal linear commodity taxation under optimal non-linear income taxation. *Journal of Public Economics*, 117:201–210, 2014.
- Laurence Jacquet, Etienne Lehmann, and Bruno Van der Linden. Optimal redistributive taxation with both intensive and extensive responses. *Journal of Economic Theory*, 148(5):1770–1805, 2013.
- Katrine Jakobsen, Kristian Jakobsen, Henrik Kleven, and Gabriel Zucman. Wealth taxation and wealth accumulation: Theory and evidence from denmark. *The Quarterly Journal of Economics*, 135(1):329–388, 2020.
- Marcin Kacperczyk, Jaromir Nosal, and Luminita Stevens. Investor sophistication and capital income inequality. *Journal of Monetary Economics*, 2018.
- Louis Kaplow. Taxation and general risk taking: A general equilibrium per. *National Tax Journal*, 47(4):789, 1994.
- Henrik Kleven. Sufficient statistics revisited. Technical report, National Bureau of Economic Research Working Paper Series w27242, 2020.
- Arnaldur S. Kristjánsson. Optimal taxation with endogenous return to capital. Mimeo, Oslo University, 2016.

- Etienne Lehmann, Sander Renes, Kevin Spiritus, and Floris T. Zoutman. Optimal taxation with multiple incomes and types. mimeo, 2019.
- Annamaria Lusardi, Pierre-Carl Michaud, and Olivia S. Mitchell. Optimal financial knowledge and wealth inequality. *Journal of Political Economy*, 125(2):431–477, 2017.
- Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.
- James Mirrlees, Stuart Adam, Tim Besley, Richard Blundell, Steve Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba. *Tax By Design: The Mirrlees Review*. Oxford University Press, Sep 2011.
- Jan Mossin. Taxation and risk-taking: an expected utility approach. *Economica*, 35(137):74–82, 1968.
- Thomas Piketty. *Capital in the twenty-first century*. Belknap Press of Harvard University Press Cambridge, MA, 2014.
- Emmanuel Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68(1): 205–29, January 2001.
- Emmanuel Saez. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics*, 83(2):217–230, 2002.
- Emmanuel Saez and Gabriel Zucman. Wealth inequality in the united states since 1913: Evidence from capitalized income tax data. *Quarterly Journal of Economics*, 131(2):519–578, 2016.
- Paul A Samuelson. The pure theory of public expenditure. *The Review of Economics and Statistics*, 36(4): 387–389, 1954.
- Agnar Sandmo. Portfolio theory, asset demand and taxation: Comparative statics with many assets. *The review of economic studies*, 44(2):369–379, 1977.
- Florian Scheuer and Joel Slemrod. Taxing our wealth. *Journal of Economic Perspectives*, 35(1):207–30, 2021.
- Dirk Schindler. Taxing risky capital income—a commodity taxation approach. *FinanzArchiv: Public Finance Analysis*, 64(3):311–333, 2008.
- Joseph E Stiglitz. The effects of income, wealth, and capital gains taxation on risk-taking. *The Quarterly Journal of Economics*, pages 263–283, 1969.
- Nicolas Werquin, Aleh Tsyvinski, and Mikhail Golosov. A variational approach to the analysis of tax systems. Technical report, Society for Economic Dynamics, 2015.

## Appendix A First-order conditions for the government

We prove our results for the general case that allows for both a stochastic public good and a stochastic lump sum (described in subsection 5) to return the uncertainty of the government’s revenues. The government uses two linear instruments ( $t^n$  and  $t^e$ ) and three nonlinear instruments ( $t^\ell(z)$ ,  $K(\tilde{r}^m)$  and  $P(\tilde{r}^m)$ ) to maximize social welfare subject to a budget constraint. For the optimization of the linear instruments we derive standard first-order conditions. The optimization of the nonlinear instruments is more challenging.

A standard approach to optimize a nonlinear function is the Euler-Lagrange formalism. A standard assumption of the Euler-Lagrange method is that the arguments of the functions that are being optimized, are exogenous. This assumption is clearly violated when optimizing a tax function: individuals change their behaviour when the tax system changes. The tax base is endogenous. Saez (2001) develops a heuristic perturbation method to overcome this difficulty in a deterministic setting where the government only levies a nonlinear tax on labour income. For a given perturbation of the tax function, he lists the effects on the government objective and requires that the total effect is zero. Recent contributions by Werquin et al. (2015), Boháček and Kejak (2018) and Lehmann et al. (2019) further formalize the perturbation method, enabling its application to more complex policy questions.

We adapt a standard proof for the Euler-Lagrange equation (see e.g. Arfken and Weber (2005) chapter 17), to incorporate behavioural responses to perturbations of the policy instruments. We start in Subsection A.1 by formally defining our perturbations to the policy instruments. In Subsection A.2 we study the effects of our perturbations on individual behaviour. This allows setting up a Lagrange equation for our problem in Subsection A.3, and deriving the first-order conditions for the government in the remaining subsections.

### A.1 Perturbations to the policy instruments

If a given policy instrument is optimal, then any perturbation to it leaves the government objective unchanged. To find optimal values for the linear instruments, we perturb  $t^n$  and  $t^e$  and demand that the effects on the government objective sum to zero. For the nonlinear instruments, we introduce perturbation functions. For the tax on labour income, for example, we introduce a perturbation function  $\epsilon^z \eta^z(z)$ . The function  $\eta^z(z)$  models an arbitrary, nonlinear but sufficiently smooth perturbation. The parameter  $\epsilon^z$  is an infinitesimal that allows varying the size of the perturbation. After the perturbation, the tax liability for any labour income  $z$  equals  $t^\ell(z) + \epsilon^z \eta^z(z)$ . Together, the entities  $\epsilon^z$  and  $\eta^z$  allow modelling any small perturbation to the tax function  $t^\ell(z)$ . If the value of  $\epsilon^z$  equals zero, then the unperturbed tax function is in place. If the optimal value of  $\epsilon^z$  is zero for every function  $\eta^z$ , then we know that the unperturbed tax function is optimal. Similarly, we introduce perturbations  $\epsilon^K \eta^K(\tilde{r}^m)$  and  $\epsilon^P \eta^P(\tilde{r}^m)$  respectively for the state-dependent lump sum and the state-dependent public good.

The general perturbations  $\epsilon^z \eta^z(z)$ ,  $\epsilon^K \eta^K(\tilde{r}^m)$  and  $\epsilon^P \eta^P(\tilde{r}^m)$  look different from the perturbations that we introduce in the main text. To verify the optimality of the government policies, we must verify that *any*

perturbation leaves the government objective unchanged. However, the perturbation parameters introduced in the main text suffice to describe the comparative statics of such general perturbations. For example, we will show in the next subsection that the effects on individual behaviour of any perturbation of the labour income tax schedule can be decomposed into income and substitution effects, and can thus be described using the simpler perturbation parameters  $\rho$  and  $\sigma$ .<sup>21</sup> Furthermore, we will show that the general perturbations to the lump sum  $\epsilon^K \eta^K(\tilde{r}^m)$  and to the public good  $\epsilon^P \eta^P(\tilde{r}^m)$  can be decomposed into local perturbations for specific realizations of the state of the world  $r^m$ . In the main text we model such local perturbations as direct changes to the values  $K(r^m)$  and  $P(r^m)$ . We now model these local perturbations more formally using a Dirac delta function  $\delta(\tilde{r}^m - r^m)$ , and we let  $\kappa^m$  and  $\pi^m$  parameterize the size of the reforms.<sup>22</sup> The stochastic lump sum after the perturbations then equals  $K(\tilde{r}^m) + \epsilon^K \eta^K(\tilde{r}^m) + \kappa^m \delta(\tilde{r}^m - r^m)$ . Similarly, the provision of the stochastic public good after the perturbations equals  $P(\tilde{r}^m) + \epsilon^P \eta^P(\tilde{r}^m) + \pi^m \delta(\tilde{r}^m - r^m)$ .

The budget constraints taking into account all perturbation parameters are then:

$$\mathcal{C}^1(z, s, \epsilon^z, \sigma, \rho) \equiv z - s - t^\ell(z) - \sigma z - \epsilon^z \eta^z(z) + \rho, \quad (40)$$

$$\begin{aligned} \tilde{\mathcal{C}}^2(s, f, p, t^n, t^e, \epsilon^K, \kappa^m, \rho_2) \equiv & s + (1 - t^n)r^b s + (1 - t^e)\tilde{y}^e(f, p) + \rho_2 \\ & + K(\tilde{r}^m) + \kappa^m \delta(\tilde{r}^m - r^m) + \epsilon^K \eta^K(\tilde{r}^m), \end{aligned} \quad (41)$$

where  $\tilde{y}^e(f, p)$  is defined by (3), and  $\rho_2$  denotes a deterministic lump sum that is given to all individuals in the second period.

## A.2 Individual behaviour

In this subsection we derive properties of individual supply and demand, which we will use in the following subsections to characterize the optimal policies. It is mathematically convenient to redefine utility function (5) by substituting budget constraints (40) and (41) for  $c^1$  and  $c^2$ . An individual facing the perturbed government policies thus maximizes the expected value of the following utility function:

$$\begin{aligned} \tilde{U}(z, s, f, p, \sigma, \epsilon^z, \rho, t^n, t^e, \epsilon^K, \kappa^m, \epsilon^P, \pi^m, \rho_2) \\ \equiv U\left(u(\mathcal{C}^1(z, s, \epsilon^z, \sigma, \rho), \tilde{\mathcal{C}}^2(s, f, p, t^n, t^e, \epsilon^K, \kappa^m, \rho_2)), \frac{z}{w}, P(\tilde{r}^m, \epsilon^P, \pi^m, r^m)\right), \end{aligned} \quad (42)$$

where we write  $P(\tilde{r}^m, \epsilon^P, \pi^m, r^m)$  as shorthand for  $P(\tilde{r}^m) + \epsilon^P \eta^P(\tilde{r}^m) + \pi^m \delta(\tilde{r}^m - r^m)$ .

<sup>21</sup>Saez (2001) considers one specific perturbation, where the function  $\eta^z$  alters the slope of the tax schedule in a small interval around a specific income level, it leaves the tax schedule unaltered below that income level, and it changes the level but leaves the slope of the tax schedule unaltered for individuals with higher incomes. Given the complexity of our model, we consider more general perturbations, and formally derive their effects on the government objective.

<sup>22</sup>The Dirac delta function  $\delta(r^m - R^m)$  is defined such that  $\delta(r^m - R^m) = 0$  whenever  $r^m \neq R^m$ , and it integrates to one on the set of real numbers:  $\int_{-\infty}^{\infty} \delta(r^m - R^m) dr^m = 1$ . This setup allows modelling a localized reform to the lump sum with non-zero effects on social welfare.

The first-order conditions for the individual are:

$$\mathbb{E}[\tilde{\mathcal{U}}_z] = \mathbb{E}[\tilde{\mathcal{U}}_s] = \mathbb{E}[\tilde{\mathcal{U}}_f] = \mathbb{E}[\tilde{\mathcal{U}}_p] = 0. \quad (43)$$

Denote the supply function for labour income as  $z(\cdot)$  and denote the demand functions for assets as  $s(\cdot)$ ,  $f(\cdot)$  and  $p(\cdot)$ . Each of the supply and demand functions has as arguments type  $\omega$  and the perturbation parameters  $\sigma$ ,  $\epsilon^z$ ,  $\rho$ ,  $t^n$ ,  $t^e$ ,  $\epsilon^K$ ,  $\kappa^m$ ,  $\epsilon^P$ ,  $\pi^m$  and  $\rho_2$ .

Substituting the supply and demand functions into the utility function yields indirect utility  $\tilde{\mathcal{V}}(\cdot)$ , which also is a function of type and of the perturbation parameters. Applying the envelope theorem to problem (42) yields the following properties:

$$\mathbb{E}[\tilde{\mathcal{V}}_{\epsilon^z}] = -\mathbb{E}[\tilde{U}_1]\eta^z(z), \quad \mathbb{E}[\tilde{\mathcal{V}}_\sigma] = -\mathbb{E}[\tilde{U}_1]z, \quad (44)$$

$$\mathbb{E}[\tilde{\mathcal{V}}_{t^e}] = 0, \quad \mathbb{E}[\tilde{\mathcal{V}}_{t^n}] = -\mathbb{E}[\tilde{U}_2]y^n, \quad \mathbb{E}[\tilde{\mathcal{V}}_\rho] = \mathbb{E}[\tilde{U}_1], \quad \mathbb{E}[\tilde{\mathcal{V}}_{\rho_2}] = \mathbb{E}[\tilde{U}_2], \quad (45)$$

$$\mathbb{E}[\tilde{\mathcal{V}}_{\epsilon^K}] = \mathbb{E}[\tilde{U}_2\eta^K(\tilde{r}^m)], \quad \mathbb{E}[\tilde{\mathcal{V}}_{\kappa^m}] = \mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m], \quad \mathbb{E}[\tilde{\mathcal{V}}_{\epsilon^P}] = \mathbb{E}[\tilde{U}_P\eta^P(\tilde{r}^m)]. \quad (46)$$

Combining these envelope conditions allows deriving Slutsky properties, as we do in the following Lemma.

**Lemma 8** (Slutsky properties). *We find the following Slutsky properties for individual behaviour:*

1. for any  $k = z, s, f, p$ , the behavioural effects of perturbations  $\sigma$ ,  $t^n$  and  $\kappa^m$  can be decomposed into income effects and compensated effects:

$$k_\sigma^* = k_\sigma + zk_\rho, \quad k_{t^n}^* = k_{t^n} + y^n k_{\rho_2}, \quad k_{\kappa^m}^* = k_{\kappa^m} - \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_2]}k_{\rho_2}. \quad (47)$$

2. individual behaviour complies to the following Slutsky symmetry between labour income and risk-free capital income:

$$z_{t^n}^* = \frac{\mathbb{E}[\tilde{U}_2]}{\mathbb{E}[\tilde{U}_1]}y_\sigma^{n*}. \quad (48)$$

3. individual behaviour complies to the following Slutsky symmetry between labour income and the stochastic lump sum:

$$z_{\kappa^m}^* = -\frac{\partial}{\partial \sigma} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right)^*, \quad (49)$$

4. individual behaviour complies to the following Slutsky symmetry between risk-free capital income and the stochastic lump sum:

$$y_{\kappa^m}^{n*} = -\frac{\partial}{\partial t^n} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_2]} \right)^*. \quad (50)$$

The latter Slutsky symmetry remains valid conditional on labour income  $z$ .

*Proof.* Denote the value of any demand or supply function  $k$  in the situation before any reforms as  $k^i$ . Introduce

the indirect utility in terms of the different compensated reforms as:

$$\mathcal{V}^*(\omega, \sigma, t^n, \kappa^n) \equiv \mathcal{V}\left(\omega, \sigma, t^n, \kappa^n, \rho = z^i \sigma, \rho_2 = y^{ni} t^n - \frac{E_{\mathcal{E}}[\tilde{U}_2^i | r^m]}{E[\tilde{U}_2^i]} \kappa^m\right),$$

where we omit irrelevant function arguments. Use envelope conditions (44)–(46) to find:

$$E[\tilde{\mathcal{V}}_{\sigma}^*] = E[\tilde{\mathcal{V}}_{\sigma}] + E[\tilde{\mathcal{V}}_{\rho}] z^i = -E[\tilde{U}_1](z - z^i), \quad (51)$$

$$E[\tilde{\mathcal{V}}_{t^n}^*] = E[\tilde{\mathcal{V}}_{t^n}] + E[\tilde{\mathcal{V}}_{\rho_2}] y^{ni} = -E[\tilde{U}_2](y^n - y^{ni}), \quad (52)$$

$$E[\tilde{\mathcal{V}}_{\kappa^m}^*] = E[\tilde{\mathcal{V}}_{\kappa^m}] - E[\tilde{\mathcal{V}}_{\rho_2}] \frac{E_{\mathcal{E}}[\tilde{U}_2 | r^m]}{E[\tilde{U}_2]} = E[\tilde{U}_2] \left( \frac{E_{\mathcal{E}}[\tilde{U}_2 | r^m]}{E[\tilde{U}_2]} - \frac{E_{\mathcal{E}}[\tilde{U}_2^i | r^m]}{E[\tilde{U}_2^i]} \right). \quad (53)$$

1. In the situation without any reforms, (51)–(53) show that  $E[\tilde{\mathcal{V}}_{\sigma}^*] = E[\tilde{\mathcal{V}}_{t^n}^*] = E[\tilde{\mathcal{V}}_{\kappa^m}^*] = 0$ . It follows that the combined reforms in (47) are indeed compensated.
2. Evaluate the partial derivative of (51) with respect to  $t^n$ , and that of (52) with respect to  $\sigma$ , each time in the situation before any reforms:

$$E[\tilde{\mathcal{V}}_{\sigma t^n}^*] = -E[\tilde{U}_1] z_{t^n}^*, \quad (54)$$

$$E[\tilde{\mathcal{V}}_{t^n \sigma}^*] = -E[\tilde{U}_2] y_{\sigma}^{n*}. \quad (55)$$

*Young's theorem* demands that the second-order partial derivatives of any function are symmetric. Apply this requirement to (54) and (55) to find Slutsky symmetry (48).

3. Evaluate the partial derivative of (51) with respect to  $\kappa^m$ , and of (53) with respect to  $\sigma$ , each time in the situation before any reforms:

$$E[\tilde{\mathcal{V}}_{\sigma \kappa^m}^*] = -E[\tilde{U}_1] z_{\kappa^m}^*, \quad (56)$$

$$E[\tilde{\mathcal{V}}_{\kappa^m \sigma}^*] = E[\tilde{U}_2] \frac{\partial}{\partial \sigma} \left( \frac{E_{\mathcal{E}}[\tilde{U}_2 | r^m]}{E[\tilde{U}_2]} \right)^*. \quad (57)$$

Applying *Young's theorem* and using the fact that the marginal rate of substitution  $E[\tilde{U}_2]/E[\tilde{U}_1]$  does not depend on the tax on labour income, yields (49).

4. Evaluate the partial derivative of (52) with respect to  $\kappa^m$ , and of (53) with respect to  $t^n$ , each time in the situation before any reforms:

$$E[\tilde{\mathcal{V}}_{t^n \kappa^m}^*] = -E[\tilde{U}_2] y_{\kappa^m}^{n*}, \quad (58)$$

$$E[\tilde{\mathcal{V}}_{\kappa^m t^n}^*] = E[\tilde{U}_2] \frac{\partial}{\partial t^n} \left( \frac{E_{\mathcal{E}}[\tilde{U}_2 | r^m]}{E[\tilde{U}_2]} \right)^*. \quad (59)$$



Applying *Young's theorem* and using the fact that the marginal rate of substitution  $E[\tilde{U}_2]/E[\tilde{U}_1]$  does not depend on the tax on labour income, yields (49).

□

If there is a marginal change to any of the type or perturbation parameters  $\nu$ , then individuals change their behaviour, such that first-order conditions (43) remain satisfied. We thus obtain for any  $k = z, s, f, p$ :<sup>23</sup>

$$0 = \frac{dE[\tilde{\mathcal{U}}_k]}{d\nu} = E[\tilde{\mathcal{U}}_{kz}]z_\nu + E[\tilde{\mathcal{U}}_{ks}]s_\nu + E[\tilde{\mathcal{U}}_{kf}]f_\nu + E[\tilde{\mathcal{U}}_{kp}]p_\nu + E[\tilde{\mathcal{U}}_{k\nu}]. \quad (60)$$

Write equation (60) in matrix notation to find the following lemma.

**Lemma 9.** *The effects of a change of any parameter  $\nu = w, \sigma, \epsilon^z, \rho, t^n, t^e, \epsilon^K, \kappa^m, \epsilon^P, \pi^m$  or  $\rho_2$  on the choices of the individuals are given by:*

$$\begin{pmatrix} z_\nu \\ s_\nu \\ f_\nu \\ p_\nu \end{pmatrix} = -(\mathbb{E}[\tilde{\mathcal{H}}])^{-1} \cdot \mathbb{E} \left[ \begin{pmatrix} \tilde{\mathcal{U}}_{z\nu} \\ \tilde{\mathcal{U}}_{s\nu} \\ \tilde{\mathcal{U}}_{f\nu} \\ \tilde{\mathcal{U}}_{p\nu} \end{pmatrix} \right],$$

where  $\tilde{\mathcal{H}}$  denotes the Hessian of utility function (42).

Lemma 9 reduces the task of finding relations between the effects of different perturbations, to finding relations between the partial derivatives of the functions  $\tilde{\mathcal{U}}_z, \tilde{\mathcal{U}}_s, \tilde{\mathcal{U}}_f$  and  $\tilde{\mathcal{U}}_p$ . The following Lemma uses Lemma 9 to find relations between the income effects in both periods.

**Lemma 10.** *The behavioural effects of perturbations to the tax intercepts in the two periods are related as follows:*

$$z_\rho = \frac{E[\tilde{U}_1]}{E[\tilde{U}_2]} z_{\rho_2}, \quad s_\rho = \frac{E[\tilde{U}_1]}{E[\tilde{U}_2]} s_{\rho_2} + 1, \quad f_\rho = \frac{E[\tilde{U}_1]}{E[\tilde{U}_2]} f_{\rho_2}, \quad p_\rho = \frac{E[\tilde{U}_1]}{E[\tilde{U}_2]} p_{\rho_2}. \quad (61)$$

*Proof.* Note the following second-order derivatives, evaluated in the situation before any reforms:

$$E[\tilde{\mathcal{U}}_{z\rho}] = (1 - \tau_z^\ell)E[\tilde{U}_{11}], \quad E[\tilde{\mathcal{U}}_{z\rho_2}] = -(1 - \tau_z^\ell)E[\tilde{U}_{12}], \quad (62)$$

$$E[\tilde{\mathcal{U}}_{s\rho}] = -E[\tilde{U}_{11}] + E[\tilde{U}_{12}] \frac{E[\tilde{U}_1]}{E[\tilde{U}_2]}, \quad E[\tilde{\mathcal{U}}_{s\rho_2}] = -E[\tilde{U}_{12}] + E[\tilde{U}_{22}] \frac{E[\tilde{U}_1]}{E[\tilde{U}_2]},$$

$$\forall k = f, p : E[\tilde{\mathcal{U}}_{k\rho}] = (1 - t^e)E[\tilde{U}_{21}\tilde{y}_k^e], \quad E[\tilde{\mathcal{U}}_{k\rho_2}] = (1 - t^e)E[\tilde{U}_{22}\tilde{y}_k^e], \quad (63)$$

<sup>23</sup>The behavioural effects  $z_\nu, s_\nu, f_\nu$  and  $p_\nu$  are total effects. This means that they include not only the direct effects of a change in the parameter  $\nu$ . They also include second-round effects due to the nonlinearity of the tax on labour income. If a change in the parameter  $\nu$  causes a change in labour income  $z$ , this will cause a change in the marginal tax on labour income  $\tau_z^\ell(z)$ , which will cause further effects on the choices of the individuals. Jacquet et al. (2013) and Jacobs and Boadway (2014) include similar second-round effects in their elasticities. The functions  $E[\tilde{\mathcal{U}}_z], E[\tilde{\mathcal{U}}_s], E[\tilde{\mathcal{U}}_f]$  and  $E[\tilde{\mathcal{U}}_p]$  are equivalent to the *shift functions* introduced by Jacquet et al. (2013).

and thus:

$$\forall k = z, s, f, p : \mathbb{E}[\tilde{u}_{ks}] = \mathbb{E}[\tilde{u}_{k\rho_2}] \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} - \mathbb{E}[\tilde{u}_{k\rho}]. \quad (64)$$

Write (64) in matrix notation:

$$\frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} \mathbb{E} \begin{bmatrix} \tilde{u}_{z\rho_2} \\ \tilde{u}_{s\rho_2} \\ \tilde{u}_{f\rho_2} \\ \tilde{u}_{p\rho_2} \end{bmatrix} - \mathbb{E} \begin{bmatrix} \tilde{u}_{z\rho} \\ \tilde{u}_{s\rho} \\ \tilde{u}_{f\rho} \\ \tilde{u}_{p\rho} \end{bmatrix} = \mathbb{E} \begin{bmatrix} \tilde{u}_{zs} \\ \tilde{u}_{ss} \\ \tilde{u}_{fs} \\ \tilde{u}_{ps} \end{bmatrix}.$$

Substitute Lemma 9 on the left-hand side and fully write the Hessian  $\mathbb{E}[\tilde{\mathcal{H}}]$ :

$$-\mathbb{E} \begin{bmatrix} \tilde{u}_{zz} & \tilde{u}_{zs} & \tilde{u}_{zf} & \tilde{u}_{zp} \\ \tilde{u}_{sz} & \tilde{u}_{ss} & \tilde{u}_{sf} & \tilde{u}_{sp} \\ \tilde{u}_{fz} & \tilde{u}_{fs} & \tilde{u}_{ff} & \tilde{u}_{fp} \\ \tilde{u}_{pz} & \tilde{u}_{ps} & \tilde{u}_{pf} & \tilde{u}_{pp} \end{bmatrix} \cdot \mathbb{E} \begin{bmatrix} \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} \begin{pmatrix} z_{\rho_2} \\ s_{\rho_2} \\ f_{\rho_2} \\ p_{\rho_2} \end{pmatrix} - \begin{pmatrix} z_{\rho} \\ s_{\rho} \\ f_{\rho} \\ p_{\rho} \end{pmatrix} \end{bmatrix} = \mathbb{E} \begin{bmatrix} \tilde{u}_{zs} \\ \tilde{u}_{ss} \\ \tilde{u}_{fs} \\ \tilde{u}_{ps} \end{bmatrix},$$

or equivalently:

$$-\mathbb{E} \begin{bmatrix} \tilde{u}_{zz} & \tilde{u}_{zs} & \tilde{u}_{zf} & \tilde{u}_{zp} \\ \tilde{u}_{sz} & \tilde{u}_{ss} & \tilde{u}_{sf} & \tilde{u}_{sp} \\ \tilde{u}_{fz} & \tilde{u}_{fs} & \tilde{u}_{ff} & \tilde{u}_{fp} \\ \tilde{u}_{pz} & \tilde{u}_{ps} & \tilde{u}_{pf} & \tilde{u}_{pp} \end{bmatrix} \cdot \mathbb{E} \begin{bmatrix} \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} \begin{pmatrix} z_{\rho_2} \\ s_{\rho_2} \\ f_{\rho_2} \\ p_{\rho_2} \end{pmatrix} - \begin{pmatrix} z_{\rho} \\ s_{\rho} \\ f_{\rho} \\ p_{\rho} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \end{bmatrix} = 0.$$

This last equation implies Eq. (61).  $\square$

When there is a marginal increase in the perturbation parameter  $\varepsilon^z$ , then at any labour income  $z$ , the tax liability will increase by  $\eta^z(z)$ , and the marginal tax rate will increase by  $\eta_z^z(z)$ . The following Lemma then shows how the behavioural effects such a perturbation can be decomposed into substitution effects and income effects.

**Lemma 11.** *The behavioural effects of any perturbation function  $\varepsilon^z \eta^z$  can be decomposed into substitution and income effects:*

$$k_{\varepsilon^z} = k_{\sigma}^* \eta_z^z - k_{\rho} \eta^z. \quad (65)$$

*Proof.* Note the following second-order derivatives of (42), evaluated in the situation before any reforms, and substitute Eqs. (62) and (63):

$$\mathbb{E}[\tilde{u}_{z\sigma}] = -\mathbb{E}[\tilde{U}_1] - \mathbb{E}[\tilde{u}_{z\rho}]z,$$

$$\begin{aligned}\forall k = s, f, p : E[\tilde{\mathcal{U}}_{k\sigma}] &= -E[\tilde{\mathcal{U}}_{k\rho}]z, \\ \forall k = z, s, f, p : E[\tilde{\mathcal{U}}_{k\epsilon z}] &= (E[\tilde{\mathcal{U}}_{k\sigma}] + zE[\tilde{\mathcal{U}}_{k\rho}])\eta_z^z - E[\tilde{\mathcal{U}}_{k\rho}]\eta^z.\end{aligned}\tag{66}$$

Substitute (66) into Lemma 9 and use Slutsky decomposition (47) to find Eq. (65).  $\square$

In the following Lemma, we apply Lemma 9 to find the effects of a change in the tax on excess returns  $t^e$  on the behaviour of the individuals.

**Lemma 12.** *A change in the tax on excess returns has no effect on labour supply and on savings:*

$$z_{t^e} = s_{t^e} = 0.$$

*A proportional change in the tax on excess returns, causes a proportional change in the investment in both risky assets:*

$$f_{t^e} = \frac{1-t^e}{f}f, \quad p_{t^e} = \frac{1-t^e}{p}p.$$

*Proof.* Note the following second-order derivatives, evaluated in the situation before any reforms, and substitute  $E[\tilde{\mathcal{U}}_{fs}]$  and  $E[\tilde{\mathcal{U}}_{ps}]$  from Eq. (64):

$$E[\tilde{\mathcal{U}}_{fz}] = \frac{1-t^e}{f} \frac{1}{w} E \left[ \left( \tilde{U}_{2\ell} - \tilde{U}_{21} \frac{E[\tilde{U}_\ell]}{E[\tilde{U}_1]} \right) (\tilde{r}^m - r^b) f \right],\tag{67}$$

$$E[\tilde{\mathcal{U}}_{ff}] = \frac{1-t^e}{f} E[(1-t^e)\tilde{U}_{22}(\tilde{r}^m - r^b)^2 f],\tag{68}$$

$$E[\tilde{\mathcal{U}}_{fp}] = \frac{1-t^e}{f} \frac{1-t^e}{p} E[\tilde{U}_{22}(\tilde{r}^m - r^b) f(\alpha + \tilde{\epsilon} - r^b) p],\tag{69}$$

$$E[\tilde{\mathcal{U}}_{pz}] = \frac{1-t^e}{p} \frac{1}{w} E \left[ \left( \tilde{U}_{2\ell} - \tilde{U}_{21} \frac{E[\tilde{U}_\ell]}{E[\tilde{U}_1]} \right) (\alpha + \tilde{\epsilon} - r^b) p \right],\tag{70}$$

$$E[\tilde{\mathcal{U}}_{pp}] = \frac{1-t^e}{p} E[(1-t^e)\tilde{U}_{22}(\alpha + \tilde{\epsilon} - r^b)^2 p],\tag{71}$$

$$\forall k = z, s, f, p : E[\tilde{\mathcal{U}}_{t^e k}] = -\frac{f}{1-t^e} E[\tilde{\mathcal{U}}_{kf}] - \frac{p}{1-t^e} E[\tilde{\mathcal{U}}_{kp}].\tag{72}$$

Apply Cramer's rule to Lemma 9 and substitute  $E[\tilde{\mathcal{U}}_{t^e z}]$  and  $E[\tilde{\mathcal{U}}_{t^e s}]$  from Eq. (72) to find the effects on labour income and savings:

$$z_{t^e} = \frac{1}{\det(E[\tilde{\mathcal{H}}])} \det \left( E \left[ \begin{array}{cccc} \left( \frac{f\tilde{\mathcal{U}}_{zf} + p\tilde{\mathcal{U}}_{zp}}{1-t^e} \right) & \tilde{\mathcal{U}}_{zs} & \tilde{\mathcal{U}}_{zf} & \tilde{\mathcal{U}}_{zp} \\ \left( \frac{f\tilde{\mathcal{U}}_{sf} + p\tilde{\mathcal{U}}_{sp}}{1-t^e} \right) & \tilde{\mathcal{U}}_{ss} & \tilde{\mathcal{U}}_{sf} & \tilde{\mathcal{U}}_{sp} \\ \left( \frac{f\tilde{\mathcal{U}}_{ff} + p\tilde{\mathcal{U}}_{fp}}{1-t^e} \right) & \tilde{\mathcal{U}}_{fs} & \tilde{\mathcal{U}}_{ff} & \tilde{\mathcal{U}}_{fp} \\ \left( \frac{f\tilde{\mathcal{U}}_{pf} + p\tilde{\mathcal{U}}_{pp}}{1-t^e} \right) & \tilde{\mathcal{U}}_{ps} & \tilde{\mathcal{U}}_{pf} & \tilde{\mathcal{U}}_{pp} \end{array} \right] \right) = 0,$$

$$s_{t^e} = \frac{1}{\det(\mathbb{E}[\tilde{\mathcal{H}}])} \det \left( \mathbb{E} \left[ \begin{array}{cccc} \tilde{u}_{zz} & \frac{f\tilde{u}_{zf} + p\tilde{u}_{zp}}{1-t^e} & \tilde{u}_{zf} & \tilde{u}_{zp} \\ \tilde{u}_{sz} & \frac{f\tilde{u}_{sf} + p\tilde{u}_{sp}}{1-t^e} & \tilde{u}_{sf} & \tilde{u}_{sp} \\ \tilde{u}_{fz} & \frac{f\tilde{u}_{ff} + p\tilde{u}_{fp}}{1-t^e} & \tilde{u}_{ff} & \tilde{u}_{fp} \\ \tilde{u}_{pz} & \frac{f\tilde{u}_{pf} + p\tilde{u}_{pp}}{1-t^e} & \tilde{u}_{pf} & \tilde{u}_{pp} \end{array} \right] \right) = 0.$$

Both equations equal zero, because the columns in the right-most determinants are linearly dependent.

To find the effects on investment in market funds and private assets, substitute  $\mathbb{E}[\tilde{u}_{t^e f}]$  and  $\mathbb{E}[\tilde{u}_{t^e p}]$  from (72):

$$f_{t^e} = \frac{1}{\det(\mathbb{E}[\tilde{\mathcal{H}}])} \det \left( \mathbb{E} \left[ \begin{array}{cccc} \tilde{u}_{zz} & \tilde{u}_{zs} & \frac{f\tilde{u}_{zf} + p\tilde{u}_{zp}}{1-t^e} & \tilde{u}_{zp} \\ \tilde{u}_{sz} & \tilde{u}_{ss} & \frac{f\tilde{u}_{sf} + p\tilde{u}_{sp}}{1-t^e} & \tilde{u}_{sp} \\ \tilde{u}_{fz} & \tilde{u}_{fs} & \frac{f\tilde{u}_{ff} + p\tilde{u}_{fp}}{1-t^e} & \tilde{u}_{fp} \\ \tilde{u}_{pz} & \tilde{u}_{ps} & \frac{f\tilde{u}_{pf} + p\tilde{u}_{pp}}{1-t^e} & \tilde{u}_{pp} \end{array} \right] \right) = \frac{f}{1-t^e} k, \quad (73)$$

$$p_{t^e} = \frac{1}{\det(\mathbb{E}[\tilde{\mathcal{H}}])} \det \left( \mathbb{E} \left[ \begin{array}{cccc} \tilde{u}_{zz} & \tilde{u}_{zs} & \tilde{u}_{zf} & \frac{f\tilde{u}_{zf} + p\tilde{u}_{zp}}{1-t^e} \\ \tilde{u}_{sz} & \tilde{u}_{ss} & \tilde{u}_{sf} & \frac{f\tilde{u}_{sf} + p\tilde{u}_{sp}}{1-t^e} \\ \tilde{u}_{fz} & \tilde{u}_{fs} & \tilde{u}_{ff} & \frac{f\tilde{u}_{ff} + p\tilde{u}_{fp}}{1-t^e} \\ \tilde{u}_{pz} & \tilde{u}_{ps} & \tilde{u}_{pf} & \frac{f\tilde{u}_{pf} + p\tilde{u}_{pp}}{1-t^e} \end{array} \right] \right) = \frac{p}{1-t^e}. \quad (74)$$

□

We now use Lemma 12 to prove Lemma 1.

*Proof of Lemma 1.*

1. We found the comparative statics  $f_{t^e} = f/(1+t^e)$  and  $p_{t^e} = p/(1+t^e)$  in Lemma 12. Substituting  $f_{t^e}$  and  $p_{t^e}$  into the uncompensated function (15) for  $\tilde{y}^e$ , yields Eq. (16).
2. We found that  $z_{t^e} = s_{t^e} = 0$  in Lemma 12. Substituting the comparative statics for  $z_{t^e}$ ,  $s_{t^e}$  and  $\tilde{y}_{t^e}^e$  into budget constraints (1) and (4) yields  $\tilde{c}_{t^e}^2 = 0$ .
3. We already showed that  $\mathbb{E}[\tilde{V}_{t^e}] = 0$  in (45).

□

Finally, we show that when the utility function is additively separable between consumption in both periods and between consumption and labour supply, and individual preferences exhibit constant absolute risk aversion, then a change in the expected return to private investment  $\alpha$  only has income effects, on savings and on labour income.

**Lemma 13.** *If the utility function is additively separable:  $U_{12} = U_{1\ell} = U_{2\ell} = 0$ , and absolute risk aversion is*

constant:  $U_{22}/U_2 \equiv A$ , then a change in  $\alpha$  only has income effects, both on savings  $s$  and on labour income  $z$ :

$$\forall k = s, z : \frac{dk}{d\alpha} = (1 - t^e)p \frac{dk}{d\rho_2}. \quad (75)$$

*Proof.* Note the following partial derivatives, taking into account additive separability and substituting (62)–(64):

$$\begin{aligned} \mathbb{E}[\tilde{\mathcal{U}}_{z\alpha}] &= (1 - t^e)p\mathbb{E}[\tilde{\mathcal{U}}_{z\rho_2}], & \mathbb{E}[\tilde{\mathcal{U}}_{s\alpha}] &= (1 - t^e)p\mathbb{E}[\tilde{\mathcal{U}}_{s\rho_2}], \\ \mathbb{E}[\tilde{\mathcal{U}}_{f\alpha}] &= (1 - t^e)p\mathbb{E}[\tilde{\mathcal{U}}_{f\rho_2}], & \mathbb{E}[\tilde{\mathcal{U}}_{p\alpha}] &= (1 - t^e)\mathbb{E}[\tilde{u}_2] + (1 - t^e)p\mathbb{E}[\tilde{\mathcal{U}}_{s\rho_2}]. \end{aligned}$$

Like in the proof of Lemma 12, use Cramer's rule and Lemma 9 to find the following comparative statics:

$$\frac{dz}{d\alpha} = (1 - t^e)p \frac{ds}{d\rho_2} - \frac{(1 - t^e)\mathbb{E}[\tilde{u}_2]}{\det(\mathbb{E}[\tilde{\mathcal{H}}])} \det \left( \mathbb{E} \left[ \begin{array}{ccc} \tilde{\mathcal{U}}_{zs} & \tilde{\mathcal{U}}_{zf} & \tilde{\mathcal{U}}_{zp} \\ \tilde{\mathcal{U}}_{ss} & \tilde{\mathcal{U}}_{sf} & \tilde{\mathcal{U}}_{sp} \\ \tilde{\mathcal{U}}_{fs} & \tilde{\mathcal{U}}_{ff} & \tilde{\mathcal{U}}_{fp} \end{array} \right] \right), \quad (76)$$

$$\frac{ds}{d\alpha} = (1 - t^e)p \frac{ds}{d\rho_2} - \frac{(1 - t^e)\mathbb{E}[\tilde{u}_2]}{\det(\mathbb{E}[\tilde{\mathcal{H}}])} \det \left( \mathbb{E} \left[ \begin{array}{ccc} \tilde{\mathcal{U}}_{zz} & \tilde{\mathcal{U}}_{zf} & \tilde{\mathcal{U}}_{zp} \\ \tilde{\mathcal{U}}_{sz} & \tilde{\mathcal{U}}_{sf} & \tilde{\mathcal{U}}_{sp} \\ \tilde{\mathcal{U}}_{fz} & \tilde{\mathcal{U}}_{ff} & \tilde{\mathcal{U}}_{fp} \end{array} \right] \right). \quad (77)$$

Note the following partial derivatives, use the fact that relative risk aversion  $u_{22}/u_2 \equiv A$  is constant and apply first-order conditions (11)–(12) for portfolio optimization:

$$\begin{aligned} \mathbb{E}[\tilde{\mathcal{U}}_{sf}] &= (1 - t^e)\mathbb{E}[(\tilde{r}^m - r^b)\tilde{u}_{22}] \frac{\mathbb{E}[\tilde{u}_1]}{\mathbb{E}[\tilde{u}_2]} = 0, \\ \mathbb{E}[\tilde{\mathcal{U}}_{sp}] &= (1 - t^e)\mathbb{E}[(\alpha + \tilde{\varepsilon} - r^b)\tilde{u}_{22}] \frac{\mathbb{E}[\tilde{u}_1]}{\mathbb{E}[\tilde{u}_2]} = 0. \end{aligned}$$

Note furthermore that taking into account additive separability, (67) and (70) imply that also  $\mathbb{E}[\tilde{\mathcal{U}}_{zf}] = \mathbb{E}[\tilde{\mathcal{U}}_{zp}] = 0$ . Substitute  $\mathbb{E}[\tilde{\mathcal{U}}_{zf}] = \mathbb{E}[\tilde{\mathcal{U}}_{zp}] = \mathbb{E}[\tilde{\mathcal{U}}_{sf}] = \mathbb{E}[\tilde{\mathcal{U}}_{sp}] = 0$  into (76)–(77) to find that the determinants in the numerators equal zero, and thus (75) is correct.  $\square$

### A.3 Lagrangian for the government optimization problem

The Lagrangian for the government's optimization problem (19) in terms of the perturbation parameters  $\epsilon^z$ ,  $\epsilon^K$ ,  $\epsilon^P$ ,  $t^n$  and  $t^e$  is:

$$\begin{aligned} \Lambda(\epsilon^z, \epsilon^K, \epsilon^P, t^n, t^e) \\ \equiv \iint_{\Omega} \mathbb{E}[\tilde{\mathcal{V}}(\omega, \epsilon^z, t^n, t^e, \epsilon^K, \epsilon^P)] dG^\omega(\omega) - \mathbb{E}[\tilde{\lambda} \cdot \{P(\tilde{r}^m) + K(\tilde{r}^m) + \epsilon^P \eta^P(\tilde{r}^m) + \epsilon^K \eta^K(\tilde{r}^m)\}] \end{aligned}$$

$$\begin{aligned}
& + (1 + r^b) \mathbb{E}[\tilde{\lambda}] \iint_{\Omega} \{t^\ell(z(\boldsymbol{\omega}, \epsilon^z, t^n, t^e, \epsilon^K, \epsilon^P)) + \epsilon^z \eta^z(z(\boldsymbol{\omega}, \epsilon^z, t^n, t^e, \epsilon^K, \epsilon^P))\} dG^\omega(\boldsymbol{\omega}) \\
& + \iint_{\Omega} \{t^n \mathbb{E}[\tilde{\lambda}] y^n(\boldsymbol{\omega}, \epsilon^z, t^n, t^e, \epsilon^K, \epsilon^P) + t^e \mathbb{E}[\tilde{\lambda} \tilde{y}^e(\boldsymbol{\omega}, \epsilon^z, t^n, t^e, \epsilon^K, \epsilon^P)]\} dG^\omega(\boldsymbol{\omega}), \tag{78}
\end{aligned}$$

where we omit superfluous function arguments for legibility. The first-order conditions require that partial derivatives of the Lagrangian with respect to  $\epsilon^z$ ,  $\epsilon^K$ ,  $\epsilon^P$ ,  $t^n$  and  $t^e$  equal zero, for all reform functions  $\eta^z$ ,  $\eta^K$  and  $\eta^P$ . We derive the respective first-order conditions in the following subsections.

#### A.4 First-order condition for the tax on labour income

The first-order condition for  $\epsilon^z$  requires that the condition  $\partial\Lambda/\partial\epsilon^z = 0$  is satisfied for every perturbation function  $\eta^z$ . Evaluate the condition  $\partial\Lambda/\partial\epsilon^z = 0$  for the Lagrangian (78) in the situation before any reforms, substitute envelope condition (44) for  $\mathbb{E}[\tilde{\mathcal{V}}_{\epsilon^z}]$ , substitute decomposition (65) for  $\forall k = z, y^n, y^e : k_\epsilon^z$ , and substitute definition (20) for  $\beta(\boldsymbol{\omega})$  and definition (21) for  $\mathcal{W}(\boldsymbol{\omega})$ :

$$0 = \iint_{\Omega} (1 + r^b - \beta(\boldsymbol{\omega})) \eta^z(z(\boldsymbol{\omega})) dG^\omega(\boldsymbol{\omega}) - \iint_{\Omega} \mathcal{W}(\boldsymbol{\omega}) \eta_z^z(z(\boldsymbol{\omega})) dG^\omega(\boldsymbol{\omega}). \tag{79}$$

To further interpret condition (79), we will rewrite it as an integral over labour incomes  $z$ . Condition (79) becomes:

$$0 = \int_{\underline{z}}^{\bar{z}} (1 + r^b - \bar{\beta}(z)) \eta^z(z) dG^z(z) - \int_{\underline{z}}^{\bar{z}} \bar{\mathcal{W}}(z) \eta_z^z(z) dG^z(z). \tag{80}$$

Use partial integration to rewrite the second integral of (80):

$$\begin{aligned}
0 = \int_{\underline{z}}^{\bar{z}} \left\{ (1 + r^b - \bar{\beta}(z)) g^z(z) + \frac{d[\bar{\mathcal{W}}(z) g^z(z)]}{dz} \right\} \eta^z(z) dz \\
- \bar{\mathcal{W}}(\bar{z}) g^z(\bar{z}) \eta^z(\bar{z}) + \bar{\mathcal{W}}(\underline{z}) g^z(\underline{z}) \eta^z(\underline{z}). \tag{81}
\end{aligned}$$

The latter condition must be valid for every perturbation function  $\eta^z$ . Let us first focus on the perturbation functions that are zero on the boundaries,  $\eta^z(\underline{z}) = \eta^z(\bar{z}) = 0$ . The second line of (81) is then zero. Consequently, the first line of (81) must also be zero. We will prove by contradiction that the terms between curly brackets of (81) must then sum to zero. Suppose that these terms do not sum to zero on some interval between  $\underline{z}$  and  $\bar{z}$ . Choose then a perturbation function  $\eta^z$  which has the same sign everywhere as the sum of the terms between curly brackets. The integral on the first line of (81) is then strictly positive. This contradicts the requirement that the first line is zero. It is thus impossible that the first-order condition is satisfied for every perturbation function  $\eta^z$  if the expression between curly brackets differs from zero on some interval. The latter reasoning is an application of the *fundamental theorem of the calculus of variations*. Requiring that the terms between the curly brackets sum to zero yields the *Euler-Lagrange equation* that characterizes the optimal tax on labour

income for any labour income  $z$ :

$$(1 + r^b - \bar{\beta}(z))g^z(z) = -\frac{d[\bar{W}(z)g^z(z)]}{dz}. \quad (82)$$

Having established that the first line of (81) must be zero, let us now consider perturbation functions  $\eta^z$  which differ from zero at one of the end points,  $\eta^z(\underline{z}) \neq 0$  or  $\eta^z(\bar{z}) \neq 0$ . Since the second line of (81) must be zero, it follows that the tax wedges on the end points must be zero. We thus find the following transversality conditions:

$$\bar{W}(\bar{z}) = \bar{W}(\underline{z}) = 0. \quad (83)$$

Integrating (82) and using transversality conditions (83), we find first-order condition (23) for the tax on labour income. Substitute transversality condition (83) into first-order condition (23) to find condition (24) for the welfare weights.

## A.5 First-order condition for the tax on risk-free capital income

The first-order condition for the tax rate on risk-free capital income requires that  $\partial\Lambda/\partial t^n = 0$ :

$$\iint_{\Omega} \left\{ \left( 1 - \frac{E[\tilde{U}_2]}{E[\tilde{\lambda}]} \right) y^n + (1 + r^b)t_z^\ell z_{t^n} + t^n y_{t^n}^n + t^e \frac{E[\tilde{\lambda}\tilde{y}_{t^n}^e]}{E[\tilde{\lambda}]} \right\} dG^\omega(\omega) = 0,$$

where we use envelope condition (45), and we evaluate in the situation before any reforms. Apply Slutsky decompositions (47) to find:

$$\iint_{\Omega} \left\{ (1 + r^b - \beta(\omega)) \frac{E[\tilde{U}_2]}{E[\tilde{U}_1]} y^n + (1 + r^b)t_z^\ell z_{t^n}^* + t^n y_{t^n}^{n*} + t^e \frac{E[\tilde{\lambda}\tilde{y}_{t^n}^{e*}]}{E[\tilde{\lambda}]} \right\} dG^\omega(\omega) = 0,$$

where we use property (61) to substitute for second-period income effects, we substitute Euler equation (10) and we substitute definition (20) of  $\beta(\omega)$ .

Change the integration domain to the labour incomes and rearrange:

$$\int_{\underline{z}}^{\bar{z}} \left\{ \frac{E[\tilde{U}_2]}{E[\tilde{U}_1]} \iint_{\Omega} (1 + r^b - \beta(\omega)) y^n dG^{\omega|z}(\omega|z) + (1 + r^b)t_z^\ell \bar{z}_{t^n}^* + t^n \bar{y}_{t^n}^{n*} + t^e \frac{E[\tilde{\lambda}\tilde{y}_{t^n}^{e*}]}{E[\tilde{\lambda}]} \right\} dG^z(z) = 0, \quad (84)$$

where we use the constancy of  $E[\tilde{U}_2]/E[\tilde{U}_1]$  due to Euler equation (10).

Let us focus now on the first term in the integrand of (84). Write this term as a covariance over the population  $\Omega$ , conditional on labour income:

$$\int_{\underline{z}}^{\bar{z}} \iint_{\Omega} (1 + r^b - \beta(\omega)) y^n dG^{\omega|z}(\omega|z) dG^z(z) = - \int_{\underline{z}}^{\bar{z}} \{ \text{cov}_{\Omega}(\beta, y^n|z) - (1 + r^b - \bar{\beta}(z)) \bar{y}^n \} dG^z(z). \quad (85)$$

Use integration by parts and use transversality condition (24) to rewrite the right-hand side of (85):

$$\int_{\underline{z}}^{\bar{z}} \iint_{\Omega} (1 + r^b - \beta(\omega)) y^n dG^{\omega|z}(\omega|z) dG^z(z) = \int_{\underline{z}}^{\bar{z}} \left\{ \frac{\int_{\underline{z}}^{\bar{z}} (1 + r^b - \bar{\beta}(\hat{z})) dG^z(\hat{z}) d\bar{y}^n}{g^z(z)} - \text{cov}_{\Omega}(\beta, y^n|z) \right\} dG^z(z). \quad (86)$$

Substitute (86) into optimality condition (84):

$$\begin{aligned} \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega}(\beta, y^n|z) dG^z(z) &= \int_{\underline{z}}^{\bar{z}} \int_{\underline{z}}^{\bar{z}} (1 + r^b - \bar{\beta}(\hat{z})) dG^z(\hat{z}) \frac{d\bar{y}^n}{dz} dz \\ &\quad + \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} \int_{\underline{z}}^{\bar{z}} \left( (1 + r^b) t_z^{\ell} \bar{z}_{t^n}^* + t^n \bar{y}_{t^n}^{n*} + t^e \frac{\mathbb{E}[\tilde{\lambda} \tilde{y}_{t^n}^{e*}]}{\mathbb{E}[\tilde{\lambda}]} \right) dG^z(z). \end{aligned}$$

Substitute the optimal labour tax condition (23) and the definition of the labour wedge (21), and use the fact that preferences are separable between leisure and consumption ( $d\bar{y}^n/dz = \bar{y}_z^{nc}$ ):

$$\begin{aligned} \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega}(\beta, y^n|z) dG^z(z) &= \int_{\underline{z}}^{\bar{z}} (1 + r^b) t_z^{\ell} \left( \frac{\bar{z}_{t^n}^* \mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} - \bar{z}_{\sigma}^* \bar{y}_z^{nc} \right) dG^z(z) + t^n \int_{\underline{z}}^{\bar{z}} \left( \frac{\bar{y}_{t^n}^{n*} \mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} - \bar{y}_{\sigma}^{n*} \bar{y}_z^{nc} \right) dG^z(z) \\ &\quad + t^e \frac{\mathbb{E} \left[ \tilde{\lambda} \int_{\underline{z}}^{\bar{z}} \left( \frac{\tilde{y}_{t^n}^{e*} \mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} - \tilde{y}_{\sigma}^{e*} \bar{y}_z^{nc} \right) dG^z(z) \right]}{\mathbb{E}[\tilde{\lambda}]}. \end{aligned} \quad (87)$$

We rewrite the terms on the right-hand side of (87) one by one. First, use Slutsky symmetry (48) to rewrite:

$$\bar{z}_{t^n}^* \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} - \bar{z}_{\sigma}^* \bar{y}_z^{nc} = \bar{y}_{\sigma}^* - \bar{z}_{\sigma}^* \bar{y}_z^{nc} = \overline{(y_{\sigma}^{n*} - z_{\sigma}^* y_z^{nc})} + \text{cov}_{\Omega}(z_{\sigma}^*, y_z^{nc}|z) = 0. \quad (88)$$

where the last step follows from assumption (22) and the fact that  $y_{\sigma}^{n*} = y_z^{nc} z_{\sigma}^*$ . Second, use Slutsky symmetry (48) to rewrite:

$$\bar{y}_{t^n}^{n*} \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} - \bar{y}_{\sigma}^{n*} \bar{y}_z^{nc} = \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} (\bar{y}_{t^n}^{n*} - \bar{z}_{t^n}^* \bar{y}_z^{nc}) = \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} (\overline{(y_{t^n}^{n*} - z_{t^n}^* y_z^{nc})} + \text{cov}_{\Omega}(z_{t^n}^*, y_z^{nc}|z)) = \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} \bar{y}_{t^n}^{nc*}, \quad (89)$$

where the last step follows from assumption (22). Third, note that  $y_{\sigma}^{n*} = y_z^{nc} z_{\sigma}^*$  and  $\tilde{y}_{\sigma}^{e*} = \tilde{y}_z^{ec} z_{\sigma}^*$ , and thus  $\tilde{y}_{\sigma}^{e*} y_z^{nc} = y_{\sigma}^{n*} \tilde{y}_z^{ec}$ . Use this equality and assumption (22) to rewrite:

$$\frac{\tilde{y}_{t^n}^{e*} \mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} - \tilde{y}_{\sigma}^{e*} \bar{y}_z^{nc} = \left( \frac{\tilde{y}_{t^n}^{e*} \mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} - y_{\sigma}^{n*} \tilde{y}_z^{ec} \right) = \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} \overline{(\tilde{y}_{t^n}^{e*} - z_{t^n}^* \tilde{y}_z^{ec})} = \frac{\mathbb{E}[\tilde{U}_1]}{\mathbb{E}[\tilde{U}_2]} \tilde{y}_{t^n}^{ec*}, \quad (90)$$

where we apply Slutsky symmetry (48) in the second step. Substitute (88)–(90) into (87) to find condition (25).



## A.6 First-order condition for the tax on excess capital income

The first-order condition for the tax on excess capital income requires  $\partial\Lambda/\partial t^e = 0$  in the situation before any reforms:

$$\iint_{\Omega} \{E[\tilde{\lambda}\tilde{y}^e] + t^e E[\tilde{\lambda}\tilde{y}_{t^e}^e]\} dG^{\omega}(\omega) = 0.$$

Substitute (16):

$$\iint_{\Omega} E[\tilde{\lambda}\tilde{y}^e] dG^{\omega}(\omega) = 0 \Leftrightarrow \iint_{\Omega} \frac{\text{cov}(\tilde{\lambda}, \tilde{y}^e)}{E[\tilde{\lambda}]} dG^{\omega}(\omega) = - \iint_{\Omega} E[\tilde{y}^e] dG^{\omega}(\omega). \quad (91)$$

Substitute (13) into (91) to find (26).

## A.7 First-order condition for the ex-post provision of the public good

The government's first-order condition for the ex-post provision of the public good,  $\partial\Lambda/\partial\epsilon^P = 0$ , requires that the following condition is satisfied:

$$E \left[ \left( \iint_{\Omega} \tilde{U}_P dG^{\omega}(\omega) - \tilde{\lambda} \right) \tilde{\eta}^P \right] = 0, \quad (92)$$

where we use envelope property (46) and the fact that the level of public good provision does not affect individual behaviour. The latter expression must be true for every perturbation function  $\tilde{\eta}^P$ . Applying similar reasoning as we did for the first-order condition for the labour income tax, using the fundamental theorem of the calculus of variation, we find government's first-order condition (27) for the optimal provision of the public good.

## A.8 First-order condition for the ex-post provision of the lump sum

The first-order condition  $\partial\Lambda/\partial\epsilon^K = 0$  for the ex-post provision of the lump sum requires:

$$\frac{E[\tilde{\lambda}\eta^K(\tilde{r}^m)]}{E[\tilde{\lambda}]} - \iint_{\Omega} \frac{E[\tilde{U}_2\eta^K(\tilde{r}^m)]}{E[\tilde{\lambda}]} dG^{\omega}(\omega) = \iint_{\Omega} \left\{ (1+r^b)t^{\ell}z_{\epsilon^K} + t^n y_{\epsilon^K}^n + t^e \frac{E[\tilde{\lambda}\tilde{y}_{\epsilon^K}^e]}{E[\tilde{\lambda}]} \right\} dG^{\omega}(\omega), \quad (93)$$

where we use envelope property (46).

We first rewrite the integrand on the right-hand side of (93). Write it in matrix notation:

$$\begin{aligned} & (1+r^b)t^{\ell}z_{\epsilon^K} + t^n r^b s_{\epsilon^K} + t^e \frac{E[\tilde{\lambda}(\tilde{r}^m - r^b)]}{E[\tilde{\lambda}]} f_{\epsilon^K} + t^e \frac{E[\tilde{\lambda}(\alpha + \tilde{\epsilon} - r^b)]}{E[\tilde{\lambda}]} p_{\epsilon^K} \\ & = - \left( (1+r^b)t^{\ell} \quad t^n r^b \quad t^e \frac{E[\tilde{\lambda}(\tilde{r}^m - r^b)]}{E[\tilde{\lambda}]} \quad t^e \frac{E[\tilde{\lambda}(\alpha + \tilde{\epsilon} - r^b)]}{E[\tilde{\lambda}]} \right) \cdot (E[\tilde{\mathcal{H}}])^{-1} \cdot E \left[ \left( \tilde{U}_{z\rho_2} \quad \tilde{U}_{s\rho_2} \quad \tilde{U}_{f\rho_2} \quad \tilde{U}_{p\rho_2} \right)^T \tilde{\eta}^K \right], \end{aligned} \quad (94)$$

where we apply Lemma 9, and the fact that for any  $k = z, s, f, p$ :  $\tilde{U}_{k\epsilon^K} = \tilde{U}_{k\rho_2}\tilde{\eta}^K$ .

Substitute (94) into (93), apply the fundamental theorem of the calculus of variations as we did in the proof for the first-order condition for the tax on labour income, substitute for any  $k = z, s, f, p$ :  $\tilde{U}_{k\kappa^m} = \tilde{U}_{k\rho_2}\delta(\tilde{r}^m - r^m)$  and apply Lemma 9 to find the following condition for any realization  $r^m$  of the market risk:

$$\frac{\lambda^{r^m}}{\mathbb{E}[\tilde{\lambda}]} - \iint_{\Omega} \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{\lambda}]} dG^{\omega}(\omega) = \iint_{\Omega} \{(1+r^b)t^{\ell}z_{\kappa^m} + t^n y_{\kappa^m}^n + t^e \mathcal{A}_{\kappa^m}\} dG^{\omega}(\omega).$$

Apply Slutsky decomposition (47), substitute property (61) for the second-period income effects and substitute definition (20) of  $\beta(\omega)$ :

$$\begin{aligned} \frac{\lambda^{r^m}}{\mathbb{E}[\tilde{\lambda}]} - \iint_{\Omega} \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_2]} dG^{\omega}(\omega) &= - \iint_{\Omega} \left\{ \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} (1+r^b - \beta(\omega)) \right\} dG^{\omega}(\omega) \\ &+ \iint_{\Omega} \{(1+r^b)t^{\ell}z_{\kappa^m}^* + t^n y_{\kappa^m}^{n*} + t^e \mathcal{A}_{\kappa^m}^*\} dG^{\omega}(\omega). \end{aligned} \quad (95)$$

Let us now focus on the first integral on the right-hand side. Rewrite the integration limits:

$$\begin{aligned} \iint_{\Omega} \left\{ \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} (1+r^b - \beta(\omega)) \right\} dG^{\omega}(\omega) &= \int_{\underline{z}}^{\bar{z}} \int_{\underline{\alpha}}^{\bar{\alpha}} \left\{ \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} (1+r^b - \beta(\omega)) \right\} dG^{\alpha|z}(\alpha|z) dG^z(z) \\ &= - \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]}, \beta \middle| z \right) dG^z(z) \\ &+ \int_{\underline{z}}^{\bar{z}} \overline{\left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right)} (1+r^b - \bar{\beta}(z)) dG^z(z). \end{aligned}$$

Apply partial integration on the last line and use transversality condition (24):

$$\int_{\underline{z}}^{\bar{z}} \overline{\left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right)} (1+r^b - \bar{\beta}(z)) dG^z(z) = \int_{\underline{z}}^{\bar{z}} \frac{d}{dz} \overline{\left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right)} \int_z^{\bar{z}} (1+r^b - \bar{\beta}(\hat{z})) dG^z(\hat{z}) dz.$$

Substitute these results into (95):

$$\begin{aligned} \frac{\lambda^{r^m}}{\mathbb{E}[\tilde{\lambda}]} - \iint_{\Omega} \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_2]} dG^{\omega}(\omega) &= \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]}, \beta \middle| z \right) dG^z(z) \\ &- \int_{\underline{z}}^{\bar{z}} \frac{d}{dz} \overline{\left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right)} \int_z^{\bar{z}} (1+r^b - \bar{\beta}(\hat{z})) dG^z(\hat{z}) dz \\ &+ \iint_{\Omega} \{(1+r^b)t^{\ell}z_{\kappa^m}^* + t^n y_{\kappa^m}^{n*} + t^e \mathcal{A}_{\kappa^m}^*\} dG^{\omega}(\omega). \end{aligned}$$

Substitute the optimal labour income tax (23) and definition (21) of the labour wedge on the second line:

$$\frac{\lambda^{r^m}}{\mathbb{E}[\tilde{\lambda}]} - \iint_{\Omega} \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_2]} dG^{\omega}(\omega) = \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]}, \beta \middle| z \right) dG^z(z)$$

$$\begin{aligned}
& + \int_{\underline{z}}^{\bar{z}} \frac{d}{dz} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right) \left( (1+r^b)t_z^\ell \bar{z}_\sigma^* + t^n \bar{y}_\sigma^{n*} + t^e \bar{\mathcal{A}}_\sigma^* \right) dG^z(z) \\
& + \int_{\underline{z}}^{\bar{z}} \left\{ (1+r^b)t_z^\ell \bar{z}_{\kappa^m}^* + t^n \bar{y}_{\kappa^m}^{n*} + t^e \bar{\mathcal{A}}_{\kappa^m}^* \right\} dG^z(z).
\end{aligned}$$

Using Slutsky symmetries (48)–(50), separability of preferences between leisure and labour supply and assumption (37), following steps analogous to those in section A.5, we find:

$$\begin{aligned}
\frac{\lambda r^m}{\mathbb{E}[\tilde{\lambda}]} - \iint_{\Omega} \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_2]} dG^\omega(\omega) &= \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]}, \beta \middle| z \right) dG^z(z) \\
& + \int_{\underline{z}}^{\bar{z}} t^n \bar{y}_{\kappa^m}^{nc*} dG^z(z) \\
& + \int_{\underline{z}}^{\bar{z}} t^e \left( \mathcal{A}_{\kappa^m}^* + \frac{d}{dz} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right) \mathcal{A}_\sigma^* \right) dG^z(z).
\end{aligned} \tag{96}$$

Note the following identities:

$$\frac{d}{d\sigma} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right)^* = \frac{d}{dz} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right) z_\sigma^*,$$

and  $\tilde{y}_\sigma^{e*} = \tilde{y}_z^{ec} z_\sigma^*$ , and thus:

$$\frac{d}{d\sigma} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right)^* \mathcal{A}_z^c = \mathcal{A}_\sigma^* \frac{d}{dz} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|r^m]}{\mathbb{E}[\tilde{U}_1]} \right).$$

Substitute this into (96) to find (38).

Substituting (38) into (26) we obtain:

$$\begin{aligned}
\iint_{\Omega} \frac{\text{cov}(\tilde{U}_2, \tilde{y}^e)}{\mathbb{E}[\tilde{U}_2]} dG^\omega(\omega) &= \text{cov} \left( \iint_{\Omega} \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|\tilde{r}^m]}{\mathbb{E}[\tilde{U}_2]} dG^\omega(\omega), \iint_{\Omega} \tilde{y}^e dG^\omega(\omega) \right) \\
& + \text{cov} \left( \iint_{\Omega} \left( t^n \bar{y}_{\kappa^m}^{nc*} + t^e \bar{\mathcal{A}}_{\kappa^m}^* \right) dG^\omega(\omega), \iint_{\Omega} \tilde{y}^e dG^\omega(\omega) \right) \\
& + \frac{\mathbb{E}[\tilde{U}_2]}{\mathbb{E}[\tilde{U}_1]} \text{cov} \left( \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|\tilde{r}^m]}{\mathbb{E}[\tilde{U}_2]}, \beta \middle| z \right) dG^z(z), \iint_{\Omega} \tilde{y}^e dG^\omega(\omega) \right).
\end{aligned} \tag{97}$$

We rewrite the terms on the right-hand side of (97) one by one. Simplify the first line:

$$\begin{aligned}
\text{cov} \left( \iint_{\Omega} \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|\tilde{r}^m]}{\mathbb{E}[\tilde{U}_2]} dG^\omega(\omega), \iint_{\Omega} \tilde{y}^e dG^\omega(\omega) \right) &= \iint_{\Omega} \frac{\text{cov}(\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|\tilde{r}^m], \tilde{r}^m)}{\mathbb{E}[\tilde{U}_2]} dG^\omega(\omega) \iint_{\Omega} f dG^\omega(\omega) \\
&= - \iint_{\Omega} \mathbb{E}[(\tilde{r}^m - r^b) f] dG^\omega(\omega),
\end{aligned}$$

where we use first-order condition (11) in the last step. Now rewrite the second line on the right-hand side of

(97) to find that it equals zero:

$$\begin{aligned}
& \text{cov} \left( \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|\tilde{r}^m]}{\mathbb{E}[\tilde{U}_2]}, \beta \middle| z \right) dG^z(z), \iint_{\Omega} \tilde{y}^e dG^{\omega}(\omega) \right) \\
&= \int_{\underline{z}}^{\bar{z}} \text{cov} \left( \text{cov}_{\Omega} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|\tilde{r}^m]}{\mathbb{E}[\tilde{U}_2]}, \beta \middle| z \right), \tilde{r}^m \right) dG^z(z) \iint_{\Omega} f dG^{\omega}(\omega) \\
&= \int_{\underline{z}}^{\bar{z}} \text{cov}_{\Omega} \left( \text{cov} \left( \frac{\mathbb{E}_{\mathcal{E}}[\tilde{U}_2|\tilde{r}^m]}{\mathbb{E}[\tilde{U}_2]}, \tilde{r}^m \right), \beta \middle| z \right) dG^z(z) \iint_{\Omega} f dG^{\omega}(\omega) = 0,
\end{aligned}$$

where the last step follows from first-order condition (11). Finally, simplify the third line on the right-hand side of (97):

$$\begin{aligned}
\text{cov} \left( \iint_{\Omega} \left( t^n \bar{y}_{\tilde{K}}^{nc*} + t^e \bar{\mathcal{A}}_{\tilde{K}}^{c*} \right) dG^{\omega}(\omega), \iint_{\Omega} \tilde{y}^e dG^{\omega}(\omega) \right) &= t^n \iint_{\Omega} \text{cov}(\bar{y}_{\tilde{K}}^{nc*}, \tilde{r}^m) dG^{\omega}(\omega) \iint_{\Omega} f dG^{\omega}(\omega) \\
&+ t^e \iint_{\Omega} \text{cov}(\bar{\mathcal{A}}_{\tilde{K}}^{c*}, \tilde{r}^m) dG^{\omega}(\omega) \iint_{\Omega} f dG^{\omega}(\omega).
\end{aligned}$$

Substitute these results into (97) to obtain (39).

## A.9 Proof of Lemma 2

*Proof of Lemma 2.* Use (15) and the fact that the private shocks are i.i.d. to find  $\text{cov}(\tilde{\lambda}, \tilde{y}_{\nu}^e) = \text{cov}(\tilde{\lambda}, \tilde{r}^m) f_{\nu}$ . Multiply and divide the right-hand side by  $\iint_{\Omega} f dG^{\omega}(\omega)$  to find:

$$\frac{\text{cov}(\tilde{\lambda}, \tilde{y}_{\nu}^e)}{\mathbb{E}[\tilde{\lambda}]} = \frac{f_{\nu}}{\iint_{\Omega} f dG^{\omega}(\omega)} \iint_{\Omega} \frac{\text{cov}(\tilde{\lambda}, \tilde{y}^e)}{\mathbb{E}[\tilde{\lambda}]} dG^{\omega}(\omega).$$

Substitute the government's optimality condition (26) and individual first-order conditions (11)–(12):

$$\frac{\text{cov}(\tilde{\lambda}, \tilde{y}_{\nu}^e)}{\mathbb{E}[\tilde{\lambda}]} = - \frac{f_{\nu}}{\iint_{\Omega} f dG^{\omega}(\omega)} \iint_{\Omega} \mathbb{E}[\tilde{y}^e] dG^{\omega}(\omega).$$

We thus find:

$$t^e \iint_{\Omega} \left( \mathbb{E}[\tilde{y}_{\nu}^e] + \frac{\text{cov}(\tilde{\lambda}, \tilde{y}_{\nu}^e)}{\mathbb{E}[\tilde{\lambda}]} \right) dG^{\omega}(\omega) = t^e \iint_{\Omega} \left( \mathbb{E}[\tilde{y}_{\nu}^e] - \frac{f_{\nu}}{\iint_{\Omega} f dG^{\omega}(\omega)} \iint_{\Omega} \mathbb{E}[\tilde{y}^e] dG^{\omega}(\omega) \right) dG^{\omega}(\omega)$$

Substitute (15) on the right-hand side:

$$t^e \iint_{\Omega} \left( \mathbb{E}[\tilde{y}_{\nu}^e] + \frac{\text{cov}(\tilde{\lambda}, \tilde{y}_{\nu}^e)}{\mathbb{E}[\tilde{\lambda}]} \right) dG^{\omega}(\omega) = t^e \left( \iint_{\Omega} p_{\nu}(\alpha - r^b) dG^{\omega}(\omega) - \frac{\iint_{\Omega} p(\alpha - r^b) dG^{\omega}(\omega)}{\iint_{\Omega} f dG^{\omega}(\omega)} \iint_{\Omega} f_{\nu} dG^{\omega}(\omega) \right).$$

Verify that the latter corresponds to Lemma 2.  $\square$

## Appendix B The Mean-Variance framework

We apply the *Mean-Variance* approach, introduced by Markowitz (1952), to our model. Suppose the returns of both risky assets are jointly normally distributed, so any linear combination of them has a univariate normal distribution. The second-period individual budget constraint (9) then implies that second-period consumption is normally distributed. Government budget constraint (34) shows that second-period consumption remains normally distributed even if the government returns the aggregate risk using a stochastic lump sum. This allows us to study the individual portfolio optimization problem in the Mean-Variance framework.

Taking labour income  $z$  and the size of the portfolio  $s$  as given and using the separability properties of individual preferences, the portfolio optimization problem consists of choosing the amounts invested in assets  $f$  and  $p$  to maximize the expected utility from consumption in the second period:

$$\max_{p,f} \mathbb{E}[u(c^1, \tilde{c}^2)|c^1], \quad (98)$$

where  $\tilde{c}^2$  is given by (4). The amount  $b$  invested in bonds follows as a residual from (2).

Assume that second-period consumption is normally distributed, with expected value  $\mathbb{E}[\tilde{c}^2]$  and standard deviation  $\text{sd}[\tilde{c}^2]$ . Then the following stochastic quantity follows the standardized normal distribution:

$$\tilde{n} \equiv \frac{\tilde{c}^2 - \mathbb{E}[\tilde{c}^2]}{\text{sd}[\tilde{c}^2]}.$$

Denote the probability density function of  $\tilde{n}$  as  $\varphi(\tilde{n})$ . The objective function (98) can then be rewritten in terms of the mean and the standard deviation of second-period consumption:

$$\mathbb{E}[u(c^1, \tilde{c}^2)|c^1] = \int_{-\infty}^{+\infty} u(c^1, \mathbb{E}[\tilde{c}^2] + n \cdot \text{sd}[\tilde{c}^2])\varphi(n)dn. \quad (99)$$

Solving this portfolio optimization problem leads to the following Lemma.

**Lemma 14.** *In the Mean-Variance framework, the ratio of investments in the risky assets is determined by:*

$$\frac{p}{f} = \left( 1 + \frac{\text{cov}(\tilde{K}, (1-t^e)f\tilde{r}^m)}{\text{var}((1-t^e)f\tilde{r}^m)} \right) \frac{\alpha - r^b}{\mathbb{E}[\tilde{r}^m] - r^b} \frac{\text{var}(\tilde{r}^m)}{\text{var}(\tilde{\varepsilon})}. \quad (100)$$

*In the case without a stochastic lump sum, so  $\text{cov}(\tilde{K}, \tilde{r}^m) = 0$ , the relative proportion  $p/f$  invested in the risky assets is increasing in private investment productivity  $\alpha$ , and is unaffected by skill, taxes or the size of the portfolio.*

*Proof.* We start from equation (99). For any asset  $k = f, p$ , we find first-order condition:

$$0 = \int_{-\infty}^{+\infty} u \left( c^1, \frac{\partial \mathbb{E}[\tilde{c}^2]}{\partial k} + n \cdot \frac{\partial \text{sd}[\tilde{c}^2]}{\partial k} \right) \varphi(n)dn. \quad (101)$$

Rewrite  $\partial \text{sd}[\tilde{c}^2]/\partial k$ , for any  $k = f, p$  :

$$\begin{aligned}
\frac{\partial \text{sd}[\tilde{c}^2]}{\partial k} &= \frac{1}{2\sqrt{\text{var}[\tilde{c}^2]}} \frac{\partial \text{var}[\tilde{c}^2]}{\partial k} \\
&= \frac{1}{2\text{sd}[\tilde{c}^2]} \frac{\partial}{\partial k} \left( \text{E} \left[ (\tilde{c}^2)^2 \right] - \text{E} [\tilde{c}^2]^2 \right) \\
&= \frac{1}{\text{sd}[\tilde{c}^2]} \left( \text{E} \left[ \tilde{c}^2 \frac{\partial \tilde{c}^2}{\partial k} \right] - \text{E} [\tilde{c}^2] \text{E} \left[ \frac{\partial \tilde{c}^2}{\partial k} \right] \right) \\
&= \frac{\text{cov} \left( \tilde{c}^2, \frac{\partial \tilde{c}^2}{\partial k} \right)}{\text{sd}[\tilde{c}^2]}. \tag{102}
\end{aligned}$$

Substitute (102) into first-order condition (101) and rearrange:

$$\frac{\text{E} \left[ \frac{\partial \tilde{c}^2}{\partial k} \right] \text{sd}[\tilde{c}^2]}{\text{cov} \left( \tilde{c}^2, \frac{\partial \tilde{c}^2}{\partial k} \right)} = - \frac{\int_{-\infty}^{+\infty} u_2 n \varphi(n) dn}{\int_{-\infty}^{+\infty} u_2 \varphi(n) dn}.$$

The right-hand side is the same for any asset. Combine the left-hand side for assets  $k = p$  and  $k = f$  to find optimality condition:

$$\frac{\text{E} \left[ \frac{\partial \tilde{c}^2}{\partial f} \right]}{\text{cov} \left( \tilde{c}^2, \frac{\partial \tilde{c}^2}{\partial f} \right)} = \frac{\text{E} \left[ \frac{\partial \tilde{c}^2}{\partial p} \right]}{\text{cov} \left( \tilde{c}^2, \frac{\partial \tilde{c}^2}{\partial p} \right)}.$$

Substitute partial derivatives  $\partial \tilde{c}^2/\partial f$  and  $\partial \tilde{c}^2/\partial p$  of the second-period budget constraint (9):

$$\frac{\text{E}[\tilde{r}^m - r^b]}{\text{cov}(\tilde{c}^2, \tilde{r}^m)} = \frac{\alpha - r^b}{\text{cov}(\tilde{c}^2, \tilde{\varepsilon})}.$$

Substitute second-period budget constraint (9) for  $\tilde{c}^2$ :

$$\frac{\text{E}[\tilde{r}^m - r^b]}{(1 - t^e)\text{var}(\tilde{r}^m) f + \text{cov}(\tilde{K}, \tilde{r}^m)} = \frac{\alpha - r^b}{(1 - t^e)\text{var}(\tilde{\varepsilon}) p}.$$

Rearrange to find (100). □

As long as the government does not return its risk by lump sums, we find the standard result that the relative proportions invested in the risky assets are determined only by the mean and the variance of the excess returns of these assets. Individuals with higher expected returns on private investment undertake more private investment. If the government returns the risk of its investments using a stochastic lump sum, then the returns to the market asset are correlated to the stochastic lump sum. Individuals will prefer to diversify their portfolio away from the market asset, towards the private asset that carries idiosyncratic risk. In this case the separation between the composition of the risky portfolio and other factors is no longer valid.

Denote the semi-elasticity of the investment  $\gamma = f, p$  with respect to some perturbation parameter  $\nu$  as  $\xi_\nu^\gamma \equiv \partial \log \gamma / \partial \nu$ . Taking logarithms on both sides of (100) and taking derivatives with respect to the perturbation parameter, results in the following lemma.

**Lemma 15.** *In the Mean-Variance framework, the semi-elasticities of investment in the market asset are related to the corresponding semi-elasticities of investment in the private asset as follows:*

$$\forall \nu = \sigma, t^n, \rho: \xi_\nu^f = \xi_\nu^p \cdot \left( 1 + \frac{\text{cov}(\tilde{K}, (1-t^e)f\tilde{r}^m)}{\text{var}((1-t^e)f\tilde{r}^m)} \right).$$

*This result extends to the corresponding compensated semi-elasticities, and to semi-elasticities conditional on labour income.*

It follows that without a stochastic lump sum, if individuals behave as predicted by the Mean-Variance framework, then the amounts invested in private and market assets move proportionally and in the same direction. This is in line with Lemma 14. As soon as the government returns part of the risk through a stochastic lump sum, so  $\text{cov}(\tilde{K}, \tilde{r}^m) > 0$ , we find  $|\xi_\nu^f| > |\xi_\nu^p|$  if  $\xi_\nu^p \neq 0$ . Investments in both assets still move in the same direction, but investment in the market assets responds more strongly to changes in an exogenous parameter  $\nu$ . One way to understand this is as follows. Suppose that due to a change in a parameter  $\nu$ , the investment in the risky portfolio, and thus in both risky assets, increases. As this happens, the relative importance of the risk of the stochastic lump sum decreases, so the term between brackets in (100) becomes relatively less important. The individual will end up investing larger share of the risky portfolio in the market asset.

Let us now verify the consequences of Lemma 15 for our model. First assume that  $\xi_\nu^p$  and  $\xi_\nu^f$  are constant over the income distribution. Lemma 2 then implies:

$$\iint_{\Omega} \mathcal{A}_\nu dG^\omega(\omega) = \iint_{\Omega} (\alpha - r^b) p dG^\omega(\omega) \cdot (\xi_\nu^f - \xi_\nu^p).$$

Without a stochastic lump sum, Lemma 15 then implies:

$$\iint_{\Omega} \mathcal{A}_\nu dG^\omega(\omega) = 0.$$

With a stochastic lump sum, we find:

$$\iint_{\Omega} \mathcal{A}_\nu dG^\omega(\omega) \gtrless 0 \iff \xi_\nu^p \gtrless 0.$$

Assume instead that  $\alpha$  is constant, so all individuals obtain the same expected rate of return to their private investments. Lemma 2 then implies:

$$\iint_{\Omega} \mathcal{A}_\nu dG^\omega(\omega) = \iint_{\Omega} (\alpha - r^b) p dG^\omega(\omega) \cdot \left( \frac{\iint_{\Omega} f \xi_\nu^f dG^\omega(\omega)}{\iint_{\Omega} f dG^\omega(\omega)} - \frac{\iint_{\Omega} p \xi_\nu^p dG^\omega(\omega)}{\iint_{\Omega} p dG^\omega(\omega)} \right). \quad (103)$$

If there is no stochastic lump sum, then Lemma 14 shows that  $p/f$  is equal for all individuals and that  $\xi_\nu^p = \xi_\nu^f$

for all individuals. It follows again that:

$$\iint_{\Omega} \mathcal{A}_{\nu} dG^{\omega}(\omega) = 0.$$

If there is a stochastic lump sum, then use Lemmas 14 and 15 to rewrite (103):

$$\begin{aligned} & \iint_{\Omega} \mathcal{A}_{\nu} dG^{\omega}(\omega) \\ &= \iint_{\Omega} (\alpha - r^b) p dG^{\omega}(\omega) \cdot \iint_{\Omega} \left( \frac{1 + \frac{\text{cov}(\tilde{K}, (1-t^e) f \tilde{r}^m)}{\text{var}((1-t^e) f \tilde{r}^m)}}{\iint_{\Omega} f dG^{\omega}(\omega)} - \frac{1 + \frac{\text{cov}(\tilde{K}, (1-t^e) f \tilde{r}^m)}{\text{var}((1-t^e) f \tilde{r}^m)}}{\iint_{\Omega} \left( 1 + \frac{\text{cov}(\tilde{K}, (1-t^e) f \tilde{r}^m)}{\text{var}((1-t^e) f \tilde{r}^m)} \right) f dG^{\omega}(\omega)} \right) f \xi_{\nu}^p dG^{\omega}(\omega) \\ &= \frac{\iint_{\Omega} (\alpha - r^b) p dG^{\omega}(\omega)}{\iint_{\Omega} f dG^{\omega}(\omega)} \cdot \frac{\iint_{\Omega} \left( 1 + \frac{\text{cov}(\tilde{K}, (1-t^e) f \tilde{r}^m)}{\text{var}((1-t^e) f \tilde{r}^m)} \right) f \xi_{\nu}^p dG^{\omega}(\omega) \cdot \iint_{\Omega} \frac{\text{cov}(\tilde{K}, (1-t^e) f \tilde{r}^m)}{\text{var}((1-t^e) f \tilde{r}^m)} f dG^{\omega}(\omega)}{\iint_{\Omega} \left( 1 + \frac{\text{cov}(\tilde{K}, (1-t^e) f \tilde{r}^m)}{\text{var}((1-t^e) f \tilde{r}^m)} \right) f dG^{\omega}(\omega)}. \end{aligned}$$

Since  $\text{cov}(\tilde{K}, \tilde{r}^m) > 0$ , we again find (assuming that  $\xi_{\nu}^p$  has the same sign for all individuals):

$$\iint_{\Omega} \mathcal{A}_{\nu} dG^{\omega}(\omega) \geq 0 \iff \xi_{\nu}^p \geq 0.$$