Optimal Taxation of Normal and Excess Returns to Risky Assets

Robin Boadway, Kevin Spiritus*

February 15, 2024

Abstract

We examine optimal linear taxes on normal and excess capital income with heterogeneous rates of return, alongside an optimal nonlinear earnings tax. Households optimize a portfolio containing three types of assets: risk-free, diversifiable risky, and private investment with idiosyncratic risk and heterogeneous expected returns. We define normal capital income as the risk-free rate times the size of the portfolio, and excess returns as any deviations from it. In this setting, taxing excess returns is ineffective for redistribution due to a Domar-Musgrave effect and only generates revenue, to be balanced against the cost of revenue uncertainty. Taxing normal returns does serve redistribution, as they reveal information about the investors' types beyond what the earnings tax base reveals.

Keywords: optimal capital taxation, risk, taxation of 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, Jo Van Biesebroeck, Bart Capéau, Annelies Casteleyn, Aart Gerritsen, Louis Kaplow, Etienne Lehmann and Erwin Ooghe for helpful comments and discussions, as well as participants to numerous seminars and conferences. We are grateful to the editor and two anonymous referees for their insightful and constructive remarks. 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. Declarations of interest: none.

1 Introduction

The current consensus among public economists is that capital income should be taxed at positive rates, despite suggestions in the earlier literature to the contrary.[1] This view has been strengthened by the observation that rates of return to capital differ among households.[2] Heterogeneous rates of return lead to above-normal—or excess—returns, which may reflect rents, luck or differences in investment ability. The taxation of rents serves both efficiency and redistributive objectives, and this has led some observers to recommend differential tax treatment of excess versus normal capital returns.

The Mirrlees Review (2011) proposed a Rate-of-Return Allowance (RRA) for the UK that would fully exempt risk-free returns from personal income taxation, while taxing excess returns at the same marginal rate as labour income. 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). Similarly, Cnossen and Jacobs (2022) have advocated the differential taxation of excess capital income for the Netherlands. In the corporate tax literature, the differential taxation of rents versus normal capital income has long been proposed, notably with the recommendation for an Allowance for Corporate Equity, which is equivalent to a tax on rents and risk (Gammie, 1991; Mirrlees Review, 2011). Recent proposals for cash-flow taxation accomplish the same result (e.g., President's Advisory Panel on Federal Tax Reform, 2005; Australian Treasury, 2010; Auerbach et al., 2017; Devereux et al., 2019).

Some Scandinavian countries already tax excess returns at a higher rate than normal returns. Norway, Finland and Sweden do so for non-incorporated businesses. Moreover, Norway uses a Rate-of-Return Allowance for corporations, whereas Finland and Sweden tax dividends above a basic exemption (see Cnossen and Sørensen, 2022). More generally, many

[1] The arguments are reviewed by, for example, Banks and Diamond (2010), Jacobs (2013), Bastani and Waldenström (2020) and Scheuer and Slemrod (2021).

[2] See, for example, Bach et al. (2020) and Fagereng et al. (2020) for empirical evidence on heterogeneous rates of return. Gahvari and Micheletto (2016), Kristjánsson (2016), Gerritsen et al. (2020) and Guvenen et al. (2023), among others, show how rate of return heterogeneity leads to positive optimal taxes on capital income.

countries shelter normal capital income from retirement savings using the so-called registered-asset treatment proposed by the Meade Report (Institute for Fiscal Studies, 1978). This entails savings in registered accounts being deductible from taxable income, while subsequent future withdrawals of principal plus accumulated capital income are taxable. As Adam and Miller (2021) note, the implication is that normal capital income is sheltered from tax while excess capital income is fully taxed as income. The rationale for taxing rents differentially is that this represents a source of tax revenue that is both efficient and equitable. Taxing normal returns, on the other hand, would impose efficiency losses, without offering distributional benefits.

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 and the government imposes an optimal nonlinear tax on labour income. We study a setting in which taxpayers choose how much they work and save, and optimize a portfolio of risk-free and risky assets. Here, normal returns are defined as the risk-free rate of return multiplied by the size of the portfolio, while excess returns are any deviations from these normal returns. Contrary to expectations, we find that linear taxation of excess returns neither reduces the expected after-tax excess returns nor the expected utility derived by investors, thus failing to redistribute the rents contained in the excess returns. The intuition follows from a generalization of the famous Domar and Musgrave (1944) result noted by Kaplow (1994) and Schindler (2008): a linear tax on excess returns with perfect loss offsets reduces the after-tax risk for investors, prompting them to increase their investments in risky assets. This increase in risky investment completely undoes the effect of the tax on after-tax income, leaving the expected utilities of the individuals unchanged.[3] Consequently, in the context of our model, the gov-

[3] Mossin (1968) and Stiglitz (1969) derived the Domar and Musgrave result in an expected utility framework for a representative agent allocating a fixed amount of wealth among risky and risk-free assets, where the risk-free asset has a zero-rate of return. In response to an increase in the capital income tax rate, the agent increases the demand for risky assets such that expected utility is kept constant. Kaplow (1994) and Schindler (2008) show that this effect applies when the return to the risk-free asset is positive as long as the capital income tax only applies to excess returns. We show that it continues to apply when the agents choose their labour supply and their amount of saving, and they compose a portfolio with multiple risky assets.

ernment is unable to use a linear tax on excess returns to redistribute between individuals with different expected rates of return.

At the same time, a positive tax on excess returns generates positive tax revenues from the increased risky investment it stimulates, including any additional rents contained therein. A drawback is that the additional tax revenues are risky. In the optimum, the government balances the benefit of additional tax revenues on excess returns against the increased uncertainty in the provision of public goods.

On the other hand, a tax on normal returns can contribute to redistributive objectives for two reasons. For one, the normal returns include information about the investment productivities of the individuals that is not already revealed by the labour income tax base. For another, if investment productivities are correlated with labour productivities, the normal returns also include information about the latter. Therefore, the tax on normal returns enhances the efficiency of redistribution. The government balances the distortions caused by a tax on normal returns against its 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 an optimal nonlinear labour income tax. We assume the government can observe normal and excess returns to capital, but it cannot observe separate components of excess returns. That is, it is practically impossible to tax rents, which may be desirable on efficiency grounds, without at the same time taxing returns to risk. To explore this issue, we allow for three types of assets: a risk-free asset yielding a normal rate of return which we assume is fixed, risky market assets that yield a competitive return and incur aggregate risk, and risky investments that yield idiosyncratic returns and possibly rents. We assume the latter assets are personal investments whose expected returns are linear and vary exogenously among individuals. Unlike market assets, personal investments cannot be insured in the capital market.

Individuals then differ in two unobservable dimensions, labour skills as in the standard optimal income tax literature and investment productivities. In our setting with imperfect information, 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. We assume that individuals invest strictly positive amounts in risk-free and risky assets.

Government revenues are stochastic because of the aggregate risk of market assets, and we follow Christiansen (1993) and Schindler (2008) in assuming that the government returns that risk to the taxpayers using a stochastic public good. Since the latter imposes a cost of risk on households, there is a social cost associated with taxing the excess returns, so 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.

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).

Main Results We first show that an analogue of the Domar-Musgrave (1944) result applies despite the inclusion of variable labour supply and saving, heterogeneous rates of return and different rates of tax on normal and excess returns. Individuals alter their portfolio compositions in response to an increase in the tax on excess returns such 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 supplies or their consumption in either period, so it does not affect their expected utilities. Therefore, the optimal tax on excess returns does not depend on the welfare weights of the individuals so does not fulfill a redistributive objective. However, although a tax on excess returns does not affect individual utilities, it does lead to an increase in expected revenues of the government by encouraging an increase in risky investments. The optimal tax rate on excess capital income balances the increase in expected tax revenues against the increase in riskiness of government revenues.

While the tax on excess returns cannot address redistributive concerns in our model, 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 optimal tax on risk-free capital income would be zero. 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 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. Gaillard and Wangner (2022) study the taxation of wealth in a macro-economic model, also in a context with type- and scale-dependent returns to capital. They show how the implications of taxing wealth depend on four statistics: the wealth Pareto tail, the degree of type- and scale-dependence, and the extent to which returns reflect investment productivity as opposed to rents. Scheuer and Slemrod (2021) also point out that differential individual returns to investment can be due to rents rather than productivity differences, in which case the argument for a capital income tax is strengthened. Schulz (2021) points out that if rates of return are scale-dependent, then taxing capital will depress rates of return, leading to a lower capital tax in the optimum. Two crucial differences between these papers and ours is that we allow for a separate tax on excess returns, which avoids distorting the intertemporal allocation of the taxpayers' consumption, and we allow for risk.

Regarding the taxation of risky capital income, the seminal contribution is that of Domar and Musgrave (1944), formalized using an expected utility approach by Mossin (1968) and Stiglitz (1969). They study a representative individual who maximizes expected utility by allocating a given amount of wealth over a safe asset and a risky asset with a constant expected return. They show that when the return on the risk-free asset is zero, a tax on capital income will induce the individual to increase investment in the risky asset such that private risk and expected utility remain unaltered. When the return on the risk-free asset is positive, a tax on all capital income will cause the individual to increase investment in the risky asset under the so-called Arrow assumptions about risk aversion, that is, absolute risk aversion is decreasing in wealth while relative risk aversion is increasing. However, expected utility will fall. Kaplow (1994) and Schindler (2008) show that when the return on the risk-free asset is positive, the analogue of the Domar-Musgrave effect will apply if the capital income tax applies only to excess returns. 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. 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.

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 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 a representative individual via changes in public goods. He does so in a model with fixed income and one risky asset with linear returns. In this case, stochastic public goods mitigate the consumption risk from holding risky assets. 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, while excess returns should be taxed, balancing public risk against private risk. We follow this latter route, assuming that the government returns part of the aggregate risk using a stochastic public good. Neither Christiansen (1993) nor Schindler (2008) consider redistributive taxation.

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 redistributional 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 a tax on excess returns should serve a redistributional purpose.

Road Map We proceeds as follows. Section 2 outlines the individual optimization problem, and finds relevant properties. We derive optimal linear tax rates on risk-free and excess returns to capital in Section 3. Section 4 concludes.

2 Individuals

Individuals are endowed with labour earning abilities, or skills, $w \in [\underline{w}, \overline{w}]$, which are distributed by the function $G^{\mathsf{w}}(w)$. The density function is $g^{\mathsf{w}}(w) = G^{\mathsf{w}}_w(w)$, where $G^{\mathsf{w}}_w(w) \equiv \mathrm{d}G^{\mathsf{w}}(w)/\mathrm{d}w$, a convention we adopt throughout the paper. 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 ℓ, yielding nonstochastic labour income $z \equiv w\ell$. Labour income is taxed according to the nonlinear tax function $t^{\ell}(z)$. Individuals consume c^1 of their first-period disposable income and save s, so:

$$c^1 = z - t^{\ell}(z) - s. \tag{1}$$

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

$$s = b + f + p. \tag{2}$$

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 $f\tilde{r}^m$. 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 with zero mean. The shock $\tilde{\varepsilon}$ is uncorrelated with the market returns. We follow the common assumption that the idiosyncratic shock $\tilde{\varepsilon}$ is added to the expected return α, which is independent of the size of investment for each individual, but differs across individuals. 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 ex ante in both their labour productivities and their expected rates of return to private investment $\alpha \in [\underline{\alpha}, \overline{\alpha}]$. We denote the two-dimensional types as $\boldsymbol{\theta} \equiv (w, \alpha)$, with domain $\Theta \equiv [\underline{w}, \overline{w}] \times [\underline{\alpha}, \overline{\alpha}]$. We denote the joint cumulative distribution of skills and expected returns to private investment by $G^{\theta}(\boldsymbol{\theta})$, with density function $g^{\theta}(\boldsymbol{\theta})$. We make no assumptions about the joint distributions of labour earning abilities w and returns to private investment α.

When choosing labour supply, savings and portfolio composition, the individuals know their type $\boldsymbol{\theta}$ and the distributions of the capital income shocks, but not the realizations of the shocks. Individuals who are ex ante equal differ ex post in the realizations of the idiosyncratic shock $\tilde{\varepsilon}$. Individuals of the same type $\boldsymbol{\theta}$ make the same decisions.

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:

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

Recall that we define *excess returns* to refer to all capital income that deviates from the risk-free return, including risk premiums, stochastic shocks and rents from private investment.

Individuals pay linear 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 savings plus second-period after-tax capital income:

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

The government chooses a nonlinear labour income tax function and linear 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. With 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:

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

The utility function displays weak separability between intertemporal consumption allocation and both labour effort and the public good, so preferences over c^1 and c^2 are independent of ℓ and P.[4]

[4]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

Individuals choose their labour supply, first period consumption, total savings and portfolio composition to maximize expected utility $\mathrm{E}[\tilde{U}]$ in (5), subject to their budget constraints (1)–(2). Second period consumption is determined as a residual by (4). The solution to this problem yields the value function $\mathrm{E}[\tilde{V}(t^\ell(\cdot), t^n, t^e, P(\cdot), \boldsymbol{\theta})]$, which is the expected maximized level of utility of an individual with type $\boldsymbol{\theta}$, given government policies.[5]

We assume throughout that second-order conditions for the individual optimization problem are met and focus on the first-order conditions. To simplify notation, we denote $\tilde{U}_i \equiv (\partial \tilde{U}/\partial u) \cdot (\partial u/\partial c^i)$, for $i = 1, 2$. The first-order conditions on earnings z and savings s are standard:

$$\frac{\mathrm{E}[\tilde{U}_\ell]}{\mathrm{E}[\tilde{U}_1]} = -(1 - t^\ell_z)w, \quad \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} = 1 + (1 - t^n)r^b. \tag{6}$$

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

$$(1 - t^e)\mathrm{E}[\tilde{r}^m - r^b] = -(1 - t^e)\frac{\mathrm{cov}(\tilde{U}_2, \tilde{r}^m)}{\mathrm{E}[\tilde{U}_2]}, \tag{7}$$

$$(1 - t^e)(\alpha - r^b) = -(1 - t^e)\frac{\mathrm{cov}(\tilde{U}_2, \tilde{\varepsilon})}{\mathrm{E}[\tilde{U}_2]}. \tag{8}$$

We assume an interior solution for the choice between risk-free and risky assets. Combining conditions (7) and (8) and using definition (3) yields:

$$(1 - t^e)\mathrm{E}[\tilde{y}^e] + (1 - t^e)\frac{\mathrm{cov}(\tilde{U}_2, \tilde{y}^e)}{\mathrm{E}[\tilde{U}_2]} = 0. \tag{9}$$

The term $\mathrm{E}[\tilde{y}^e]$ reflects the expected excess return to capital, which is positive (although some realizations of y^e are negative). The normalized covariance $\mathrm{cov}(\tilde{U}_2, \tilde{y}^e)/\mathrm{E}[\tilde{U}_2]$ is negative. The size of the second term in (9) reflects the required total risk premium from the risky investments. The left-hand side of (9) reflects the certainty equivalent of the after-tax excess capital incomes. Individuals thus invest in the risky assets up to the point where the certainty equivalent of their excess capital income equals zero. At the margin, they become indifferent

(2008) for a discussion of more general cases in a representative agent model. We verify the robustness of our results in our longer working paper (Boadway and Spiritus, 2021), assuming the government returns the risk of its revenues through a stochastic lump sum, potentially affecting the decisions of the individuals.

[5] When a function argument is a function itself rather than its value, we indicate this by adding $(\cdot)$ behind the function name.

between investing in risk-free and risky assets.

We denote the behavioural responses to policy perturbations by using subscripts. For example, the response of labour income to a marginal increase in the tax rate on normal capital income is z_{t^n}. We denote compensated responses using an asterisk, for example $z^*_{t^n}$. Following Christiansen (1984), Saez (2002) and Jacobs and Boadway (2014), we split the individual decision process into two stages. In the second stage, individuals take their labour supply as given, and they choose their expenditures on first-period consumption and savings, and their portfolio composition. In the first stage individuals choose their labour supply, anticipating the outcome of the second stage. When we study the effects of policy perturbations, we denote the effects on second-stage decisions conditional on labour income using a superscript c, for example $\tilde{y}^{ec}_{t^n}$.

We now formulate some properties of individual demand. Domar and Musgrave (1944), Mossin (1968), Stiglitz (1969) and Sandmo (1977) derive comparative statics for the effect of capital income tax changes on the demand for risky assets. In a representative agent setting where portfolio size is given and the return on the risk-free asset is zero, the agent responds to a change in the capital income tax by increasing the demand for the risky asset such that expected utility is kept constant. Kaplow (1994) and Schindler (2008) show that this Domar-Musgrave result also applies to a change in a tax on excess returns when the risk-free asset has a positive return. In the following lemma, we extend the findings of Kaplow and Schindler to our setting with heterogeneous agents with differential returns to private investment and variable savings and labour supply.

Lemma 1 (Generalized Domar-Musgrave effect)**.** *A reform to the tax on excess returns t^e has the following effects on outcomes for agents of a given type $\boldsymbol{\theta}$:*

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

$$\tilde{y}^e_{t^e} = \frac{\tilde{y}^e}{1 - t^e}. \tag{10}$$

2. *labour income, normal capital income and consumption in either period are not affected*

by a change in the tax on excess capital income:

$$z_{t^e} = y^n_{t^e} = c^1_{t^e} = \tilde{c}^2_{t^e} = 0. \tag{11}$$

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

$$\mathrm{E}[\tilde{V}_{t^e}] = 0. \tag{12}$$

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 the net-of-tax excess returns $(1-t^e)\tilde{y}^e$ decrease proportionally. This implies that the marginal utility of consumption $\tilde{U}_2$ becomes less responsive to the volatility of the excess returns. As a consequence, the net-of-tax marginal risk premiums required by the individuals (the right-hand sides of (7) and (8)) decline more than the effective decrease in the net-of-tax marginal risk premiums (the left-hand sides of (7) and (8)). 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.[6] □

An increase in the tax rate on excess returns has two effects on the tax revenues from an individual's excess returns, $t^e\tilde{y}^e$. First, there will be a mechanical increase in tax revenues, equal to $\tilde{y}^e$. Second, there will be increase in tax revenues due to the extension of investments in risky investments, equal to $t^e\tilde{y}^e_{t^e}$. Summing these two effects and using (10), we obtain the following corollary.

Corollary 1. *An increase in the tax rate on excess returns causes tax revenues to rise by*

[6]Bach et al. (2020) observe that risk aversions are heterogeneous. Lemma 1 remains valid in such a setting. We ignore this possibility because interpersonal welfare comparisons in such a setting are fraught with difficulties.

the full amount of the increase in excess capital income:

$$\frac{\mathrm{d}(t^e \tilde{y}^e)}{\mathrm{d}t^e} = \tilde{y}^e_{t^e}. \tag{13}$$

Intuitively, an increase in t^e induces an increase in excess capital income which causes tax revenues to increase by the same amount as the increase in excess capital income. As a consequence, the government bears all additional investment risk, and by (12) expected utility remains unchanged. The increase in excess capital income includes both an increase in rents from private investments and an increase in returns to risk. This result will be important in determining the optimal tax on excess capital income below.

Lemma 1 and Corollary 1 are robust to assumptions about individual preferences. A crucial assumption, however, is the linearity of the tax on excess returns on capital. 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. This would complicate our analysis, without adding much intuition. The second reason is that linear taxes on capital income might be collected by financial institutions, easing compliance.[7]

Note that Lemma 1 and Corollary 1 apply only to the tax on excess returns and not to the tax on risk-free returns. The effect of the latter on expected utility is not offset by a Domar-Musgrave effect. That would only be the case if, as in Mossin (1968) and Stiglitz (1969), portfolio size were fixed.

3 The government

3.1 The government's problem

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. The government cannot observe the allocation of savings among the different types of assets. One reason is that it is difficult

[7]Lemma 1 also crucially assumes that there is full loss-offsetting, that is symmetrical compensation for negative excess returns.

to distinguish between aggregate and idiosyncratic components of risk. Moreover, financial institutions might repackage bundles of assets, for fiscal or other reasons.

The government has access to the bond market and it balances its budget over time. The *law of large numbers* ensures that the government budget constraint is not affected by the idiosyncratic shocks. Aggregate shocks do 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:

$$P(\tilde{r}^m) = \iint_\Theta \left[(1+r^b)t^\ell(z) + t^e \mathrm{E}_\mathcal{E}[\tilde{y}^e|\tilde{r}^m] + t^n y^n\right] \mathrm{d}G^\theta(\boldsymbol{\theta}), \tag{14}$$

where $\mathrm{E}_\mathcal{E}$ denotes an average over the realizations of the private investment risk, conditional on the realization of the market rate of return.[8]

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:

$$\max_{t^\ell(\cdot),t^n,t^e,P(\cdot)} \iint_\Theta \mathrm{E}[\tilde{V}(t^\ell(\cdot), t^n, t^e, P(\cdot), \boldsymbol{\theta})]\mathrm{d}G^\theta(\boldsymbol{\theta}), \tag{15}$$

subject to the intertemporal budget constraint (14).

Note that we define social welfare in the first period, while we define the government's budget constraint in the second period. 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.

3.2 Optimal taxes

We now have all the building blocks to study the government problem of maximizing expected social welfare (15) subject to budget constraint (14), taking into account the be-

[8]The expectation operator $\mathrm{E}[\cdot]$ refers to the average of the operand over all potential realizations of the shocks (r^m, ε). The conditional expectation operator $\mathrm{E}_\mathcal{E}[\cdot|r^m]$ denotes an average over all potential realizations ε of private investment risk, conditional on the realization r^m of market risk. The quantity $\mathrm{E}_\mathcal{E}[\tilde{y}^e|\tilde{r}^m]$ contains aggregate risk, but no idiosyncratic risk.

We assume that the government 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.

havioural responses of the individuals. 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).

A difficulty in constructing the Lagrangian is that the government budget depends on the state of the world. 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 λ^{r^m}. To reflect this, we introduce a *stochastic budget Lagrange multiplier* $\tilde{\lambda}$. Each realization λ^{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.

We introduce some notation before discussing the optimality conditions. Suppose the government gives an additional unit of income in the first period to an individual of type $\boldsymbol{\theta} = (w, \alpha)$. The effect on social welfare, denoted $\beta(\boldsymbol{\theta})$, consists of two parts. The first is a direct effect on the expected utility of this individual. The second are income effects on the different tax bases, which affect the tax revenues. The monetary value of the effect on social welfare is:[9]

$$\beta(\boldsymbol{\theta}) \equiv \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{\lambda}]} + (1+r^b)t^\ell_z z_\rho + t^n y^n_\rho + t^e \mathrm{E}[\tilde{y}^e_\rho] + t^e \frac{\mathrm{cov}(\tilde{\lambda}, \tilde{y}^e_\rho)}{\mathrm{E}[\tilde{\lambda}]}, \tag{16}$$

where subscripts ρ indicate the effects of the additional income on the tax bases. The term $\beta(\boldsymbol{\theta})$ indicates the *net marginal social utility of income* for an individual of type $\boldsymbol{\theta}$, following Diamond (1975). The variation of $\beta(\boldsymbol{\theta})$ with skill captures the benefits of redistributing income between individuals with different earning abilities, whereas the variation of $\beta(\boldsymbol{\theta})$ with the expected return to private investment captures the benefits of redistributing income between individuals with different investment productivities. The social welfare weights include a social marginal risk premium $\mathrm{cov}(\tilde{\lambda}, \tilde{y}^e_\rho)/\mathrm{E}[\tilde{\lambda}]$, reflecting the welfare cost due to the uncertainty of the effects on government revenue from the tax on excess returns.

[9] We derive this term from the government's Lagrangian in Appendix A.4. We rewrite the income effects on tax revenue from excess capital income using the identity $\mathrm{E}[\tilde{\lambda}\tilde{y}^e_\rho] = \mathrm{E}[\tilde{\lambda}]\mathrm{E}[\tilde{y}^e_\rho] + \mathrm{cov}(\tilde{\lambda}, \tilde{y}^e_\rho)$.

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

$$\mathcal{W}(\boldsymbol{\theta}) \equiv -(1+r^b)t_z^\ell z_\sigma^* - t^n y_\sigma^{n*} - t^e \mathrm{E}\left[\tilde{y}_\sigma^{e*}\right] - t^e \frac{\operatorname{cov}(\tilde{\lambda}, \tilde{y}_\sigma^{e*})}{\mathrm{E}[\tilde{\lambda}]}, \tag{17}$$

where subscripts σ indicate the effects of an infinitesimal increase in the marginal tax on labour income. The 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-$\boldsymbol{\theta}$ 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, corrected for the uncertainty of the latter.

The first-order conditions for the labour income tax and the public good provision are standard, and are briefly discussed in Appendix A. In the following subsections, we focus on the optimal tax rates on the normal and excess capital incomes.

3.2.1 Optimal linear tax on excess capital income

The government chooses t^e to maximize social welfare (15) subject to the budget constraint (14). By Lemma 1, a change in t^e does not affect expected social welfare: it only affects the government budget constraint. Therefore, we can write the first-order condition for the tax on excess returns as (see Appendix A.6):

$$\iint_\Theta \left\{ \mathrm{E}[\tilde{\lambda}\tilde{y}^e] + t^e \mathrm{E}[\tilde{\lambda}\tilde{y}_{t^e}^e] \right\} dG^\theta(\boldsymbol{\theta}) = 0. \tag{18}$$

To interpret this condition, we make use of the first-order condition on the optimal provision of the public good (see Appendix A.7):

$$\lambda^{r^m} = \iint_\Theta \mathrm{E}_\mathcal{E}[\tilde{U}_P | r^m] dG^\theta(\boldsymbol{\theta}). \tag{19}$$

Using (18)–(19) and (10), we obtain the following theorem.

Theorem 1. *The optimal linear tax on excess capital income satisfies the following condition:*

$$\iint_\Theta \mathrm{E}\left[\tilde{y}_{t^e}^e\right] \mathrm{d}G^\theta(\boldsymbol{\theta}) + \iint_\Theta \frac{\operatorname{cov}\left(\iint_\Theta \mathrm{E}_\mathcal{E}[\tilde{U}_P|\tilde{r}^m] dG^\theta(\boldsymbol{\theta}), \tilde{y}_{t^e}^e\right)}{\mathrm{E}[\tilde{\lambda}]} \mathrm{d}G^\theta(\boldsymbol{\theta}) = 0. \tag{20}$$

The optimal tax on excess returns is not affected by the social welfare weights.

Eq. (20) is analogous to condition (9) determining risky investments by the household. The first term in (20) is the expected increase in tax revenue when the government increases t^e. By Corollary 1, this is the induced change in excess capital income. The second term, which is negative, is the government's risk premium associated with the uncertainty of the increase in excess capital income. At the optimum, the additional revenue just balances the increase in risk to the government. Note that Theorem 1 does not require that the government optimizes the tax on labour income.

One of the key differences between our paper and the literature is that we introduce heterogeneity in the expected excess returns. Intuitively, one might expect that this would lead to additional terms in (20), reflecting the distributional benefits of the tax on excess returns. However, optimality condition (20) does not depend on the social welfare weights β. 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, in our model, 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 (20). Recall from (15) 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 (9) the certainty equivalents of excess capital income are equal to zero for all optimizing individuals. Consequently, in our model, there is no scope for redistribution based on ex ante characteristics through the tax on excess returns.

However, the tax on excess returns still plays a supporting role in achieving redistributive objectives. As Lemma 1 shows, an increase in t^e induces agents to increase risky investments and therefore excess capital income. As we showed in Corollary 1, the increase in excess capital income is fully reflected as an increase in the tax revenue on excess capital income. In presence of private investment risk, the additional tax revenues on excess returns thus capture the increases in the rents from private investment that are caused by it, which implies that it constitutes a fully efficient source of tax revenue. Increases in t^e are limited by the fact that they give rise to riskiness in the provision of the public good.

We now show that t^e should be strictly between zero and unity. Suppose first that $t^e = 0$. The first term in (20) is positive since $\tilde{y}^e_{t^e} = \tilde{y}^e$ when $t^e = 0$ by (10). The second term 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 in (20) would be positive. Therefore, the only remaining possibility is a positive tax on excess returns $t^e > 0$.

To see that t^e should not be larger than 100%, consider first-order conditions (7)–(8) for the composition of the portfolios. If $t^e > 1$, then gains are turned into losses and losses are turned into gains. The only way for first-order conditions (7)–(8) to hold, is for individuals to undertake risky investments only if expected excess returns are negative, $\mathrm{E}[\tilde{y}^e] < 0$. Given our assumption that the expected excess rates of return to both risky assets are positive, this would imply that individuals would assume negative investments in both risky assets. In this case the first term in (20) remains positive because of (10). The covariance in (20) will also be positive, as higher market returns will lead to more negative revenues from excess returns, so lower public good provision and a larger marginal utility of the public good. Since both terms in (20) would be positive, the condition cannot hold and $t^e > 1$ cannot be optimal. In the border case where taxes on excess returns are fully taxed away, $t^e = 1$, it makes no difference for the individuals whether they invest in safe or risky assets (see first-order conditions (7)–(8)). In this case, our model is not well defined. If we make the realistic assumption that there is some net cost associated with investing in risky assets, it makes sense that individuals will only invest in safe assets when excess returns are fully taxed away, and we can realistically exclude the case where $t^e \geq 1$. We summarize our findings in Corollary 2 (see Appendix A.8).[10]

[10] In our longer working paper (Boadway and Spiritus, 2021), we study a case where contrary to the present paper there is no idiosyncratic risk and the government returns the aggregate risk of its revenues using a

Corollary 2. *The optimal linear tax rate on excess capital returns is strictly positive, and strictly smaller than 100%:*

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

If there is more aggregate risk, or if individuals are more risk averse with respect to the public good, then the optimal tax rate on excess returns is smaller.

To gain more insight into the optimal tax on the excess returns, let us briefly make some further assumptions. First, assume that the returns the risky assets are jointly normally distributed, so any linear combination of them has a univariate normal distribution. This assumption implies that individuals optimize their portfolios according to the Mean-Variance framework. Second, assume that all individuals have a constant Arrow-Pratt measure of global absolute risk aversion with respect to public good provision, so the following quantity is constant: $A \equiv -\mathrm{E}[\tilde{U}_{PP}]/\mathrm{E}[\tilde{U}_P]$. We prove the following Corollary in Appendix A.9.

Corollary 3. *Assume individuals optimize their portfolios according to the Mean-Variance framework, and they have a constant Arrow-Pratt measure of global absolute risk aversion A with respect to public good provision. Then the optimal linear tax rate on excess capital returns is:*

$$\frac{t^e}{1-t^e} = \frac{\iint_\Theta \mathrm{E}\left[\tilde{y}^e_{t^e}\right] \mathrm{d}G^\theta(\boldsymbol{\theta})}{A \mathrm{var}\left(\iint_\Theta \tilde{y}^e \mathrm{d}G^\theta(\boldsymbol{\theta})\right)}. \tag{21}$$

Corollary 3 clearly illustrates our findings in Theorem 1. On the one hand, the more risk averse individuals are with respect to the public good (larger A) or the more volatile is the tax base (larger $\mathrm{var}(\iint_\Theta \tilde{y}^e \mathrm{d}G^\theta(\boldsymbol{\theta}))$), the lower should be the tax on excess returns. On the other hand, the more the expected value of the tax base increases as the tax on excess returns increases (larger $\iint_\Theta \mathrm{E}[\tilde{y}^e_{t^e}]\mathrm{d}G^\theta(\boldsymbol{\theta})$), the higher should be the tax on excess returns.

---

uniform stochastic lump sum rather than a stochastic public good. We show that the tax on excess returns then becomes largely irrelevant. As soon as there is also uninsured idiosyncratic risk, the government should tax excess returns at high rates for insurance reasons.

3.2.2 Optimal linear tax on risk-free capital income

We derive in Appendix A the first-order condition for the tax on risk-free capital income. 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.[11] Let $G^{\mathrm{z}}(z)$ denote the cumulative distribution function for labour incomes in the tax optimum, and for any function $h(\boldsymbol{\theta})$, let $\overline{h}(z)$ denote the average of $h(\boldsymbol{\theta})$ for all types $\boldsymbol{\theta}$ who choose labour income z. The following theorem then characterizes the optimal tax on risk-free capital income.

Theorem 2. *Suppose labour income is optimally taxed, and the labour tax wedge $\mathcal{W}$ is not correlated with the propensity to save out of labour income:* $\mathrm{cov}_\Theta\left(\mathcal{W}, y^{nc}_z|z\right) = 0$. *The optimal linear tax on risk-free capital income then balances its distortions against its distributional benefits:*

$$\begin{aligned} t^n \int_{\underline{z}}^{\overline{z}} \overline{y^{nc*}_{t^n}}(z)\mathrm{d}G^{\mathrm{z}}(z) =& \frac{\mathrm{E}[\tilde{U}_2]}{\mathrm{E}[\tilde{U}_1]} \int_{\underline{z}}^{\overline{z}} \mathrm{cov}_\Theta(\beta, y^n|z)\mathrm{d}G^{\mathrm{z}}(z) \\ &- t^e \int_{\underline{z}}^{\overline{z}} \left(\mathrm{E}[\overline{\tilde{y}^{ec*}_{t^n}}(z)] - \chi^{c*}(z)\right)\mathrm{d}G^{\mathrm{z}}(z) \\ &- \int_{\underline{z}}^{\overline{z}} \overline{\mathcal{W}}(z)\left(\frac{\mathrm{d}\overline{y^n}}{\mathrm{d}z} - \frac{\overline{\partial y^{nc}}}{\partial z}\right)\frac{\mathrm{E}[\tilde{U}_2]}{\mathrm{E}[\tilde{U}_1]}\mathrm{d}G^{\mathrm{z}}(z), \end{aligned} \tag{22}$$

where $\chi(z)$ denotes the social risk premium associated with the uncertainty of tax revenues from excess capital income, for individuals with labour income z:

$$\chi(z) \equiv \frac{\mathrm{cov}(\tilde{\lambda}, \overline{\tilde{y}^e_{t^n}}|z)}{\mathrm{E}[\tilde{\lambda}]}.$$

Consider first the right-hand side of (22). The first term involves the population covariance between β and y^n conditional on labour income, and 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 distributional characteristic of risk-free capital income is zero, $\mathrm{cov}_\Theta(\beta, y^n|z) = 0$, and

[11] This corresponds to Assumption 2 of Saez (2002, p. 225), who argues there are no reasons to think that such correlations exist, and that since it is difficult to empirically check this condition it seems reasonable to assume that it holds.

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 $\text{cov}_\Theta(\beta, y^n|z) \neq 0$, 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 $\text{cov}_\Theta(\beta, y^n|z) \neq 0$. 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 α 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 $\text{cov}_\Theta(\beta, y^n|z)$ 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 (22) then puts a downward pressure on the optimal tax on risk-free capital income.

Lemma 2. *If preferences are additively separable between consumption in the two periods and absolute risk aversion is constant, then the distributional characteristic of risk-free capital income conditional on labour income is positive:*

$$\text{cov}_\Theta(\beta, y^n|z) > 0.$$

With more general preferences, one cannot unambiguously sign the covariance $\text{cov}_\Theta(\beta, y^n|z)$ 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. The sign and the size of $\text{cov}_\Theta(\beta, y^n|z)$ thus remain unclear.

The second term on the right-hand side of (22) concerns the compensated effect of an increase in t^n on tax revenue from excess capital income, conditional on labour income. The government needs to account not only for the expected effect of the tax increase on excess capital income $\mathrm{E}[\overline{\tilde{y}^{ec*}_{t^n}}]$, but also for the welfare effect of the uncertainty therein through the risk premium χ. The sign of the term χ^{c*} is opposite to that of $\mathrm{E}[\overline{\tilde{y}^{ec*}_{t^n}}]$, since λ rises as r^m falls and vice versa.[12] It is not clear a priori whether the total impact of the second term in (22) on social welfare is positive or negative. The following lemma, which we prove in Appendix A.10, shows how the second term on the right-hand side of (22) reflects the change in the relative importance of idiosyncratic and aggregate risk due to a perturbation of t^n.

Lemma 3. *The effect of a perturbation of 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:*

$$t^e \int_{\underline{z}}^{\overline{z}} \left(\mathrm{E}[\overline{\tilde{y}^{ec*}_{t^n}}(z)] - \chi^{c*}(z)\right) \mathrm{d}G^z(z) = \mathcal{T}^p \cdot \left(\frac{\mathcal{T}^{pc*}_{t^n}}{\mathcal{T}^p} - \frac{\mathcal{T}^{fc*}_{t^n}}{\mathcal{T}^f}\right),$$

where $\mathcal{T}^p \equiv t^e \iint_\Theta (\alpha - r^b) p \mathrm{d}G^\theta(\boldsymbol{\theta})$ and $\mathcal{T}^f \equiv t^e \iint_\Theta \mathrm{E}[\tilde{r}^m - r^b] f \mathrm{d}G^\theta(\boldsymbol{\theta})$ denote the respective tax revenues from excess returns from private and market investments.

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 more diversified. We show in Appendix B that various simplifications of our model lead to the conclusion that the second term on the right-hand side of (22) 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 the second term in (22) equals zero when the elasticities of p and f with respect to t^n are constant over the income distribution. In those cases, the expected revenue effects of a reform of t^n, which are typically used as sufficient statistic in optimal tax

[12] Use (3) to find: $\chi = \mathrm{cov}(\tilde{\lambda}, \overline{\tilde{y}^e_{t^n}}|z)/\mathrm{E}[\tilde{\lambda}] = \overline{f}_{t^n}\mathrm{cov}(\tilde{\lambda}, \tilde{r}^m)/\mathrm{E}[\tilde{\lambda}]$.

studies, are cancelled out by the welfare cost of their uncertainty. 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. Even if the second term on the right-hand side of (22) becomes zero under common theoretical assumptions, its sign is ultimately an unanswered empirical question.

The third term on the right-hand side of (22) concerns the difference between the cross-sectional variation and the individual variation of normal capital income. Since we assume that individuals have identical preferences that are separable between leisure and consumption, there is only one reason why this term may differ from zero: the expected rates of return to private investment could be correlated with the labour productivities. Suppose that, as argued above, individuals with higher expected returns to private investment save less, conditional in labour income, and suppose that expected returns to private investment are positively correlated with labour productivity. Then conditional on labour income, lower savings indicate higher labour productivities. It is then optimal to implement a lower tax on normal capital income, because doing so would alleviate the distortions caused by the tax on labour income.

The integral on the left-hand side of (22) is negative: a compensated increase of the tax on risk-free capital income decreases savings and thus risk-free returns, conditional on labour income. Furthermore, we showed in the previous subsection that the tax rate t^e on excess returns is never negative. If the right-hand side of (22) is positive, then Theorem 2 indicates that the optimal tax on risk-free returns to capital income is negative.[13]

[13] The optimality condition in Theorem 2 is similar to the condition for optimal linear commodity taxation of Spiritus (2024). He finds that the optimal tax on a commodity is determined by the marginal excess burden caused by that tax, the distributional characteristic of the commodity conditional on labour income, and the difference between the individual and cross-sectional variations of consumption of the commodity. The main difference between Spiritus (2024) and Theorem 2 is that the latter contains additional terms to account for the welfare cost of the riskiness of the tax revenues.

4 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 on 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, in our stylized model, a Domar-Musgrave effect nullifies such a role: individuals respond to a tax on excess returns by adjusting their portfolios, taking on more risk such that their expected utility remains unchanged. As the size of the risk premium before taxation increases, tax revenue from excess returns also increases. Part of that tax revenue results from higher rents generated from increased private investments. These are fully taxed away by the government and represent a fully efficient source of tax revenue. Part of the tax revenue represents additional risky revenues resulting from increased investment in market assets. The revenue-generating role of the tax on excess returns is balanced against the welfare cost of the uncertainty of those revenues. The higher is the taxpayers' risk aversion with respect to the public good and the more volatile are the revenues from the tax on excess returns, the lower the tax on excess returns should be.

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 labour 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 and conditional on labour income, 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.

Our stylized model has certain limitations. It assumes that taxes on capital income are linear, and it ignores any incentive effects that taxing excess returns to capital might incur. Additionally, it does not consider scale effects or general equilibrium effects, and it simplifies the life-cycle to just two periods. Our objective is not to create a comprehensive model, but rather to establish a benchmark demonstrating that optimal tax results are sensitive to the inclusion of risky assets. The generalized Domar-Musgrave (1944) effect means that results from a model with deterministic capital incomes cannot be directly extrapolated to a context with risky capital incomes. One should be particularly careful when applying a sufficient-statistics framework. If tax revenues are uncertain, the change in expected tax revenues due to a policy reform is no longer a sufficient measure for its effect on 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.

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.

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.

Australian Treasury. Australia's future tax system. *British Tax Review*, (4):307–310, 2010.

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

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.

Robin Boadway and Kevin Spiritus. Optimal taxation of normal and excess returns to risky assets. 2021.

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.

Sijbren Cnossen and Bas Jacobs, editors. *Tax by Design for the Netherlands*. Oxford University Press, 2022.

Sijbren Cnossen and Peter Birch Sørensen. Towards a true dual income tax. In Sijbren Cnossen and Bas Jacobs, editors, *Tax by design for the Netherlands*, chapter Chapter 3, pages 46–65. Oxford University Press, 2022.

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.

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

Alexandre Gaillard and Philipp Wangner. Wealth, returns, and taxation: A tale of two dependencies. *Available at SSRN 3966130*, 2022.

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. *CESifo 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 and redistributional effects of wealth taxation. *The Quarterly Journal of Economics*, 138(2):835–894, 2023.

Institute for Fiscal Studies. *The Structure and Reform of Direct Taxation: Report of a Committee Chaired by Professor J. E. Meade.* London: Allen & Unwin, 1978. that is hereafter referred to as “the Meade Report”.

Bas Jacobs. From optimal tax theory to applied tax policy. *FinanzArchiv*, 69(3):338–389, 2013.

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.

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.

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.

President's Advisory Panel on Federal Tax Reform. Simple, fair, and pro-growth: Proposals to fix america's tax system. 2005.

Mark Rubinstein. The valuation of uncertain income streams and the pricing of options. *The Bell Journal of Economics*, pages 407–425, 1976.

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.

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.

Karl Schulz. Redistribution of return inequality. Technical report, CESifo Working Paper, 2021.

Kevin Spiritus. Optimal commodity taxation when households earn multiple incomes. *International Tax and Public Finance*, 2024. Forthcoming.

Kevin Spiritus, Etienne Lehmann, Sander Renes, and Floris Zoutman. Optimal taxation with multiple incomes and types. *CEPR Discussion Paper*, (DP16797), 2022.

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

The government uses two linear instruments (t^n and t^e) and two nonlinear instruments ($t^\ell(z)$ 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) and Spiritus et al. (2022) 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 ϵ^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 ϵ^z and η^z allow modelling any small perturbation to the tax function $t^\ell(z)$. If the value of ϵ^z equals zero, then the unperturbed tax function is in place. If the optimal value of ϵ^z is zero for every function η^z, then we know that the unperturbed tax function is optimal.[14] Similarly, we introduce perturbations $\epsilon^P\eta^P(\tilde{r}^m)$ for the state-dependent public good.

To verify the optimality of the government policies, we must verify that *any* perturbation leaves the government objective unchanged. 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 simpler perturbation parameters. For this purpose, we parameterize the labour tax function after the perturbations as $t^\ell(z)+\epsilon^z\eta^z(z)+\sigma z-\rho$, where the parameter ρ shifts the intercept and σ shifts the slope of the tax function. The corresponding marginal tax rate is $t^\ell_z(z) + \epsilon^z\eta^z_z(z) + \sigma$.[15] Furthermore, we will show that the general perturbations 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 $P(r^m)$. We now model these local perturbations more formally using a Dirac delta function $\delta(\tilde{r}^m - r^m)$, and we let π^m parameterize the size of the reforms.[16] The provision of the stochastic public good after the perturbations then 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:

$$\mathcal{C}^1(z, s, \epsilon^z, \sigma, \rho) \equiv z - s - t^\ell(z) - \sigma z - \epsilon^z\eta^z(z) + \rho, \tag{23}$$

$$\tilde{\mathcal{C}}^2(s, f, p, t^n, t^e, \rho_2) \equiv s + (1 - t^n)r^b s + (1 - t^e)\tilde{y}^e(f, p) + \rho_2, \tag{24}$$

where $\tilde{y}^e(f, p)$ is defined by (3), and ρ_2 denotes a deterministic lump sum that is given to all individuals in the second period.

[14] Saez (2001) considers one specific perturbation, where the function η^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.

[15] For previous papers that parameterize nonlinear tax schedules to study comparative statics, see, e.g., Christiansen (1981) and Jacquet et al. (2013).

[16] 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: $\int_{-\infty}^{\infty} \delta(r^m - R^m)\mathrm{d}r^m = 1$. This setup allows modelling a localized reform to the lump sum with non-zero effects on social welfare.

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 (23) and (24) for c^1 and c^2. An individual facing the perturbed government policies thus maximizes the expected value of the following utility function:

$$\begin{aligned}\tilde{\mathcal{U}}(z,&s,f,p,\sigma,\epsilon^z,\rho,t^n,t^e,\epsilon^P,\pi^m,\rho_2)\\ &\equiv U\left(u(\mathcal{C}^1(z,s,\epsilon^z,\sigma,\rho),\tilde{\mathcal{C}}^2(s,f,p,t^n,t^e,\rho_2)),\frac{z}{w},P(\tilde{r}^m,\epsilon^P,\pi^m,r^m)\right),\end{aligned} \tag{25}$$

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

The first-order conditions for the individual are:

$$\mathrm{E}[\tilde{\mathcal{U}}_z]=\mathrm{E}[\tilde{\mathcal{U}}_s]=\mathrm{E}[\tilde{\mathcal{U}}_f]=\mathrm{E}[\tilde{\mathcal{U}}_p]=0. \tag{26}$$

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 $\boldsymbol{\theta}$ and the perturbation parameters σ, ϵ^z, ρ, t^n, t^e, ϵ^P, π^m and ρ_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 (25) yields the following properties:

$$\mathrm{E}[\tilde{\mathcal{V}}_{\epsilon^z}]=-\mathrm{E}[\tilde{U}_1]\eta^z(z),\quad \mathrm{E}[\tilde{\mathcal{V}}_\sigma]=-\mathrm{E}[\tilde{U}_1]z, \tag{27}$$

$$\mathrm{E}[\tilde{\mathcal{V}}_{t^e}]=0,\quad \mathrm{E}[\tilde{\mathcal{V}}_{t^n}]=-\mathrm{E}[\tilde{U}_2]y^n,\quad \mathrm{E}[\tilde{\mathcal{V}}_\rho]=\mathrm{E}[\tilde{U}_1],\quad \mathrm{E}[\tilde{\mathcal{V}}_{\rho_2}]=\mathrm{E}[\tilde{U}_2], \tag{28}$$

$$\mathrm{E}[\tilde{\mathcal{V}}_{\epsilon^P}]=\mathrm{E}[\tilde{U}_P\eta^P(\tilde{r}^m)]. \tag{29}$$

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

Lemma 4 (Slutsky properties)**.** *We find the following Slutsky properties for individual behaviour:*

- *for any $k=z,s,f,p$, the behavioural effects of perturbations σ and t^n can be decomposed into income effects and compensated effects:*

$$k^*_\sigma=k_\sigma+zk_\rho,\quad k^*_{t^n}=k_{t^n}+y^nk_{\rho_2}. \tag{30}$$

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

$$z^*_{t^n}=\frac{\mathrm{E}[\tilde{U}_2]}{\mathrm{E}[\tilde{U}_1]}y^{n*}_\sigma. \tag{31}$$

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:

$$\mathcal{V}^*(\boldsymbol{\theta}, \sigma, t^n) \equiv \mathcal{V}\left(\boldsymbol{\theta}, \sigma, t^n, \rho = z^i\sigma, \rho_2 = y^{ni}t^n\right),$$

where we omit irrelevant function arguments. Use envelope conditions (27)–(29) to find:

$$\mathrm{E}[\tilde{\mathcal{V}}^*_\sigma] = \mathrm{E}[\tilde{\mathcal{V}}_\sigma] + \mathrm{E}[\tilde{\mathcal{V}}_\rho]z^i = -\mathrm{E}[\tilde{U}_1](z - z^i), \tag{32}$$

$$\mathrm{E}[\tilde{\mathcal{V}}^*_{t^n}] = \mathrm{E}[\tilde{\mathcal{V}}_{t^n}] + \mathrm{E}[\tilde{\mathcal{V}}_{\rho_2}]y^{ni} = -\mathrm{E}[\tilde{U}_2](y^n - y^{ni}), \tag{33}$$

1. In the situation without any reforms, (32)–(33) show that $\mathrm{E}[\tilde{\mathcal{V}}^*_\sigma] = \mathrm{E}[\tilde{\mathcal{V}}^*_{t^n}] = 0$. It follows that the combined reforms in (30) are indeed compensated.

2. Evaluate the partial derivative of (32) with respect to t^n, and that of (33) with respect to σ, each time in the situation before any reforms:

$$\mathrm{E}[\tilde{\mathcal{V}}^*_{\sigma t^n}] = -\mathrm{E}[\tilde{U}_1]z^*_{t^n}, \tag{34}$$

$$\mathrm{E}[\tilde{\mathcal{V}}^*_{t^n\sigma}] = -\mathrm{E}[\tilde{U}_2]y^{n*}_\sigma. \tag{35}$$

   Young's theorem demands that the second-order partial derivatives of any function are symmetric. Apply this requirement to (34) and (35) to find Slutsky symmetry (31).

□

If there is a marginal change to any of the type or perturbation parameters ν, then individuals change their behaviour, such that first-order conditions (26) remain satisfied. We thus obtain for any $k = z, s, f, p$:[17]

$$0 = \frac{\mathrm{dE}[\tilde{\mathcal{U}}_k]}{\mathrm{d}\nu} = \mathrm{E}[\tilde{\mathcal{U}}_{kz}]z_\nu + \mathrm{E}[\tilde{\mathcal{U}}_{ks}]s_\nu + \mathrm{E}[\tilde{\mathcal{U}}_{kf}]f_\nu + \mathrm{E}[\tilde{\mathcal{U}}_{kp}]p_\nu + \mathrm{E}[\tilde{\mathcal{U}}_{k\nu}]. \tag{36}$$

Write equation (36) in matrix notation to find the following lemma.

Lemma 5. *The effects of a change of any parameter* $\nu = w$, σ, ϵ^z, ρ, t^n, t^e, ϵ^P, π^m *or* ρ_2

[17] The behavioural effects z_ν, s_ν, f_ν and p_ν are total effects. This means that they include not only the direct effects of a change in the parameter ν. They also include second-round effects due to the nonlinearity of the tax on labour income. If a change in the parameter ν causes a change in labour income z, this will cause a change in the marginal tax on labour income $t^\ell_z(z)$, 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 $\mathrm{E}[\tilde{\mathcal{U}}_z]$, $\mathrm{E}[\tilde{\mathcal{U}}_s]$, $\mathrm{E}[\tilde{\mathcal{U}}_f]$ and $\mathrm{E}[\tilde{\mathcal{U}}_p]$ are equivalent to the *shift functions* introduced by Jacquet et al. (2013).

on the choices of the individuals are given by:

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

where $\tilde{\mathcal{H}}$ denotes the Hessian of utility function (25).

Lemma 5 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 5 to find relations between the income effects in both periods.

Lemma 6. *The behavioural effects of perturbations to the tax intercepts in the two periods are related as follows:*

$$z_\rho = \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} z_{\rho_2}, \quad s_\rho = \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} s_{\rho_2} + 1, \quad f_\rho = \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} f_{\rho_2}, \quad p_\rho = \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} p_{\rho_2}. \tag{37}$$

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

$$\mathrm{E}[\tilde{\mathcal{U}}_{z\rho}] = (1-t_z^\ell)\mathrm{E}[\tilde{U}_{11}], \qquad \mathrm{E}[\tilde{\mathcal{U}}_{z\rho_2}] = -(1-t_z^\ell)\mathrm{E}[\tilde{U}_{12}], \tag{38}$$

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

$$\forall k = f, p : \mathrm{E}[\tilde{\mathcal{U}}_{k\rho}] = (1-t^e)\mathrm{E}[\tilde{U}_{21}\tilde{y}_k^e], \qquad \mathrm{E}[\tilde{\mathcal{U}}_{k\rho_2}] = (1-t^e)\mathrm{E}[\tilde{U}_{22}\tilde{y}_k^e], \tag{39}$$

and thus:

$$\forall k = z, s, f, p : \mathrm{E}[\tilde{\mathcal{U}}_{ks}] = \mathrm{E}[\tilde{\mathcal{U}}_{k\rho_2}]\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} - \mathrm{E}[\tilde{\mathcal{U}}_{k\rho}]. \tag{40}$$

Write (40) in matrix notation:

$$\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}\mathrm{E}\left[\begin{pmatrix} \tilde{\mathcal{U}}_{z\rho_2} \\ \tilde{\mathcal{U}}_{s\rho_2} \\ \tilde{\mathcal{U}}_{f\rho_2} \\ \tilde{\mathcal{U}}_{p\rho_2} \end{pmatrix}\right] - \mathrm{E}\left[\begin{pmatrix} \tilde{\mathcal{U}}_{z\rho} \\ \tilde{\mathcal{U}}_{s\rho} \\ \tilde{\mathcal{U}}_{f\rho} \\ \tilde{\mathcal{U}}_{p\rho} \end{pmatrix}\right] = \mathrm{E}\left[\begin{pmatrix} \tilde{\mathcal{U}}_{zs} \\ \tilde{\mathcal{U}}_{ss} \\ \tilde{\mathcal{U}}_{fs} \\ \tilde{\mathcal{U}}_{ps} \end{pmatrix}\right].$$

Substitute Lemma 5 on the left-hand side and fully write the Hessian $\mathrm{E}[\widetilde{\mathcal{H}}]$:

$$-\mathrm{E}\left[\begin{pmatrix}\tilde{\mathcal{U}}_{zz} & \tilde{\mathcal{U}}_{zs} & \tilde{\mathcal{U}}_{zf} & \tilde{\mathcal{U}}_{zp}\\ \tilde{\mathcal{U}}_{sz} & \tilde{\mathcal{U}}_{ss} & \tilde{\mathcal{U}}_{sf} & \tilde{\mathcal{U}}_{sp}\\ \tilde{\mathcal{U}}_{fz} & \tilde{\mathcal{U}}_{fs} & \tilde{\mathcal{U}}_{ff} & \tilde{\mathcal{U}}_{fp}\\ \tilde{\mathcal{U}}_{pz} & \tilde{\mathcal{U}}_{ps} & \tilde{\mathcal{U}}_{pf} & \tilde{\mathcal{U}}_{pp}\end{pmatrix}\right]\cdot\mathrm{E}\left[\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}\begin{pmatrix}z_{\rho 2}\\ s_{\rho 2}\\ f_{\rho 2}\\ p_{\rho 2}\end{pmatrix}-\begin{pmatrix}z_\rho\\ s_\rho\\ f_\rho\\ p_\rho\end{pmatrix}\right]=\mathrm{E}\left[\begin{pmatrix}\tilde{\mathcal{U}}_{zs}\\ \tilde{\mathcal{U}}_{ss}\\ \tilde{\mathcal{U}}_{fs}\\ \tilde{\mathcal{U}}_{ps}\end{pmatrix}\right],$$

or equivalently:

$$-\mathrm{E}\left[\begin{pmatrix}\tilde{\mathcal{U}}_{zz} & \tilde{\mathcal{U}}_{zs} & \tilde{\mathcal{U}}_{zf} & \tilde{\mathcal{U}}_{zp}\\ \tilde{\mathcal{U}}_{sz} & \tilde{\mathcal{U}}_{ss} & \tilde{\mathcal{U}}_{sf} & \tilde{\mathcal{U}}_{sp}\\ \tilde{\mathcal{U}}_{fz} & \tilde{\mathcal{U}}_{fs} & \tilde{\mathcal{U}}_{ff} & \tilde{\mathcal{U}}_{fp}\\ \tilde{\mathcal{U}}_{pz} & \tilde{\mathcal{U}}_{ps} & \tilde{\mathcal{U}}_{pf} & \tilde{\mathcal{U}}_{pp}\end{pmatrix}\right]\cdot\mathrm{E}\left[\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}\begin{pmatrix}z_{\rho 2}\\ s_{\rho 2}\\ f_{\rho 2}\\ p_{\rho 2}\end{pmatrix}-\begin{pmatrix}z_\rho\\ s_\rho\\ f_\rho\\ p_\rho\end{pmatrix}+\begin{pmatrix}0\\ 1\\ 0\\ 0\end{pmatrix}\right]=0.$$

This last equation implies Eq. (37). □

When there is a marginal increase in the perturbation parameter ε^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 7. *The behavioural effects of any perturbation function $\epsilon^z\eta^z$ can be decomposed into substitution and income effects:*

$$k_{\epsilon^z} = k^*_\sigma \eta^z_z - k_\rho \eta^z. \tag{41}$$

Proof. Note the following second-order derivatives of (25), evaluated in the situation before any reforms, and substitute Eqs. (38) and (39):

$$\begin{aligned}\mathrm{E}[\tilde{\mathcal{U}}_{z\sigma}] &= -\mathrm{E}[\tilde{U}_1] - \mathrm{E}[\tilde{\mathcal{U}}_{z\rho}]z,\\ \forall k = s, f, p : \mathrm{E}[\tilde{\mathcal{U}}_{k\sigma}] &= -\mathrm{E}[\tilde{\mathcal{U}}_{k\rho}]z,\\ \forall k = z, s, f, p : \mathrm{E}[\tilde{\mathcal{U}}_{k\epsilon^z}] &= (\mathrm{E}[\tilde{\mathcal{U}}_{k\sigma}] + z\mathrm{E}[\tilde{\mathcal{U}}_{k\rho}])\eta^z_z - \mathrm{E}[\tilde{\mathcal{U}}_{k\rho}]\eta^z.\end{aligned} \tag{42}$$

Substitute (42) into Lemma 5 and use Slutsky decomposition (30) to find Eq. (41). □

In the following Lemma, we apply Lemma 5 to find the effects of a change in the tax on excess returns t^e on the behaviour of the individuals.

Lemma 8. *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:

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

Proof. Note the following second-order derivatives, evaluated in the situation before any reforms, and substitute $\mathrm{E}[\tilde{\mathcal{U}}_{fs}]$ and $\mathrm{E}[\tilde{\mathcal{U}}_{ps}]$ from Eq. (40):

$$\mathrm{E}[\tilde{\mathcal{U}}_{fz}] = \frac{1-t^e}{f}\frac{1}{w}\mathrm{E}\left[\left(\tilde{U}_{2\ell} - \tilde{U}_{21}\frac{\mathrm{E}[\tilde{U}_\ell]}{\mathrm{E}[\tilde{U}_1]}\right)(\tilde{r}^m - r^b)f\right], \tag{43}$$

$$\mathrm{E}[\tilde{\mathcal{U}}_{ff}] = \frac{1-t^e}{f}\mathrm{E}[(1-t^e)\tilde{U}_{22}(\tilde{r}^m - r^b)^2 f], \tag{44}$$

$$\mathrm{E}[\tilde{\mathcal{U}}_{fp}] = \frac{1-t^e}{f}\frac{1-t^e}{p}\mathrm{E}[\tilde{U}_{22}(\tilde{r}^m - r^b)f(\alpha+\tilde{\varepsilon}-r^b)p], \tag{45}$$

$$\mathrm{E}[\tilde{\mathcal{U}}_{pz}] = \frac{1-t^e}{p}\frac{1}{w}\mathrm{E}\left[\left(\tilde{U}_{2\ell} - \tilde{U}_{21}\frac{\mathrm{E}[\tilde{U}_\ell]}{\mathrm{E}[\tilde{U}_1]}\right)(\alpha+\tilde{\varepsilon}-r^b)p\right], \tag{46}$$

$$\mathrm{E}[\tilde{\mathcal{U}}_{pp}] = \frac{1-t^e}{p}\mathrm{E}\left[(1-t^e)\tilde{U}_{22}(\alpha+\tilde{\varepsilon}-r^b)^2 p\right], \tag{47}$$

$$\forall k = z,s,f,p : \mathrm{E}[\tilde{\mathcal{U}}_{t^e k}] = -\frac{f}{1-t^e}\mathrm{E}[\tilde{\mathcal{U}}_{kf}] - \frac{p}{1-t^e}\mathrm{E}[\tilde{\mathcal{U}}_{kp}]. \tag{48}$$

Apply Cramer's rule to Lemma 5 and substitute $\mathrm{E}[\tilde{\mathcal{U}}_{t^e z}]$ and $\mathrm{E}[\tilde{\mathcal{U}}_{t^e s}]$ from Eq. (48) to find the effects on labour income and savings:

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

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

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 $\mathrm{E}[\tilde{\mathcal{U}}_{t^e f}]$

and $\mathrm{E}[\tilde{\mathcal{U}}_{t^e p}]$ from (48):

$$f_{t^e} = \frac{1}{\det\left(\mathrm{E}[\tilde{\mathcal{H}}]\right)} \det\left(\mathrm{E}\left[\begin{pmatrix} \tilde{\mathcal{U}}_{zz} & \tilde{\mathcal{U}}_{zs} & \frac{f\tilde{\mathcal{U}}_{zf}+p\tilde{\mathcal{U}}_{zp}}{1-t^e} & \tilde{\mathcal{U}}_{zp} \\ \tilde{\mathcal{U}}_{sz} & \tilde{\mathcal{U}}_{ss} & \frac{f\tilde{\mathcal{U}}_{sf}+p\tilde{\mathcal{U}}_{sp}}{1-t^e} & \tilde{\mathcal{U}}_{sp} \\ \tilde{\mathcal{U}}_{fz} & \tilde{\mathcal{U}}_{fs} & \frac{f\tilde{\mathcal{U}}_{ff}+p\tilde{\mathcal{U}}_{fp}}{1-t^e} & \tilde{\mathcal{U}}_{fp} \\ \tilde{\mathcal{U}}_{pz} & \tilde{\mathcal{U}}_{ps} & \frac{f\tilde{\mathcal{U}}_{pf}+p\tilde{\mathcal{U}}_{pp}}{1-t^e} & \tilde{\mathcal{U}}_{pp} \end{pmatrix}\right]\right) = \frac{f}{1-t^e}, \tag{49}$$

$$p_{t^e} = \frac{1}{\det\left(\mathrm{E}[\tilde{\mathcal{H}}]\right)} \det\left(\mathrm{E}\left[\begin{pmatrix} \tilde{\mathcal{U}}_{zz} & \tilde{\mathcal{U}}_{zs} & \tilde{\mathcal{U}}_{zf} & \frac{f\tilde{\mathcal{U}}_{zf}+p\tilde{\mathcal{U}}_{zp}}{1-t^e} \\ \tilde{\mathcal{U}}_{sz} & \tilde{\mathcal{U}}_{ss} & \tilde{\mathcal{U}}_{sf} & \frac{f\tilde{\mathcal{U}}_{sf}+p\tilde{\mathcal{U}}_{sp}}{1-t^e} \\ \tilde{\mathcal{U}}_{fz} & \tilde{\mathcal{U}}_{fs} & \tilde{\mathcal{U}}_{ff} & \frac{f\tilde{\mathcal{U}}_{ff}+p\tilde{\mathcal{U}}_{fp}}{1-t^e} \\ \tilde{\mathcal{U}}_{pz} & \tilde{\mathcal{U}}_{ps} & \tilde{\mathcal{U}}_{pf} & \frac{f\tilde{\mathcal{U}}_{pf}+p\tilde{\mathcal{U}}_{pp}}{1-t^e} \end{pmatrix}\right]\right) = \frac{p}{1-t^e}. \tag{50}$$

□

We now use Lemma 8 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 8. Substituting f_{t^e} and p_{t^e} into the uncompensated function (3) for $\tilde{y}^e$, yields Eq. (10).

2. We found that $z_{t^e} = s_{t^e} = 0$ in Lemma 8. Substituting the comparative statics for z_{t^e}, s_{t^e} and $\tilde{y}^e_{t^e}$ into budget constraints (1) and (4) yields $\tilde{c}^2_{t^e} = 0$.

3. We already showed that $\mathrm{E}[\tilde{V}_{t^e}] = 0$ in (28).

□

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 α only has income effects, on savings and on labour income.

Lemma 9. *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 α only has income effects, both on savings s and on labour income z:*

$$\forall k = s, z : \frac{\mathrm{d}k}{\mathrm{d}\alpha} = (1-t^e)p\frac{\mathrm{d}k}{\mathrm{d}\rho_2}. \tag{51}$$

Proof. Note the following partial derivatives, taking into account additive separability and substituting (38)–(40):

$$\mathrm{E}[\tilde{\mathcal{U}}_{z\alpha}] = (1-t^e)p\mathrm{E}[\tilde{\mathcal{U}}_{z\rho_2}], \qquad \mathrm{E}[\tilde{\mathcal{U}}_{s\alpha}] = (1-t^e)p\mathrm{E}[\tilde{\mathcal{U}}_{s\rho_2}],$$

$$\mathrm{E}[\tilde{\mathcal{U}}_{f\alpha}] = (1-t^e)p\mathrm{E}[\tilde{\mathcal{U}}_{f\rho_2}], \qquad \mathrm{E}[\tilde{\mathcal{U}}_{p\alpha}] = (1-t^e)\mathrm{E}[\tilde{u}_2] + (1-t^e)p\mathrm{E}[\tilde{\mathcal{U}}_{s\rho_2}].$$

Like in the proof of Lemma 8, use Cramer's rule and Lemma 5 to find the following comparative statics:

$$\frac{\mathrm{d}z}{\mathrm{d}\alpha} = (1-t^e)p\frac{\mathrm{d}s}{\mathrm{d}\rho_2} - \frac{(1-t^e)\mathrm{E}\left[\tilde{u}_2\right]}{\det\left(\mathrm{E}[\tilde{\mathcal{H}}]\right)}\det\left(\mathrm{E}\left[\begin{pmatrix}\tilde{\mathcal{U}}_{zs} & \tilde{\mathcal{U}}_{zf} & \tilde{\mathcal{U}}_{zp}\\ \tilde{\mathcal{U}}_{ss} & \tilde{\mathcal{U}}_{sf} & \tilde{\mathcal{U}}_{sp}\\ \tilde{\mathcal{U}}_{fs} & \tilde{\mathcal{U}}_{ff} & \tilde{\mathcal{U}}_{fp}\end{pmatrix}\right]\right), \tag{52}$$

$$\frac{\mathrm{d}s}{\mathrm{d}\alpha} = (1-t^e)p\frac{\mathrm{d}s}{\mathrm{d}\rho_2} - \frac{(1-t^e)\mathrm{E}\left[\tilde{u}_2\right]}{\det\left(\mathrm{E}[\tilde{\mathcal{H}}]\right)}\det\left(\mathrm{E}\left[\begin{pmatrix}\tilde{\mathcal{U}}_{zz} & \tilde{\mathcal{U}}_{zf} & \tilde{\mathcal{U}}_{zp}\\ \tilde{\mathcal{U}}_{sz} & \tilde{\mathcal{U}}_{sf} & \tilde{\mathcal{U}}_{sp}\\ \tilde{\mathcal{U}}_{fz} & \tilde{\mathcal{U}}_{ff} & \tilde{\mathcal{U}}_{fp}\end{pmatrix}\right]\right). \tag{53}$$

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 (7)–(8) for portfolio optimization:

$$\mathrm{E}[\tilde{\mathcal{U}}_{sf}] = (1-t^e)\mathrm{E}\left[(\tilde{r}^m - r^b)\tilde{u}_{22}\right]\frac{\mathrm{E}\left[\tilde{u}_1\right]}{\mathrm{E}\left[\tilde{u}_2\right]} = 0,$$

$$\mathrm{E}[\tilde{\mathcal{U}}_{sp}] = (1-t^e)\mathrm{E}\left[(\alpha + \tilde{\varepsilon} - r^b)\tilde{u}_{22}\right]\frac{\mathrm{E}\left[\tilde{u}_1\right]}{\mathrm{E}\left[\tilde{u}_2\right]} = 0.$$

Note furthermore that taking into account additive separability, (43) and (46) imply that also $\mathrm{E}[\tilde{\mathcal{U}}_{zf}] = \mathrm{E}[\tilde{\mathcal{U}}_{zp}] = 0$. Substitute $\mathrm{E}[\tilde{\mathcal{U}}_{zf}] = \mathrm{E}[\tilde{\mathcal{U}}_{zp}] = \mathrm{E}[\tilde{\mathcal{U}}_{sf}] = \mathrm{E}[\tilde{\mathcal{U}}_{sp}] = 0$ into (52)–(53) to find that the determinants in the numerators equal zero, and thus (51) is correct. □

A.3 Lagrangian for the government optimization problem

The Lagrangian for the government's optimization problem in terms of the perturbation parameters ϵ^z, ϵ^P, t^n and t^e is:

$$\begin{aligned}\Lambda(\epsilon^z, &\epsilon^P, t^n, t^e)\\ &\equiv \iint_\Theta \mathrm{E}[\tilde{\mathcal{V}}(\boldsymbol{\theta}, \epsilon^z, t^n, t^e, \epsilon^P)]\mathrm{d}G^\theta(\boldsymbol{\theta}) - \mathrm{E}[\tilde{\lambda}\cdot\{P(\tilde{r}^m) + \epsilon^P\eta^P(\tilde{r}^m)\}]\\ &\quad + (1+r^b)\mathrm{E}[\tilde{\lambda}]\iint_\Theta \{t^\ell(z(\boldsymbol{\theta}, \epsilon^z, t^n, t^e, \epsilon^P)) + \epsilon^z\eta^z(z(\boldsymbol{\theta}, \epsilon^z, t^n, t^e, \epsilon^P))\}\mathrm{d}G^\theta(\boldsymbol{\theta})\\ &\quad + \iint_\Theta \{t^n\mathrm{E}[\tilde{\lambda}]y^n(\boldsymbol{\theta}, \epsilon^z, t^n, t^e, \epsilon^P) + t^e\mathrm{E}[\tilde{\lambda}\tilde{y}^e(\boldsymbol{\theta}, \epsilon^z, t^n, t^e, \epsilon^P)]\}\mathrm{d}G^\theta(\boldsymbol{\theta}),\end{aligned} \tag{54}$$

where we omit superfluous function arguments for legibility. The first-order conditions require that partial derivatives of the Lagrangian with respect to ϵ^z, ϵ^P, t^n and t^e equal zero, for all reform functions η^z and η^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 ϵ^z requires that the condition $\partial\Lambda/\partial\epsilon^z = 0$ is satisfied for every perturbation function η^z. Evaluate the condition $\partial\Lambda/\partial\epsilon^z = 0$ for the Lagrangian (54) in the situation before any reforms, substitute envelope condition (27) for $\mathrm{E}[\tilde{\mathcal{V}}_{\epsilon^z}]$, substitute decomposition (41) for $\forall k = z, y^n, y^e : k^z_\epsilon$, and substitute definition (16) for $\beta(\boldsymbol{\theta})$:

$$0 = \iint_\Theta (1 + r^b - \beta(\boldsymbol{\theta}))\eta^z(z(\boldsymbol{\theta}))\mathrm{d}G^\theta(\boldsymbol{\theta}) - \iint_\Theta \mathcal{W}(\boldsymbol{\theta})\eta^z_z(z(\boldsymbol{\theta}))\mathrm{d}G^\theta(\boldsymbol{\theta}). \tag{55}$$

To further interpret condition (55), we will rewrite it as an integral over labour incomes z. Condition (55) becomes:

$$0 = \int_{\underline{z}}^{\overline{z}} \left(1 + r^b - \overline{\beta}(z)\right)\eta^z(z)\mathrm{d}G^z(z) - \int_{\underline{z}}^{\overline{z}} \overline{\mathcal{W}}(z)\eta^z_z(z)\mathrm{d}G^z(z). \tag{56}$$

Use partial integration to rewrite the second integral of (56):

$$\begin{aligned} 0 = \int_{\underline{z}}^{\overline{z}} &\left\{\left(1 + r^b - \overline{\beta}(z)\right)g^z(z) + \frac{\mathrm{d}[\overline{\mathcal{W}}(z)g^z(z)]}{\mathrm{d}z}\right\}\eta^z(z)\mathrm{d}z \\ &- \overline{\mathcal{W}}(\overline{z})g^z(\overline{z})\eta^z(\overline{z}) + \overline{\mathcal{W}}(\underline{z})g^z(\underline{z})\eta^z(\underline{z}). \end{aligned} \tag{57}$$

The latter condition must be valid for every perturbation function η^z. Let us first focus on the perturbation functions that are zero on the boundaries, $\eta^z(\underline{z}) = \eta^z(\overline{z}) = 0$. The second line of (57) is then zero. Consequently, the first line of (57) must also be zero. We will prove by contradiction that the terms between curly brackets of (57) must then sum to zero. Suppose that these terms do not sum to zero on some interval between $\underline{z}$ and $\overline{z}$. Choose then a perturbation function η^z which has the same sign everywhere as the sum of the terms between curly brackets. The integral on the first line of (57) 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 η^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:

$$(1 + r^b - \overline{\beta}(z))g^z(z) = -\frac{\mathrm{d}[\overline{\mathcal{W}}(z)g^z(z)]}{\mathrm{d}z}. \tag{58}$$

Having established that the first line of (57) must be zero, let us now consider perturbation functions η^z which differ from zero at one of the end points, $\eta^z(\underline{z}) \neq 0$ or $\eta^z(\overline{z}) \neq 0$. Since the second line of (57) must be zero, it follows that the tax wedges on the end points must be zero. We thus find the following transversality conditions:

$$\overline{\mathcal{W}}(\overline{z}) = \overline{\mathcal{W}}(\underline{z}) = 0. \tag{59}$$

Integrating (58) and using transversality conditions (59), we find the first-order condition for the tax on labour income:

$$\overline{\mathcal{W}}(z)g^{\mathrm{z}}(z) = \int_{z}^{\overline{z}} (1 + r^b - \overline{\beta}(z))\mathrm{d}G^{\mathrm{z}}(z). \tag{60}$$

The optimal marginal labour income tax (60) 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. Substitute transversality condition (59) into first-order condition (60) to find condition the standard optimality condition for the welfare weights in second-period values:

$$\iint_\Theta \beta(\boldsymbol{\theta})\mathrm{d}G^\theta(\boldsymbol{\theta}) = 1 + r^b. \tag{61}$$

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 $\partial\Lambda/\partial t^n = 0$:

$$\iint_\Theta \left\{ \left(1 - \frac{\mathrm{E}[\tilde{U}_2]}{\mathrm{E}[\tilde{\lambda}]}\right) y^n + (1+r^b)t_z^\ell z_{t^n} + t^n y_{t^n}^n + t^e \frac{\mathrm{E}[\tilde{\lambda}\tilde{y}_{t^n}^e]}{\mathrm{E}[\tilde{\lambda}]} \right\} \mathrm{d}G^\theta(\boldsymbol{\theta}) = 0,$$

where we use envelope condition (28), and we evaluate in the situation before any reforms. Apply Slutsky decompositions (30) to find:

$$\iint_\Theta \left\{ (1 + r^b - \beta(\boldsymbol{\theta})) \frac{\mathrm{E}[\tilde{U}_2]}{\mathrm{E}[\tilde{U}_1]} y^n + (1+r^b)t_z^\ell z_{t^n}^* + t^n y_{t^n}^{n*} + t^e \frac{\mathrm{E}[\tilde{\lambda}\tilde{y}_{t^n}^{e*}]}{\mathrm{E}[\tilde{\lambda}]} \right\} \mathrm{d}G^\theta(\boldsymbol{\theta}) = 0, \tag{62}$$

where we use property (37) to substitute for second-period income effects, we substitute Euler equation (6) and we substitute definition (16) of $\beta(\boldsymbol{\theta})$.[18]

Change the integration domain to the labour incomes and rearrange:

$$\int_{\underline{z}}^{\overline{z}} \left\{ \frac{\mathrm{E}[\tilde{U}_2]}{\mathrm{E}[\tilde{U}_1]} \overline{(1 + r^b - \beta)\, y^n} + (1+r^b)t_z^\ell z_{t^n}^* + t^n y_{t^n}^{n*} + t^e \frac{\mathrm{E}[\tilde{\lambda}\overline{\tilde{y}_{t^n}^{e*}}]}{\mathrm{E}[\tilde{\lambda}]} \right\} \mathrm{d}G^{\mathrm{z}}(z) = 0, \tag{63}$$

[18] In the remainder of this proof we will employ the optimality of the tax schedule on labour income. Eq. (62) remains valid if the tax schedule on labour income is not optimal. In that case, supposing the intercept of the tax system is still optimized and transversality condition (61) remains valid, the marginal excess burden of taxing normal capital income should be balanced against a distributional characteristic of normal capital income:

$$\mathrm{cov}_\Theta(\beta, y^n) = \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} \iint_\Theta \left\{ (1+r^b)t_z^\ell z_{t^n}^* + t^n y_{t^n}^{n*} + t^e \frac{\mathrm{E}[\tilde{\lambda}\tilde{y}_{t^n}^{e*}]}{\mathrm{E}[\tilde{\lambda}]} \right\} \mathrm{d}G^\theta(\boldsymbol{\theta}).$$

where we use the constancy of $\mathrm{E}[\tilde{U}_2]/\mathrm{E}[\tilde{U}_1]$ due to Euler equation (6).

Let us focus now on the first term in the integrand of (63). Write this term as a covariance over the population Θ, conditional on labour income:

$$\int_{\underline{z}}^{\overline{z}} \overline{(1+r^b-\beta)\, y^n} \mathrm{d}G^z(z) = -\int_{\underline{z}}^{\overline{z}} \left\{ \mathrm{cov}_\Theta(\beta, y^n|z) - \left(1+r^b-\overline{\beta}(z)\right)\overline{y^n}\right\} \mathrm{d}G^z(z). \tag{64}$$

Use integration by parts and use transversality condition (61) to rewrite the right-hand side of (64):

$$\int_{\underline{z}}^{\overline{z}} \overline{(1+r^b-\beta)\, y^n} \mathrm{d}G^z(z) = \int_{\underline{z}}^{\overline{z}} \left\{ \frac{\int_z^{\overline{z}} \left(1+r^b-\overline{\beta}(\hat{z})\right) \mathrm{d}G^z(\hat{z})}{g^z(z)} \frac{\mathrm{d}\overline{y}^n}{\mathrm{d}z} - \mathrm{cov}_\Theta(\beta, y^n|z)\right\} \mathrm{d}G^z(z). \tag{65}$$

Substitute (65) into optimality condition (63) and substitute the optimal labour tax condition (60):[19]

$$\begin{aligned}\int_{\underline{z}}^{\overline{z}} \mathrm{cov}_\Theta(\beta, y^n|z)\mathrm{d}G^z(z) = &\int_{\underline{z}}^{\overline{z}} \overline{\mathcal{W}}(z)\left(\frac{\mathrm{d}\overline{y^n}}{\mathrm{d}z} - \frac{\overline{\partial y^{nc}}}{\partial z}\right)\mathrm{d}G^z(z) - \int_{\underline{z}}^{\overline{z}} \mathrm{cov}_\Theta\left(\mathcal{W}, y^{nc}_z|z\right)\mathrm{d}G^z(z) \\ &+ \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}\int_{\underline{z}}^{\overline{z}} \left(\overline{\mathcal{W}y^{nc}_z}\frac{\mathrm{E}[\tilde{U}_2]}{\mathrm{E}[\tilde{U}_1]} + (1+r^b)t^\ell_z \overline{z^*_{t^n}} + t^n \overline{y}^{n*}_{t^n} + t^e \frac{\mathrm{E}\left[\tilde{\lambda}\overline{\tilde{y}^{e*}_{t^n}}\right]}{\mathrm{E}[\tilde{\lambda}]}\right)\mathrm{d}G^z(z).\end{aligned}$$

Substitute the definition of the labour wedge (17):

$$\begin{aligned}\int_{\underline{z}}^{\overline{z}} \mathrm{cov}_\Theta(\beta, y^n|z)\mathrm{d}G^z(z) = &\int_{\underline{z}}^{\overline{z}} \overline{\mathcal{W}}(z)\left(\frac{\mathrm{d}\overline{y^n}}{\mathrm{d}z} - \frac{\overline{\partial y^{nc}}}{\partial z}\right)\mathrm{d}G^z(z) - \int_{\underline{z}}^{\overline{z}} \mathrm{cov}_\Theta\left(\mathcal{W}, y^{nc}_z|z\right)\mathrm{d}G^z(z) \\ &+ \int_{\underline{z}}^{\overline{z}} \left((1+r^b)t^\ell_z \left(\overline{z^*_{t^n}\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} - z^*_\sigma y^{nc}_z}\right)\right)\mathrm{d}G^z(z) \\ &+ t^n \int_{\underline{z}}^{\overline{z}} \left(\overline{y^{n*}_{t^n}\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} - y^{n*}_\sigma y^{nc}_z}\right)\mathrm{d}G^z(z) \\ &+ t^e \frac{\mathrm{E}\left[\tilde{\lambda}\int_{\underline{z}}^{\overline{z}}\left(\overline{\tilde{y}^{e*}_{t^n}\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} - \tilde{y}^{e*}_\sigma y^{nc}_z}\right)\mathrm{d}G^z(z)\right]}{\mathrm{E}[\tilde{\lambda}]}.\end{aligned} \tag{66}$$

We rewrite some of the terms on the right-hand side of (66). First, use Slutsky symmetry (31) to rewrite:

$$z^*_{t^n}\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]} - z^*_\sigma y^{nc}_z = y^*_\sigma - z^*_\sigma y^{nc}_z = 0, \tag{67}$$

where the last step follows from the fact that, keeping constant labour income z, a compensated perturbation of the tax on labour income does not directly affect the savings choices.

[19] Add and subtract $\int_{\underline{z}}^{\overline{z}} \overline{\mathcal{W}}(z)\overline{y}^{nc}_z \mathrm{d}G^z(z)$ to find the result.

Second, use Slutsky symmetry (31) to rewrite:

$$y_{t^n}^{n*}\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}-y_\sigma^{n*}y_z^{nc}=\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}\left(y_{t^n}^{n*}-z_{t^n}^{*}y_z^{nc}\right)=\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}y_{t^n}^{nc*} \tag{68}$$

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 to rewrite:

$$\tilde{y}_{t^n}^{e*}\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}-\tilde{y}_\sigma^{e*}y_z^{nc}=\tilde{y}_{t^n}^{e*}\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}-y_\sigma^{n*}\tilde{y}_z^{ec}=\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}\left(\tilde{y}_{t^n}^{e*}-z_{t^n}^{*}\tilde{y}_z^{ec}\right)=\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}\tilde{y}_{t^n}^{ec*}, \tag{69}$$

where we apply Slutsky symmetry (31) in the second step. Substitute (67)–(69) into (66) to find the optimality condition:

$$\int_{\underline{z}}^{\overline{z}}\mathrm{cov}_\Theta(\beta,y^n|z)\mathrm{d}G^{\mathrm{z}}(z)=\int_{\underline{z}}^{\overline{z}}\overline{\mathcal{W}}(z)\left(\frac{\mathrm{d}\overline{y^n}}{\mathrm{d}z}-\frac{\overline{\partial y^{nc}}}{\partial z}\right)\mathrm{d}G^{\mathrm{z}}(z)-\int_{\underline{z}}^{\overline{z}}\mathrm{cov}_\Theta\left(\mathcal{W},y_z^{nc}|z\right)\mathrm{d}G^{\mathrm{z}}(z)$$
$$+t^n\int_{\underline{z}}^{\overline{z}}\overline{y_{t^n}^{nc*}}\mathrm{d}G^{\mathrm{z}}(z)\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}+t^e\frac{\mathrm{E}\left[\tilde{\lambda}\int_{\underline{z}}^{\overline{z}}\overline{\tilde{y}_{t^n}^{ec*}}\mathrm{d}G^{\mathrm{z}}(z)\right]}{\mathrm{E}[\tilde{\lambda}]}\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{U}_2]}.$$

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. This immediately yields condition (18).

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:

$$\mathrm{E}\left[\left(\iint_\Theta\tilde{U}_P\mathrm{d}G^\theta(\boldsymbol{\theta})-\tilde{\lambda}\right)\tilde{\eta}^P\right]=0, \tag{70}$$

where we use envelope property (29) 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 for the optimal provision of the public good:

$$\lambda^{r^m}=\iint_\Theta\mathrm{E}_\mathcal{E}[\tilde{U}_P|r^m]\mathrm{d}G^\theta(\boldsymbol{\theta}). \tag{71}$$

Take the expected value of (71), divide by $\mathrm{E}[\tilde{\lambda}]$ and rearrange to find:

$$1=\iint_\Theta(1+r^b)\frac{\mathrm{E}[\tilde{U}_P]}{\mathrm{E}[\tilde{U}_1]}\mathrm{d}G^\theta(\boldsymbol{\theta})+\iint_\Theta\left(\frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{\lambda}]}-(1+r^b)\right)\frac{\mathrm{E}[\tilde{U}_P]}{\mathrm{E}[\tilde{U}_1]}\mathrm{d}G^\theta(\boldsymbol{\theta}). \tag{72}$$

The term $\mathrm{E}[\tilde{U}_P]/\mathrm{E}[\tilde{U}_1]$ 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 (72) reflects the effect on social welfare of a marginal increase in public good provision when all individuals have equal social weights, so $\mathrm{E}[\tilde{U}_1]/\mathrm{E}[\tilde{\lambda}] = 1 + r^b$ 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 $\mathrm{E}[\tilde{U}_P]/\mathrm{E}[\tilde{U}_1]$ increases sufficiently with the social weight $\mathrm{E}[\tilde{U}_1]/\mathrm{E}[\tilde{\lambda}]$, then the second term on right-hand side of (72) is positive. This implies that the first term on the right-hand side of (72) must be smaller than one, and thus that the provision of the public good must be larger than if distributional motives were ignored.

A.8 Proof of Corollary 2

Proof. Our reasoning in Subsection 3.2.1 clarifies why the optimal tax on excess returns must be in the range $0 < t^e < 1$. To see how the optimal t^e responds to the importance of aggregate risk or the degree of risk aversion towards the public good, note that (10) is a differential equation defining excess returns as a function of the tax rate. Its solution is $\tilde{y}^e(t^e) = \tilde{y}^e(0)/(1-t^e)$. Substitute this solution and (10) into (20):

$$\iint_\Theta \mathrm{E}\left[\tilde{y}^e(0)\right] \mathrm{d}G^\theta(\boldsymbol{\theta}) = -\iint_\Theta \frac{\mathrm{cov}\left(\iint_\Theta \mathrm{E}_\mathcal{E}[\tilde{U}_P|\tilde{r}^m]dG^\theta(\boldsymbol{\theta}), \tilde{y}^e(0)\right)}{\mathrm{E}[\tilde{\lambda}]}\mathrm{d}G^\theta(\boldsymbol{\theta}).$$

The left side is positive and independent of t^e, while the right side ranges from zero to infinity as t^e varies from 0 to 1 and individuals increase their risky investments. The value of t^e where both sides are equal decreases with more aggregate risk or increased risk aversion towards the public good. □

A.9 Proof of Corollary 3

Proof. Rearrange the integrals in (20) and substitute (19):

$$\iint_\Theta \mathrm{E}\left[\tilde{y}^e_{t^e}\right] \mathrm{d}G^\theta(\boldsymbol{\theta}) + \frac{\mathrm{cov}\left(\iint_\Theta \mathrm{E}_\mathcal{E}[\tilde{U}_P|\tilde{r}^m]dG^\theta(\boldsymbol{\theta}), \iint_\Theta \tilde{y}^e_{t^e}\mathrm{d}G^\theta(\boldsymbol{\theta})\right)}{\iint_\Theta \mathrm{E}[\tilde{U}_P]dG^\theta(\boldsymbol{\theta})} = 0.$$

Given our assumption that the returns to the risky assets are jointly normally distributed, we can apply a result by Rubinstein (1976):[20]

$$\iint_\Theta \mathrm{E}\left[\tilde{y}^e_{t^e}\right] \mathrm{d}G^\theta(\boldsymbol{\theta}) + \frac{\iint_\Theta \mathrm{E}[\tilde{U}_{PP}] dG^\theta(\boldsymbol{\theta})}{\iint_\Theta \mathrm{E}[\tilde{U}_P] dG^\theta(\boldsymbol{\theta})} \mathrm{cov}\left(\tilde{P}, \iint_\Theta \tilde{y}^e_{t^e} \mathrm{d}G^\theta(\boldsymbol{\theta})\right) = 0.$$

Substitute the individual budget constraint (4) and the government's budget constraint (14), and use (10):

$$\iint_\Theta \mathrm{E}\left[\tilde{y}^e_{t^e}\right] \mathrm{d}G^\theta(\boldsymbol{\theta}) + \frac{t^e}{1-t^e} \frac{\iint_\Theta \mathrm{E}[\tilde{U}_{PP}] dG^\theta(\boldsymbol{\theta})}{\iint_\Theta \mathrm{E}[\tilde{U}_P] dG^\theta(\boldsymbol{\theta})} \mathrm{var}\left(\iint_\Theta \tilde{y}^e \mathrm{d}G^\theta(\boldsymbol{\theta})\right) = 0.$$

Finally, use our assumption of constant global absolute risk aversion with respect to the provision of the public good and rearrange to find (21). □

A.10 Proof of Lemma 3

Proof of Lemma 3. Use (3) and the fact that the private shocks are i.i.d. to find $\mathrm{cov}(\tilde{\lambda}, \tilde{y}^e_\nu) = \mathrm{cov}(\tilde{\lambda}, \tilde{r}^m) f_\nu$. Multiply and divide the right-hand side by $\iint_\Theta f \mathrm{d}G^\theta(\boldsymbol{\theta})$ to find:

$$\frac{\mathrm{cov}(\tilde{\lambda}, \tilde{y}^e_\nu)}{\mathrm{E}[\tilde{\lambda}]} = \frac{f_\nu}{\iint_\Theta f \mathrm{d}G^\theta(\boldsymbol{\theta})} \iint_\Theta \frac{\mathrm{cov}(\tilde{\lambda}, \tilde{y}^e)}{\mathrm{E}[\tilde{\lambda}]} \mathrm{d}G^\theta(\boldsymbol{\theta}).$$

Substitute the government's optimality condition (18) and use (10):

$$\frac{\mathrm{cov}(\tilde{\lambda}, \tilde{y}^e_\nu)}{\mathrm{E}[\tilde{\lambda}]} = -\frac{f_\nu}{\iint_\Theta f \mathrm{d}G^\theta(\boldsymbol{\theta})} \iint_\Theta \mathrm{E}[\tilde{y}^e] \mathrm{d}G^\theta(\boldsymbol{\theta}).$$

We thus find:

$$t^e \iint_\Theta \left(\mathrm{E}\left[\tilde{y}^e_\nu\right] + \frac{\mathrm{cov}\left(\tilde{\lambda}, \tilde{y}^e_\nu\right)}{\mathrm{E}[\tilde{\lambda}]}\right) \mathrm{d}G^\theta(\boldsymbol{\theta}) = t^e \iint_\Theta \left(\mathrm{E}\left[\tilde{y}^e_\nu\right] - \frac{f_\nu}{\iint_\Theta f \mathrm{d}G^\theta(\boldsymbol{\theta})} \iint_\Theta \mathrm{E}[\tilde{y}^e] \mathrm{d}G^\theta(\boldsymbol{\theta})\right) \mathrm{d}G^\theta(\boldsymbol{\theta})$$

Substitute (3) for $\tilde{y}^e$ on the right-hand side:

$$t^e \iint_\Theta \left(\mathrm{E}\left[\tilde{y}^e_\nu\right] + \frac{\mathrm{cov}\left(\tilde{\lambda}, \tilde{y}^e_\nu\right)}{\mathrm{E}[\tilde{\lambda}]}\right) \mathrm{d}G^\theta(\boldsymbol{\theta}) = t^e \left(\iint_\Theta p_\nu(\alpha - r^b) \mathrm{d}G^\theta(\boldsymbol{\theta}) - \frac{\iint_\Theta p(\alpha - r^b) \mathrm{d}G^\theta(\boldsymbol{\theta})}{\iint_\Theta f \mathrm{d}G^\theta(\boldsymbol{\theta})} \iint_\Theta f_\nu \mathrm{d}G^\theta(\boldsymbol{\theta})\right).$$

These derivations extend to compensated effects and effects conditional on labour income. Verify that the latter corresponds to Lemma 3. □

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

[20] Rubinstein (1976, , p.421) finds that for any variables $\tilde{x}$ and $\tilde{y}$ that are bivariate normal, and any function $g(\tilde{y})$ that is at least once differentiable, we have the property $\mathrm{cov}(\tilde{x}, g(\tilde{y})) = \mathrm{E}[g'(\tilde{y})]\mathrm{cov}(\tilde{x}, \tilde{y})$.

combination of them has a univariate normal distribution. The second-period individual budget constraint (4) then implies that second-period consumption is normally distributed. Government budget constraint (14) shows that second-period consumption remains normally distributed. 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:

$$\max_{p,f} \mathrm{E}[u(c^1, \tilde{c}^2)|c^1], \tag{73}$$

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 $\mathrm{E}[\tilde{c}^2]$ and standard deviation $\mathrm{sd}[\tilde{c}^2]$. Then the following stochastic quantity follows the standardized normal distribution:

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

Denote the probability density function of $\tilde{n}$ as $\varphi(\tilde{n})$. The objective function (73) can then be rewritten in terms of the mean and the standard deviation of second-period consumption:

$$\mathrm{E}[u(c^1, \tilde{c}^2)|c^1] = \int_{-\infty}^{+\infty} u(c^1, \mathrm{E}[\tilde{c}^2] + n \cdot \mathrm{sd}[\tilde{c}^2])\varphi(n)\mathrm{d}n. \tag{74}$$

Solving this portfolio optimization problem leads to the following Lemma.

Lemma 10. *In the Mean-Variance framework, the ratio of investments in the risky assets is determined by:*

$$\frac{p}{f} = \frac{\alpha - r^b}{\mathrm{E}\left[\tilde{r}^m\right] - r^b} \frac{\mathrm{var}\left(\tilde{r}^m\right)}{\mathrm{var}\left(\tilde{\varepsilon}\right)}. \tag{75}$$

The relative proportion p/f invested in the risky assets is increasing in private investment productivity α, and is unaffected by skill, taxes, or the size of the portfolio.

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.

Denote the semi-elasticity of the investment $\gamma = f, p$ with respect to some perturbation parameter ν as $\xi^\gamma_\nu \equiv \partial \log \gamma / \partial \nu$. Taking logarithms on both sides of (75) and taking derivatives with respect to the perturbation parameter, results in the following lemma.

Lemma 11. *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:*

$$\forall \nu = \sigma, t^n, \rho : \; \xi_\nu^f = \xi_\nu^p.$$

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

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 10.

Let us now verify the consequences of Lemma 11 for our model. First assume that ξ_ν^p and ξ_ν^f are constant over the income distribution. Lemma 3 then implies:

$$\iint_\Theta \left(\mathrm{E}\left[\tilde{y}_\nu^e\right] + \frac{\mathrm{cov}\left(\tilde{\lambda}, \tilde{y}_\nu^e\right)}{\mathrm{E}[\tilde{\lambda}]} \right) \mathrm{d}G^\theta(\boldsymbol{\theta}) = \iint_\Theta (\alpha - r^b) p \mathrm{d}G^\theta(\boldsymbol{\theta}) \cdot (\xi_\nu^p - \xi_\nu^f).$$

Lemma 11 then implies:

$$\iint_\Theta \left(\mathrm{E}\left[\tilde{y}_\nu^e\right] + \frac{\mathrm{cov}\left(\tilde{\lambda}, \tilde{y}_\nu^e\right)}{\mathrm{E}[\tilde{\lambda}]} \right) \mathrm{d}G^\theta(\boldsymbol{\theta}) = 0.$$

It follows that when the elasticities ξ_ν^p and ξ_ν^f are constant over the income distribution and the mean-variance framework applies, the welfare effects of changes in tax revenue from excess capital income due to a perturbation are exactly cancelled out by the welfare effects of the uncertainty of those changes in tax revenue.

Assume instead that α is constant, so all individuals obtain the same expected rate of return to their private investments. Lemma 3 then implies:

$$\iint_\Theta \left(\mathrm{E}\left[\tilde{y}_\nu^e\right] + \frac{\mathrm{cov}\left(\tilde{\lambda}, \tilde{y}_\nu^e\right)}{\mathrm{E}[\tilde{\lambda}]} \right) \mathrm{d}G^\theta(\boldsymbol{\theta}) = \iint_\Theta (\alpha - r^b) p \mathrm{d}G^\theta(\boldsymbol{\theta}) \cdot \left(\frac{\iint_\Theta p \xi_\nu^p \mathrm{d}G^\theta(\boldsymbol{\theta})}{\iint_\Theta p \mathrm{d}G^\theta(\boldsymbol{\theta})} - \frac{\iint_\Theta f \xi_\nu^f \mathrm{d}G^\theta(\boldsymbol{\theta})}{\iint_\Theta f \mathrm{d}G^\theta(\boldsymbol{\theta})} \right). \tag{76}$$

Then Lemma 10 shows that p/f is equal for all individuals and that $\xi_\nu^p = \xi_\nu^f$ for all individuals. It follows again that:

$$\iint_\Theta \left(\mathrm{E}\left[\tilde{y}_\nu^e\right] + \frac{\mathrm{cov}\left(\tilde{\lambda}, \tilde{y}_\nu^e\right)}{\mathrm{E}[\tilde{\lambda}]} \right) \mathrm{d}G^\theta(\boldsymbol{\theta}) = 0.$$

It follows that also when all individuals have the same expected return α from private investment and the mean-variance framework applies, the welfare effects of changes in tax revenue from excess capital income due to a perturbation are exactly cancelled out by the welfare effects of the uncertainty of those changes in tax revenue.