The Optimal Taxation of Risky Capital Income: The Rate-of-Return Allowance

Kevin Spiritus,* Robin Boadway†

January 2, 2017

PRELIMINARY VERSION

We study the optimality of taxing capital income by the Rate-of-Return Allowance (RRA) proposed by the Mirrlees Review. In a mean-variance framework with constant returns to scale in private investment, the optimal tax on risk-free returns is zero, but positive (negative) with decreasing (increasing) returns. The optimal tax on excess returns to risky assets is positive if the stochastic tax revenue is returned to the household by variable public good provision. If it is returned as a stochastic lump sum, the optimal tax on excess returns is irrelevant with only aggregate risk, and approaches 100% if there is also idiosyncratic risk.

Keywords: optimal capital taxation, Rate-of-Return Allowance, risk, excess returns
JEL Classification: H21, H23, H24

1. Introduction

The Mirrlees Review [2011] proposed a novel set of tax policies for capital income, aimed at taxing excess returns while exempting risk-free returns. Interest on interest-bearing accounts, which are acquired out of after-tax income (T), would be exempt from tax as it accumulates (E) and when the asset is sold (E). This tax-prepaid treatment is denoted as TEE, indicating the times at which taxes are due. Risky assets and personal business assets on the other hand would receive TtE treatment. Only returns up to a Rate-of-Return Allowance (RRA) would be tax-free where the RRA is a risk-free interest rate. Those above the RRA, which include above-normal returns and returns to risk, would be taxed (t). Furthermore, pension savings for retirement would continue to be taxed on a registered or EET basis. They

---

*Research Centre of Public Economics, KU Leuven, Naamsestraat 69 box 3565, 3000 Leuven, Belgium
    E-mail: kevin.spiritus@kuleuven.be.
†Department of Economics, Queen's University, K7L 3N6 Kingston, Ontario, Canada
    E-mail: boadwayr@econ.queensu.ca.

would be tax-deductible when the asset is acquired (E) and tax-exempt as returns accumulate (E), with accumulated asset values taxable when cashed in (T). Such a system accepts that risk-free normal returns to assets should be tax-exempt as in a classical progressive consumption tax system, but that excess returns to capital should be taxed, presumably on the argument that they include windfall gains.

The RRA system is an administratively complicated way to tax excess returns to assets. The same objective is achieved by the EET system proposed for pension savings, since accumulated returns that are taxed on withdrawal include any risk-premiums and above-normal gains that have accrued. The difference is mostly one of timing: government revenues accrue later under EET than under RRA. But the bigger issue concerns the public economic rationale for exempting risk-free returns to assets while taxing excess returns. This paper addresses that question in an optimal income tax framework.

It is well established that under certain circumstances, an optimal nonlinear income tax would exempt capital income. (Banks and Diamond [2010] summarize the arguments.) This would be the case in a two-period life-cycle setting where individuals with different wage rates and identical preferences supply labour in the first period, save in risk-free assets to finance consumption over the two periods, and where preferences are weakly separable between goods and labour. The case for taxing capital income arises once these assumptions are violated, for example, if utility discount rates are correlated with wage rates, if wage rates are stochastic, or if there are borrowing constraints.

In this paper, we adopt the above assumptions that would lead to zero taxation of capital income in a risk-free setting with no excess returns. We then consider whether the exemption of the risk-free component of capital income continues to hold when some assets earn excess returns. An important consideration is that besides return to risk, excess returns could include other elements such as inframarginal rents, advantages of scale or skill premiums, and it is impossible to separate these components for tax purposes. It is therefore practically impossible to tax rents, which may be desirable on efficiency grounds, without at the same time taxing returns to risk. To confront this issue, we allow for three types of assets: a risk-free asset, risky assets that yield a competitive return, and those that yield an idiosyncratic and possibly above-normal return. 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.

The presence of assets with idiosyncratic risk has generally been avoided in the recent dynamic optimal tax literature, where the focus has been on idiosyncratic wage risk. Buchholz and Konrad [2014] summarize the arguments, noting the private sector may choose not to insure all risks for incentive reasons, in which case government insurance will introduce inefficiencies. Only in specific cases, e.g. when there are positive spillovers from the activities of inventors or when an asset cannot be traded because it is productive only in the hands of a specific individual, can there be good reasons to provide public insurance. Be that as it may, we show that responses of portfolio allocations between the two types of risk have an impact on the optimal tax on risk-free returns. Since the non-corporate sector forms a significant part of the economy, with startups relying on private equity and owners of family businesses preferring to hold their shares closely, the presence of idiosyncratic risk cannot simply be ignored.

We derive circumstances under which the tax on the risk-free component is zero, while that on the excess component is positive, which would correspond to the RRA system. In the classic case where there are no above-normal returns, so all excess returns are returns to risk, a tax on the excess returns leads to portfolio choice effects similar to Domar and Musgrave [1944], Mossin [1968] and Stiglitz [1969]. In our setting, the ultimate effect of the tax on excess returns, and by extension our optimal tax results, depend on two key considerations.

The first is how the government disposes of stochastic tax revenues. If all risk were idiosyncratic, tax revenues would be deterministic. An optimizing government would then let the tax rate on excess returns approach 100%, providing pure insurance.[1] As soon as there is some aggregate risk though, government revenues become stochastic. This uncertainty somehow needs to be returned to the population, generally removing the possibility of full insurance. We consider two possibilities. In one we assume, following Atkinson and Stiglitz [1980] and Gordon [1985], that stochastic revenues are returned to individuals by state-contingent equal per capita lump sums. If there is only aggregate risk, returning stochastic tax revenues to the households undoes the insurance effect of the capital income tax in equilibrium, and the tax rate on excess returns would be irrelevant. As soon as there is also some idiosyncratic risk, the government would provide full insurance by taxing excess returns at 100%.

The other possibility is that the government returns the risky tax revenues to the households via public goods. In this case, as long as the public good is not a perfect substitute for consumption, stochastic public goods mitigate the consumption risk from holding market assets. An increase in the tax on excess returns reduces uncertain capital income and therefore consumption risk, but it increases risk associated with consuming the public good. This assumption is consistent with observed government behaviour, e.g. during the 2008 recession.[2] Christiansen [1993] follows this approach in studying the effect of taxes on portfolio choice in a representative individual model with fixed income and identifies the trade-off between public and private consumption. Schindler [2008] also uses this assumption to study the effect of taxes on portfolio choice in a representative individual model with fixed income and one risky asset with linear returns. He finds that risk-free assets should not be taxed in an efficient tax system, while excess returns should be taxed, balancing public risk against private risk. Our model confirms the last finding in a more general setting, favouring a positive tax on excess returns.

If we allow the government to return the uncertainty of its revenues via a combination of a stochastic lump sum and a stochastic public good, we find that in the case without idiosyncratic risk, the tax on excess returns is again redundant. As soon as there is some idiosyncratic risk, the optimal tax on excess returns again approaches 100%, providing pure insurance. The trade-off against private and public risk remains, now balancing the lump

---

[1]To keep the model tractable, we disregard incentive and moral hazard effects on investment. Moreover, returns to capital do not depend on investment skill. Gerritsen et al. [2016] look at skill premiums in the absence of uncertainty, and conclude that returns to capital should be taxed at a positive rate when individuals with higher labour productivity obtain higher rents. They do not consider the possibility of an RRA.

[2]Alesina et al. [2015] provide a list of austerity measures for a number of OECD countries, distinguishing between tax increases and spending cuts, and Skidmore and Scorsone [2011] and Chernick et al. [2011] document the impact of the recession on local expenditures in the United States.

sum and the public good to reach the optimum.

The second consideration concerns the technology of investment returns. In the case of market assets, we assume that returns are independent of scale and the same for all investors. However, for private investments, we allow returns to be non-constant in the size of the investment. With decreasing returns, which is standard in the literature (e.g., Mintz [1981]), the marginal expected return is decreasing and less than the average return, so inframarginal rents are generated. The case where returns are increasing in portfolio size was suggested by Piketty [2014], and has since been empirically validated by Kacperczyk et al. [2016] and Fagereng et al. [2016]. The latter showed that the correlation cannot be explained by risk premiums alone. One possible explanation is the existence of threshold effects: minimum wealth required by private banks, minimum investments required by venture capital funds, and so on. In this case, the expected rate of return is increasing in portfolio size, so higher-income investors with larger investments will obtain a higher return. Whether returns to private investment are increasing or decreasing influences the optimal tax on risk-free returns.

In our setting with heterogeneous individuals and a nonlinear labour income tax, the optimality of an RRA-equivalent tax system depends on these two considerations. We show that the sign of the optimal tax on risk-free returns depends on whether investment is shifted towards private or market assets in case of a reform to that tax. In the mean-variance framework, if private investments exhibit constant returns to scale, then the composition of the risky portfolio is unaffected by policy reforms, and the optimal tax on risk-free returns on all assets should be zero. Taken together with the above cases where the optimal tax on excess returns is positive, this represents the RRA system. However, if expected returns on private investments are increasing, the tax on risk-free returns should be negative, and vice versa if returns are decreasing. These results hold regardless of how the risky tax revenues are returned to the taxpayers.

We begin in section 2 by outlining a general model with both risky and risk-free assets, in which the risky assets can take two forms. As mentioned, one is the purchase of a market portfolio where competitive capital markets ensure that all remaining risk is aggregate. The other is a personal investment, such as a private business or private equity investments where risks are assumed to be idiosyncratic. We allow for increasing or decreasing marginal rates of return on the private assets. Individuals differ in wage rates and face a nonlinear tax on their labour income. We present a solution method suitable for our problem that involves deriving the optimal tax system using a perturbation method analogous to that of Saez [2001]. Section 3 then solves the individual optimization problem, and finds relevant Slutsky properties. We derive optimal linear tax rates on risk-free and excess returns to capital, optimal nonlinear tax rates on labour income and public good decision rules in section 4. Section 5 concludes.

2. The Model

In this section we outline the basic features of the model. We begin by describing the characteristics of individuals, including their preferences and the three decisions they make:

labour supply, saving and portfolio choice. We then turn to the government and the labour and capital income tax policies it chooses to finance a public good from a stochastic budget constraint. Finally, we outline the perturbation methodology we use to analyze government policies.

2.1. Individuals

There is a continuum of individuals endowed with skills $w \in \left[\underline{w}, \overline{w}\right]$, which are distributed by the function $G^{\mathsf{w}}(w)$, with corresponding density function $g^{\mathsf{w}}(w) = G_w^{\mathsf{w}}(w)$, where $G_w^{\mathsf{w}}(w) \equiv \mathrm{d}G^{\mathsf{w}}(w)/\mathrm{d}w$, a convention we adopt throughout the paper. Skills determine the effective labour generated by a given effort.

Outcomes are stochastic. Individuals differ ex ante only in their skill levels, and those with the same skill level w make the same decisions. We denote equilibrium values of variables shared by individuals of skill w by the superscript $\cdot^w$. Ex post there will be differences among individuals with the same skill level.

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

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

Savings a are invested in three kinds of assets: bonds b, market funds f and private investment opportunities p such that:

$$a = b + f + p.$$

Bonds are risk-free and yield a normal return r^b, so bond income in period two is $r^b b$. Market funds yield a stochastic market rate of return $\tilde{r}^m$, impacted by aggregate shocks, where stochastic variables are denoted by a tilde. Total returns from market funds are then $\tilde{r}^m f$. The rate $\tilde{r}^m$ is drawn from a distribution function $G^{\mathsf{m}}(r^m)$ with corresponding density function $g^{\mathsf{m}}(r^m) \equiv G_{r^m}^{\mathsf{m}}(r^m)$. The domain for these functions is denoted $\mathscr{R}$. Investment in private investment opportunities p yields a return $\tilde{F}(p) \equiv \alpha(p) + \tilde{\varepsilon} \cdot p$, where $\tilde{\varepsilon}$ is an idiosyncratic shock that is independent and identically distributed, thus uncorrelated to the market returns, and drawn from a distribution function $G^{\mathsf{p}}(\varepsilon)$, with corresponding density function $g^{\mathsf{p}}(\varepsilon) \equiv G_\varepsilon^{\mathsf{p}}(\varepsilon)$. The domain for these distribution functions is denoted $\mathscr{E}$. Although we assume that the function α is strictly increasing ($\alpha_p > 0$), we allow the expected marginal rate of return α_p to be either constant, strictly increasing or strictly decreasing everywhere ($\alpha_{pp} \gtreqless 0$). If $\alpha_{pp} < 0$, so returns to scale are decreasing, inframarginal rents will be obtained. If $\alpha_{pp} > 0$, e.g. due to threshold effects, the marginal rate of return will increase with the scale of the investment. The assumption that market risks are aggregate while private investment opportunities are subject to idiosyncratic risk captures the assumption that capital markets fully insure idiosyncratic risk on market portfolios but private investments are not fully insured.

When choosing labour supply, savings and portfolio composition, the individual knows his skill w and the distribution functions G^m and G^p of the capital income shocks, but not the realization of the shocks $\tilde{r}^m$ and $\tilde{\varepsilon}$.

Total capital income realized in the second period is denoted as:

$$\tilde{y} \equiv r^b \cdot (a - f - p) + \tilde{r}^m f + \tilde{F}(p).$$

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 y^e such that $\tilde{y} \equiv y^n + \tilde{y}^e$, with:

$$y^n \equiv r^b a, \quad \text{and} \quad \tilde{y}^e \equiv \tilde{r}^m f + \tilde{F}(p) - r^b \cdot (f + p).$$

Recall that we define *excess returns* to refer to all capital income that deviates from the risk-free return, including risk premiums, economies of scale and stochastic shocks. This corresponds with the definition used in the Mirrlees Report [2011] in their proposed RRA system for risky assets. Individuals pay taxes $t^n y^n$ on the risk-free part of their capital income, and $t^e \tilde{y}^e$ on excess returns. As discussed below, the government may return part of the risk of its tax revenues to the households in the second period using an equal per capita lump sum $\tilde{K}$. Second-period consumption thus equals assets saved plus second period after-tax capital income and the lump sum:

$$\tilde{c}^2 \equiv a + (1 - t^n) y^n + (1 - t^e) \tilde{y}^e + \tilde{K}.$$

The government chooses its income tax function and capital income tax rates in the first period. It obtains labour income tax revenues in the first period and capital income tax revenues in the second, and must satisfy an intertemporal budget constraint described below. If there is aggregate risk in the capital markets, tax revenues will be stochastic. The government returns this risk to the individuals in the second period, using the stochastic lump sum $\tilde{K}$, stochastic provision of a pure public good $\tilde{P}$, or a combination of both. The realization of these instruments is not known when individuals make their decisions and the government chooses its tax structure.

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

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

This utility function displays weak separability between intertemporal allocation and labour effort (so preferences over c^1 and c^2 are independent of ℓ). As well, we assume the public good is weakly separable from all individual choice variables, so does not affect them. For ease of notation, we did not write the corresponding subutility function. Absolute consumption risk aversion is decreasing, as is standard and in line with empirical evidence (see e.g. Friend and Blume, 1975).

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

2.2. The Government

Each individual reports three variables to the government: labour income z^w, risk-free capital income $y^{n,w} \equiv r^b a^w$ and excess capital income $y^{e,w} \equiv y^w - r^b a^w$.[3] These variables allow the government to impose a tax on income z^w, which we assume to take the nonlinear form $t^\ell(z^w)$, and taxes on risk-free capital income $y^{n,w}$ and excess capital income $y^{e,w}$, which we assume for simplicity are at the linear rates t^n and t^e. One reason to use linear tax rates on capital income is that they might be levied through financial institutions, easing compliance. If this is the case, the observability requirements are less stringent than those in our model: the variables no longer need to be observed on an individual basis. This does not pose a problem for the implementation of our results, since the optimal tax equilibrium derived depends only on sufficient statistics. These are weighted population averages which can be inferred from surveys or administrative data. These statistics include labour supply responses to tax reforms, savings responses, and also income densities and probability distributions of the returns to capital.

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

The government sets these tax instruments, together with the levels of the lump sum $K(r^m)$ and spending on the public good $P(r^m)$, corresponding to each potential realization of the market rate of return $\tilde{r}^m$, to maximize an additive social welfare function:

$$\max_{\{t^\ell(\cdot), t^n, t^e, P(\cdot), K(\cdot)\}} \int_{\underline{w}}^{\overline{w}} \mathrm{E}\big[\tilde{V}^w \big(t^\ell(\cdot), t^n, t^e, P(\tilde{r}^m), K(\tilde{r}^m) \big) \big] \mathrm{d} G^w(w), \tag{1}$$

subject to the intertemporal budget constraint in second-period values:

$$K(\tilde{r}^m) + P(\tilde{r}^m) = \int_{\underline{w}}^{\overline{w}} \left[(1 + r^b) t^\ell(z^w) + t^e \int_{\mathscr{E}} y^{e,w}(\tilde{r}^m, \epsilon) g^i(\epsilon) \mathrm{d}\epsilon + t^n y^{n,w} \right] \mathrm{d} G^w(w), \tag{2}$$

where $\mathrm{E}\big[\tilde{V}^w \big(t^\ell(\cdot), t^n, t^e, P(\tilde{r}^m), K(\tilde{r}^m) \big) \big]$ is the expected maximized level of utility of an individual with skill w, given the government policies, and $y^{e,w}(\tilde{r}^m, \epsilon)$ is the realized excess capital income for given capital income shocks.[4] The government has access to the bond market between the periods. Over time its budget is balanced. The *law of large numbers* ensures that, with sufficiently large populations at each skill level, the government budget constraint is not affected by the idiosyncratic shocks. Aggregate shocks do however cause government revenue to be stochastic.

---

[3]Equivalently, the government could require reporting of savings a^w and capital income y^w, allowing inference of risk-free and excess returns.

[4]We avoid differences between ex-ante and ex-post objectives by using a utilitarian social welfare function. Our results remain valid when individual expected utilities are transformed by a concave, increasing transformation Ψ, so with ex-ante government objective $\int_{\underline{w}}^{\overline{w}} \Psi \big(\mathrm{E}\big[\tilde{V}^w \big\{ t^\ell(\cdot), t^n, t^e, P(\tilde{r}^m), K(\tilde{r}^m) \big\} \big] \big) \mathrm{d} G^w(w)$.

The lump sum and the provision of public goods vary with the aggregate shock. We treat the levels to be provided for each potential realization of the market shock—that is, $K(\tilde{r}^m)$ and $P(\tilde{r}^m)$—as policy instruments in the government's optimization problem.

2.3. Tax Reforms

To solve our model using a mechanism design approach would be difficult, given the ex-post heterogeneity of the individuals, given the restrictions to the tax functions, and given the way the uncertainty of government revenues needs to be returned to the individuals. We analyze optimal policies using a perturbation method analogous to Saez [2001], but adopted to our context with linear capital income taxes, a stochastic lump sum and a stochastic public good. Given the complexity of our model, we work with separate perturbations on each of the different policy instruments, rather than following the usual approach of simultaneously perturbing various instruments in a budget-neutral way (e.g. Piketty and Saez, 2013). We model local reforms of non-linear policy instruments more rigorously than is usually done, making sure that all effects are taken into account. The latter is important, given the additional terms that are brought about by uncertainty. Our starting point is that when the tax and expenditure policies have been set optimally, a marginal reform to any of them has no impact on expected social welfare.

Begin by decomposing the labour income tax function at a given value of z into an intercept term plus an integral term over the marginal tax rates:

$$t^\ell(z) \;=\; t^\ell(0) + \int_0^z t_z^\ell(\zeta)\,\mathrm{d}\zeta.$$

Note that even though we defined social welfare as an integral over the skill distribution, we introduce the tax reforms at a given level of labour income. This contrasts with the original approach of Mirrlees [1971], who studies the optimal bundle to be assigned to each skill type.

The effect on $t^\ell(z)$ of any labour income tax reform can be decomposed into a linear combination of reforms to the tax intercept and to any of the marginal tax rates at different income levels. In what follows we focus on two reforms: a reform of the tax intercept, and a reform of the marginal tax rate at one particular level of the tax base. We derive optimality conditions for the effects of such reforms. The conditions for the marginal tax rate at one particular level of z can be generalized to any other level.

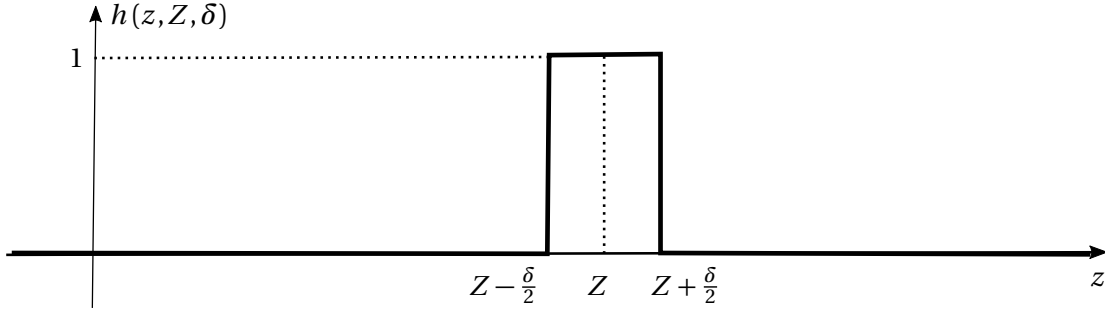Define the following rectangular function, for some level Z of the labour income tax base and for any arbitrarily small parameter δ:

$$h : \mathbb{R}^3 \to \mathbb{R} : (z, Z, \delta) \mapsto \begin{cases} Z - \frac{\delta}{2} < z < Z + \frac{\delta}{2} : & 1, \\ \text{elsewhere}: & 0. \end{cases} \tag{3}$$

This function is shown in Figure 1. It will help us introduce a reform of the marginal tax rate at level Z.

Figure 1: The rectangular function $h(z, Z, \delta)$.

Denote by $H(z, Z, \delta) \equiv \int_0^z h(\zeta, Z, \delta) \,\mathrm{d}\zeta$ the primitive of the function $h(z, Z, \delta)$ defined in (3). It has the following property:

$$\forall z \in \mathbb{R}: H(z, Z, \delta) = \begin{cases} z \leq Z - \frac{\delta}{2}: & 0, \\ Z - \frac{\delta}{2} < z < Z + \frac{\delta}{2}: & z - Z + \frac{\delta}{2}, \\ z \geq Z + \frac{\delta}{2}: & \delta. \end{cases} \tag{4}$$

Note that $\partial H(z, Z, \delta) / \partial z = h(z, Z, \delta)$. The function $H(z, Z, \delta)$ is shown in Figure 2.

Figure 2: The function $H(z, Z, \delta)$.
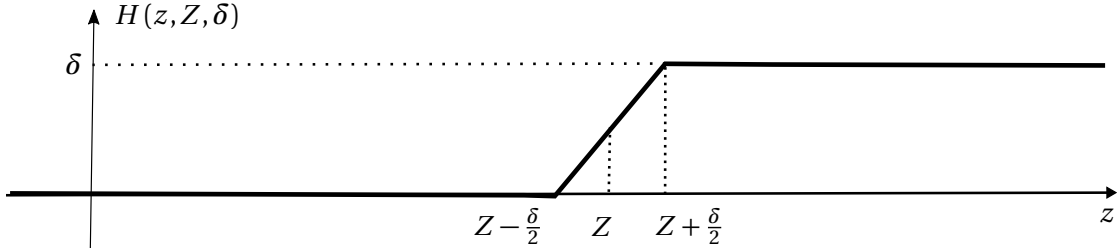
Now, introduce reform parameter ρ for the tax intercept and reform parameter σ^z for the marginal tax rate at some level Z of the labour income tax base. Applying these leads to the reformed tax function:

$$\tau^\ell\left(z | \rho, \sigma^z, Z, \delta\right) \equiv t^\ell(z) + \rho + \sigma^z H(z, Z, \delta), \tag{5}$$

with corresponding marginal tax rate:

$$\tau_z^\ell\left(z | \rho, \sigma^z, Z, \delta\right) \equiv t_z^\ell(z) + \sigma^z h(z, Z, \delta).$$

The original and reformed tax functions are shown in Figure 3.

A reform of the σ^z-parameter will have a local effect on marginal tax rates, and an income effect at higher levels of the tax base. A reform of the ρ-parameter will have income effects

Figure 3: The functions $t^{\ell}(z)$ and $\tau^{\ell}\left(z|\rho,\sigma^z,Z,\delta\right)$.

only, for all individuals in the population. If the tax function $t^{\ell}(\cdot)$ is optimally set, the optimal values of the reform parameters ρ and σ^z should be zero. This is studied in the next sections.
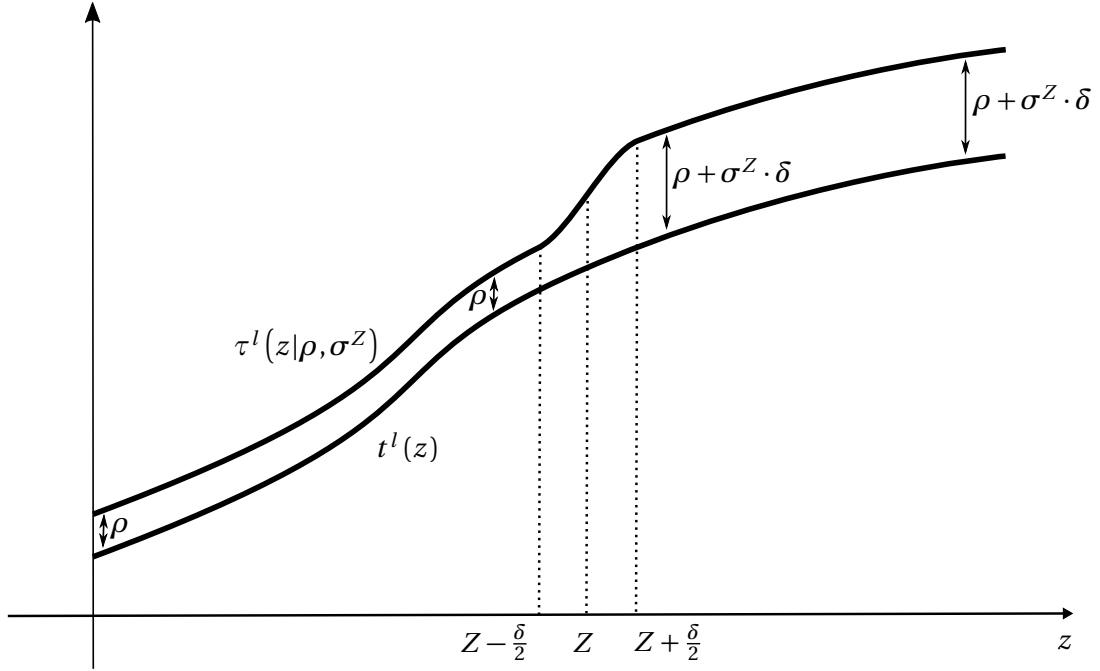
For the linear capital income taxes we work directly with the parameters t^n and t^e. The vector of tax reform parameters is summarized as $\overrightarrow{\sigma} \equiv \left(\rho, \sigma^z, t^e, t^n\right)$.

2.4. Reforms to the Public Good and the Lump Sum

When setting its tax policies and knowing that its intertemporal budget constraint (2) has to be satisfied for any realization of the shock to the market returns $\tilde{r}^m$, the government implicitly sets the level of the public good and the lump sum as functions of the shocks:

$$P : \mathscr{R} \to \mathbb{R} : r^m \mapsto P\left(r^m\right),$$
$$K : \mathscr{R} \to \mathbb{R} : r^m \mapsto K\left(r^m\right).$$

It sets its policies before the realization of the shocks, choosing levels of $P\left(r^m\right)$ and $K\left(r^m\right)$ for each potential realization of the market shock $\tilde{r}^m$.

Analogous to the reform of the labour income tax function, we consider reforms of these functions on an interval of width δ around some arbitrarily chosen realization R^m of the market shock. Denoting the sizes of the reforms respectively as π^m and κ^m, the reformed

public expenditure and lump sum functions can be written as follows:

$$\mathscr{P}\left(r^m|\pi^m, R^m, \delta\right) \equiv P\left(r^m\right) + \pi^m \cdot h\left(r^m, R^m, \delta\right), \tag{6}$$

$$\mathscr{K}\left(r^m|\pi^m, R^m, \delta\right) \equiv K\left(r^m\right) + \kappa^m \cdot h\left(r^m, R^m, \delta\right). \tag{7}$$

When the public good function P and the lump sum function L are optimal, the optimal sizes π^m and κ^m of the reforms are zero at all values of R^m.

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

3. Individual Optimization Problem

Consider the individual two-period optimization problem in a situation where the policy reform parameters, ρ, σ^z, π^m and κ^m, are not necessarily zero. An individual of type w chooses effective labour supply z, first-period consumption c^1, total savings a and investment in assets with excess returns f and p to maximize expected utility. The Lagrangian for this maximization problem is (omitting the reform parameters for the public good and the lump sum):

$$\mathscr{L}^w\left(c^1, z, a, f, p, \mu | \overrightarrow{\sigma}\right) \equiv \mathrm{E}\left[U\left(u\left(c^1, c^2\left(a, f, p | \overrightarrow{\sigma}, \tilde{r}^m, \tilde{\varepsilon}\right)\right), \frac{z}{w}, \mathscr{P}\left(\tilde{r}^m\right)\right)\right]$$
$$- \mu \cdot \left[c^1 - z + \tau^\ell\left(z | \rho, \sigma^z, Z, \delta\right) + a\right], \tag{8}$$

with second-period consumption given by:

$$c^2\left(a, f, p | \overrightarrow{\sigma}, \tilde{r}^m, \tilde{\varepsilon}\right) = a + \left(1 - t^n\right) r^b a + \left(1 - t^e\right)\left[\tilde{r}^m f + \tilde{F}\left(p\right) - r^b \cdot \left(f + p\right)\right] + \mathscr{K}\left(\tilde{r}^m\right). \tag{9}$$

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

We indicate partial derivatives of any function using subscripts. To facilitate notation, we denote $\tilde{U}_1 \equiv \left(\partial \tilde{U}/\partial u\right) \cdot \left(\partial u/\partial c_1\right)$ and $\tilde{U}_2 \equiv \left(\partial \tilde{U}/\partial u\right) \cdot \left(\partial u/\partial c_2\right)$. The first-order conditions on earnings z and savings a are standard and are as follows:

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

The first-order conditions on portfolio choices f and p can be written:

$$\left(1 - t^e\right)\mathrm{E}\left[\tilde{U}_2 \cdot \left(\tilde{r}^m - r^b\right)\right] = 0, \quad \left(1 - t^e\right)\mathrm{E}\left[\tilde{U}_2 \cdot \left(\tilde{F}_p - r^b\right)\right] = 0. \tag{11}$$

11

Rearrange conditions (11) to find the marginal risk premiums required by the individuals:

$$\left(1-t^e\right)\mathrm{E}\left[\tilde{r}^m-r^b\right]=-\left(1-t^e\right)\frac{\mathrm{cov}\left(\tilde{U}_2,\tilde{r}^m\right)}{\mathrm{E}\left[\tilde{U}_2\right]}, \tag{12}$$

$$\left(1-t^e\right)\mathrm{E}\left[\tilde{F}_p-r^b\right]=-\left(1-t^e\right)\frac{\mathrm{cov}\left(\tilde{U}_2,\tilde{\varepsilon}\right)}{\mathrm{E}\left[\tilde{U}_2\right]}. \tag{13}$$

We assume throughout that second-order conditions are met.

The solution to this problem gives uncompensated demand functions for a type$-w$ individual, which depend on the reform parameters for the taxes and the lump sum, but by assumption not on the public good: $c^{1,w}\left(\overrightarrow{\sigma},\kappa^m\right)$, $z^w\left(\overrightarrow{\sigma},\kappa^m\right)$, $a^w\left(\overrightarrow{\sigma},\kappa^m\right)$, $f^w\left(\overrightarrow{\sigma},\kappa^m\right)$ and $p^w\left(\overrightarrow{\sigma},\kappa^m\right)$. Corresponding to these demand functions are the following uncompensated tax base functions:

$$y^{n,w}\left(\overrightarrow{\sigma},\kappa^m\right)=r^b a^w\left(\overrightarrow{\sigma},\kappa^m\right), \tag{14}$$

$$\tilde{y}^{e,w}\left(\overrightarrow{\sigma},\kappa^m\right)=\tilde{r}^m f^w\left(\overrightarrow{\sigma},\kappa^m\right)+\tilde{F}\left(p^w\left(\overrightarrow{\sigma},\kappa^m\right)\right)-r^b f^w\left(\overrightarrow{\sigma},\kappa^m\right)-r^b p^w\left(\overrightarrow{\sigma},\kappa^m\right). \tag{15}$$

The indirect expected utility function $\mathrm{E}\left[\tilde{V}^w\left(\overrightarrow{\sigma},\pi^m,\kappa^m\right)\right]$ is the maximum function for this problem for given policy reform parameters. Apply the envelope theorem to Lagrangian (8) and use the definition of the reformed tax function (5), definitions (6)-(7) of reformed public expenditures and lump sums, and the individual second-period budget constraint (9) to find partial derivatives of the indirect expected utility function with respect to the policy parameters:

$$\mathrm{E}\left[\tilde{V}_\rho\right]=-\mathrm{E}\left[\tilde{U}_1\right];\quad \mathrm{E}\left[\tilde{V}_{\sigma^z}\right]=-\mathrm{E}\left[\tilde{U}_1\right]H\left(z,Z,\delta\right); \tag{16}$$

$$\mathrm{E}\left[\tilde{V}_{t^n}\right]=-\mathrm{E}\left[\tilde{U}_2\right]y^n;\quad \mathrm{E}\left[\tilde{V}_{t^e}\right]=-\mathrm{E}\left[\tilde{U}_2\tilde{y}^e\right]; \tag{17}$$

$$\mathrm{E}\left[\tilde{V}_{\pi^m}\right]=\mathrm{E}\left[\tilde{U}_P h\left(\tilde{r}^m,R^m,\delta\right)\right];\quad \mathrm{E}\left[\tilde{V}_{\kappa^m}\right]=\mathrm{E}\left[\tilde{U}_2 h\left(\tilde{r}^m,R^m,\delta\right)\right]. \tag{18}$$

The effect of a tax reform on excess returns can be decomposed as follows:

$$\mathrm{E}\left[\tilde{V}_{t^e}\right]=-\mathrm{E}\left[\tilde{U}_2\right]\left(\mathrm{E}\left[\tilde{y}^e\right]+\frac{\mathrm{cov}\left(\tilde{U}_2,\tilde{y}^e\right)}{\mathrm{E}\left[\tilde{U}_2\right]}\right). \tag{19}$$

There is not only an adverse effect from the expected change in the tax liability on excess returns. The total effect is attenuated by the fact the uncertainty of private consumption is decreased by a tax increase.

The remainder of this section develops some of the building blocks that are used in the optimal-tax analysis. Subsection 3.1 derives the Slutsky properties, taking into account the local nature of the policy reforms and the fact that different reforms occur in different periods, subsection 3.2 derives effects of policy reforms conditional on labour earnings, and subsection 3.3 introduces the mean-variance framework.

3.1. Slutsky Properties

To find Slutsky properties consider the first-period cost minimization problem for a given level of expected utility $\mathrm{E}\left[\tilde{U}^w\right]=\overline{V}$. The Lagrangian for the dual optimization problem is as follows, where the asterisk superscript is used to indicate compensated values (omitting some reform parameters for ease of notation):

$$\mathscr{L}^{w*}\left(c^1,a,f,z,p,\mu^*|\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)\equiv c^1-z+t^\ell(z)+\rho+\sigma^z H(z,Z,\delta)+a$$
$$-\mu^*\mathrm{E}\left[U\left(u\left(c^1,c^2\left(a,f,p|\overrightarrow{\sigma},\kappa^m,\tilde{r}^m,\tilde{\varepsilon}\right)\right),\frac{z}{w},\mathscr{P}\left(\tilde{r}^m|\pi^m\right)\right)\right]$$
$$+\mu^*\overline{V}. \tag{20}$$

The solutions are Hicksian demand functions, which depend on tax parameters $\overrightarrow{\sigma}$, the public good parameter π^m, the lump sum parameter κ^m and the required expected level of utility $\overline{V}$. The minimum function defines the expenditure function $X^w\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$.

A desirable property of these Hicksian demand functions would be that they do not respond to a inframarginal changes in any of the tax functions since these give rise to pure income effects. Indeed, if there is such a change, and the individual is compensated in that same period, the size of the compensation will exactly be equal to the size of the reform. The original situation is restored and individuals will not change their behaviour. If the compensation happens in a different period though, individuals would alter their savings in order to restore their original consumption pattern. So, while it is natural to compensate for labour tax reforms in the first period, for capital tax reforms it is more natural to do so in the second period.

We thus introduce two sets of compensated demand functions. The first, with compensation in the first period, is denoted with a single superscript asterisk: $c^{1,w*}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$, $\ell^{w*}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$, $a^{w*}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$, $f^{w*}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$ and $p^{w*}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$, with corresponding tax bases $z^{w*}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$, $y^{n,w*}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$ and $y^{e,w*}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\overline{V}\right)$. The second set of compensated demand functions, with compensation in the second period, is denoted with a double superscript asterisk $\cdot^{**}$.

To derive the Slutsky equations, note that in the individual's optimum compensated demands should be equal to the uncompensated demands. In the case of first-period compensation this means:

$$\forall b=c^1,a,f,p,\ell,z:\quad b\left(\overrightarrow{\sigma},\kappa^m\right)=b^*\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\mathrm{E}\left[\tilde{V}\left(\overrightarrow{\sigma},\pi^m,\kappa^m\right)\right]\right).$$

Combining the partial derivatives of this identity with respect to ρ and σ^z, using (14) and (15) and the envelope properties (16), yields Slutsky equations:

$$\forall b=c^1,a,f,p,\ell,z,y^n,\tilde{y}^e:\quad b_{\sigma^z}^w=b_{\sigma^z}^*\cdot h(z,Z,\delta)+b_\rho\cdot H(z,Z,\delta). \tag{21}$$

The added term $h(z,Z,\delta)$ reflects the fact that inframarginal reforms and reforms above the current level of the tax base do not have compensated effects. Eq. (21) would still be correct without this term: we add it to facilitate notation later on.

The situation is more complex for reforms to the capital income tax rates or the stochastic lump sum. As mentioned, since these policy changes occur in the second period, it is

natural to define compensated effects as if they were compensated in the second period. The most direct way to model this would be to introduce a reform parameter ρ^2 for the second-period tax intercept. This would lead to traditional Slutsky equations, for example:

$$\forall b = c^1, a, f, \ell, p, z, y^n, \tilde{y}^e: \quad b_{t^n} = b_{t^n}^{**} + b_{\rho^2} \cdot y^n.$$

However, doing this makes it difficult to compare with effects of reforms in the first period, complicating our efforts to characterize the optimal tax equilibrium. It is useful to find Slutsky equations in terms of first-period income effects, even if compensation is defined to occur in the second period. In appendix A we derive the relevant Slutsky equations:

$$y_{t^n}^n = y_{t^n}^{n**} + \left(y_\rho^n + r^b \right) \frac{\mathrm{E}\left[\tilde{U}_2 \right]}{\mathrm{E}\left[\tilde{U}_1 \right]} y^n; \quad y_{t^e}^n = y_{t^e}^{n**} + \left(y_\rho^n + r^b \right) \frac{\mathrm{E}\left[\tilde{U}_2 \tilde{y}^e \right]}{\mathrm{E}\left[\tilde{U}_1 \right]}; \tag{22}$$

$$\tilde{y}_{t^n}^e = \tilde{y}_{t^n}^{e**} + \tilde{y}_\rho^e \frac{\mathrm{E}\left[\tilde{U}_2 \right]}{\mathrm{E}\left[\tilde{U}_1 \right]} y^n; \quad \tilde{y}_{t^e}^e = \tilde{y}_{t^e}^{e**} + \tilde{y}_\rho^e \frac{\mathrm{E}\left[\tilde{U}_2 \tilde{y}^e \right]}{\mathrm{E}\left[\tilde{U}_1 \right]}; \tag{23}$$

$$z_{t^n} = z_{t^n}^{**} + z_\rho \frac{\mathrm{E}\left[\tilde{U}_2 \right]}{\mathrm{E}\left[\tilde{U}_1 \right]} y^n; \quad z_{t^e} = z_{t^e}^{**} + z_\rho \frac{\mathrm{E}\left[\tilde{U}_2 \tilde{y}^e \right]}{\mathrm{E}\left[\tilde{U}_1 \right]}; \tag{24}$$

$$y_{\kappa^m}^n = y_{\kappa^m}^{n**} - \left(y_\rho^n + r^b \right) \frac{\mathrm{E}\left[\tilde{U}_2 h \left(\tilde{r}^m, R^m, \delta \right) \right]}{\mathrm{E}\left[\tilde{U}_1 \right]}; \tag{25}$$

$$\tilde{y}_{\kappa^m}^e = \tilde{y}_{\kappa^m}^{e**} - \frac{\mathrm{E}\left[\tilde{U}_2 h \left(\tilde{r}^m, R^m, \delta \right) \right]}{\mathrm{E}\left[\tilde{U}_1 \right]} \tilde{y}_\rho^e; \tag{26}$$

$$z_{\kappa^m} = z_{\kappa^m}^{**} - \frac{\mathrm{E}\left[\tilde{U}_2 h \left(\tilde{r}^m, R^m, \delta \right) \right]}{\mathrm{E}\left[\tilde{U}_1 \right]} z_\rho. \tag{27}$$

In Appendix B we show that a reform to the tax rate on excess returns has no compensated effects.

Lemma 1. *The compensated effects on labour income, savings and portfolio composition of a reform in the tax rate on excess capital income are zero:*[5]

$$z_{t^e}^{**} = a_{t^e}^{**} = f_{t^e}^{**} = p_{t^e}^{**} = 0.$$

To derive Slutsky symmetry relations for the compensated demands, first apply the envelope theorem to the Lagrangian (20) to find partial derivatives of the expenditure function:

$$X_{\sigma^z} = H(z, Z, \delta); \quad X_{t^n} = \frac{\mathrm{E}\left[\tilde{U}_2 \right]}{\mathrm{E}\left[\tilde{U}_1 \right]} \cdot y^n; \quad X_{t^e} = \frac{\mathrm{E}\left[\tilde{U}_2 \cdot \tilde{y}^e \right]}{\mathrm{E}\left[\tilde{U}_1 \right]}; \quad X_{\kappa^m} = -\frac{\mathrm{E}\left[\tilde{U}_2 h \left(\tilde{r}^m, R^m, \delta \right) \right]}{\mathrm{E}\left[\tilde{U}_1 \right]}.$$

---

[5]These results extend those of Schindler [2008], who finds that reforms to the tax on excess returns have no compensated effects in a model with fixed labour incomes.

Taking second-order derivatives then yields the Slutsky symmetries (applying lemma 1 in the second equation):

$$X_{\sigma^z t^n} = X_{t^n \sigma^z} \Longleftrightarrow h(z, Z, \delta) \cdot z_{t^n}^{**} = \frac{\mathrm{E}\big[\tilde{U}_2\big]}{\mathrm{E}\big[\tilde{U}_1\big]} \cdot y_{\sigma^z}^{n*}, \tag{28}$$

$$X_{\sigma^z t^e} = X_{t^e \sigma^z} \Longleftrightarrow \frac{\partial}{\partial \sigma^z}\left(\frac{\mathrm{E}\big[\tilde{U}_2 \tilde{y}^{e*}\big]}{\mathrm{E}\big[\tilde{U}_1\big]}\right) = 0, \tag{29}$$

$$X_{\sigma^z \kappa^m} = X_{\kappa^m \sigma^z} \Longleftrightarrow h(z, Z, \delta) \cdot z_{\kappa^m}^{**} = -\frac{\partial}{\partial \sigma^z}\left(\frac{\mathrm{E}\big[\tilde{U}_2^* h(\tilde{r}^m, R^m, \delta)\big]}{\mathrm{E}\big[\tilde{U}_1^*\big]}\right). \tag{30}$$

These results are similar to the traditional Slutsky equations, adapted though to the local nature of our labour tax reforms and our reforms to the stochastic lump sums.

3.2. Conditional Effects of Marginal Policy Reforms

When working out the optimal tax equations we will encounter a number of terms that turn out to be responses to policy reforms *conditional on labour income*. To find properties of these conditional responses it is useful to think of the individual problem as a two-stage optimization problem in which labour supply and therefore earnings are chosen first, and then earnings are allocated to consumption, savings and portfolio composition.

We study the problem starting from the second stage. Taking labour supply and first-period disposable income as given, individuals optimize utility conditional on these quantities by choosing first-period consumption, savings and portfolio composition. In the first stage they optimize their overall utility by choosing their labour supply, anticipating the outcome of second-period choices.

Consider a reform to the marginal labour income tax function. Because the reform does not affect relevant prices in the second-stage optimization problem, and because preferences are weakly separable between consumption and leisure, its only impact on second-stage choices is through changes in labour income:

$$\forall b = c^1, a, f, p, y^n, \tilde{y}^e: \quad b_{\sigma^z}^* = b_z^c\big(\overrightarrow{\sigma}, \kappa^m, \overline{V} | z\big(\overrightarrow{\sigma}, \kappa^m, \overline{V}\big)\big) \cdot z_{\sigma^z}^*, \tag{31}$$

where superscript $\cdot^c$ denotes variables conditional on labour income.

From this we obtain the following properties for reforms to the tax rate on risk-free returns and the stochastic lump sum:

$$\forall b = c^1, a, f, p, y^n, \tilde{y}^e: \forall v = t^n, \kappa^m:$$
$$b_v^{c**} = b_v^{**} - b_z^c \cdot z_v^{**} = b_v^{**} - \frac{b_{\sigma^z}^*}{z_{\sigma^z}^*} \cdot z_v^{**}. \tag{32}$$

In appendix C we show that property (31) also leads to the following lemma, which will be useful in what follows.

Lemma 2. *With weak separability of preferences between consumption and labour supply, the tax bases have the following properties:*

$$\frac{\mathrm{E}\left[\tilde{U}_2\right]}{\mathrm{E}\left[\tilde{U}_1\right]} \cdot \frac{\mathrm{d}y^n}{\mathrm{d}w} - \frac{z_{t^n}^{**}}{z_{\sigma^z}^*} \cdot \frac{\mathrm{d}z^w}{\mathrm{d}w} = 0, \tag{33}$$

$$\frac{\mathrm{d}}{\mathrm{d}w}\left(\frac{\mathrm{E}\left[\tilde{U}_2 \tilde{y}^e\right]}{\mathrm{E}\left[\tilde{U}_1\right]}\right) = 0, \tag{34}$$

$$\frac{\mathrm{d}}{\mathrm{d}w}\left(\frac{\mathrm{E}\left[\tilde{U}_2 h\left(\tilde{r}^m, R^m, \delta\right)\right]}{\mathrm{E}\left[\tilde{U}_1\right]}\right) + \frac{z_{\kappa^m}^{**}}{z_{\sigma^z}^*}\frac{\mathrm{d}z}{\mathrm{d}w} = 0. \tag{35}$$

To interpret these equations, it suffices to see that the left-hand sides are always responses to a change in ability w conditional on labour income. These effects are zero because of the separability of preferences between consumption and labour supply.

In Appendix B we derive the following properties for the compensated effects of the tax rate on risk-free returns, conditional on labour income.

Lemma 3. *The compensated effects on portfolio composition of a reform in the tax rate on risk-free capital income, conditional on labour income, work only through changes in the amount of savings:*

$$f_{t^n}^{c**}\big|_{\mathrm{d}a=0} = p_{t^n}^{c**}\big|_{\mathrm{d}a=0} = 0.$$

From lemma 1 it also follows that the compensated effects of a reform to the tax on excess returns, conditional on labour income, are zero.

3.3. The Mean-Variance Framework

As the above results show, reforms to the tax on excess returns have no compensated effects, while reforms to the tax on risk-free returns have compensated effects on the size of the portfolio, and through that channel only on the composition of the portfolio. It is thus useful to revisit the portfolio optimization problem, for given labour income z and for given size of the portfolio a.

A case of special interest is the *mean-variance* or *modern-portfolio-theory* approach, introduced by Markowitz [1952]. Suppose the returns of both risky assets are jointly normally distributed, so any linear combination of them has a univariate normal distribution. The second-period budget constraint (9) then implies that second-period consumption is normally distributed. This allows studying the individual portfolio optimization problem in the mean-variance framework. Whenever we work in this framework, we will explicitly mention it.

Taking labour income z and the size of the portfolio a as given, 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 following objective:

$$\max_{p,f} \mathrm{E}\left[u\left(c^1, \tilde{c}^2\right)\big|c^1\right].$$

Define the following stochastic quantity:

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

where $\mathrm{sd}[\tilde{c}^2]$ denotes the standard deviation of second-period consumption.

Assuming second-period consumption $\tilde{c}^2$ is normally distributed, $\tilde{n}$ follows the standardized normal distribution, with probability density function $\varphi(\tilde{n})$. The objective function can then be written:[6]

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

The following properties follow from second-period budget constraint (9):

$$\mathrm{E}\left[\tilde{c}_p^2\right] = \left(1 - t^e\right) \cdot \mathrm{E}\left[\tilde{F}_p(p) - r^b\right], \quad \mathrm{E}\left[\tilde{c}_f^2\right] = \left(1 - t^e\right) \cdot \mathrm{E}\left[\tilde{r}^m - r^b\right],$$

and since for each of the risky assets $v = p, f$:

$$\frac{\partial \mathrm{sd}\left[\tilde{c}^2\right]}{\partial v} = \frac{1}{2 \cdot \mathrm{sd}\left[\tilde{c}^2\right]} \cdot \frac{\partial \mathrm{var}\left[\tilde{c}^2\right]}{\partial v} = \frac{\mathrm{cov}\left(\tilde{c}^2, \tilde{c}_v^2\right)}{\mathrm{sd}\left[\tilde{c}^2\right]},$$

we find:

$$\frac{\partial \mathrm{sd}\left[\tilde{c}^2\right]}{\partial p} = \left(1 - t^e\right) \cdot \frac{\mathrm{var}(\tilde{\varepsilon})p}{\mathrm{sd}\left[\tilde{r}^m f + \tilde{\varepsilon}p\right]}, \quad \frac{\partial \mathrm{sd}\left[\tilde{c}^2\right]}{\partial f} = \left(1 - t^e\right) \cdot \frac{\mathrm{var}(\tilde{r}^m)f}{\mathrm{sd}\left[\tilde{r}^m f + \tilde{\varepsilon}p\right]}.$$

The first-order condition with respect to the amount invested in the market asset f then is:

$$\frac{\mathrm{E}\left[\tilde{r}^m - r^b\right] \cdot \mathrm{sd}\left[\tilde{r}^m f + \tilde{\varepsilon}p\right]}{\mathrm{var}(\tilde{r}^m) \cdot f} \quad = \quad -\frac{\int_{-\infty}^{+\infty} \tilde{u}_2 \cdot n \cdot \varphi(n)\,\mathrm{d}n}{\int_{-\infty}^{+\infty} \tilde{u}_2 \cdot \varphi(n)\,\mathrm{d}n}, \tag{36}$$

where we used the assumption that the shocks to the private asset are i.i.d., and thus uncorrelated with the shocks to the market asset. The first-order condition for the amount invested in the private asset p:

$$\frac{\mathrm{E}\left[\tilde{F}_p(p) - r^b\right] \cdot \mathrm{sd}\left[\tilde{r}^m f + \tilde{\varepsilon}p\right]}{\mathrm{var}(\tilde{\varepsilon}) p} \quad = \quad -\frac{\int_{-\infty}^{+\infty} \tilde{u}_2 \cdot n \cdot \varphi(n)\,\mathrm{d}n}{\int_{-\infty}^{+\infty} \tilde{u}_2 \cdot \varphi(n)\,\mathrm{d}n}. \tag{37}$$

If both first-order conditions are fulfilled, we find that $p/\left(f \cdot \left(\alpha_p(p) - r^b\right)\right)$ is a constant. Combine first-order conditions (36) and (37) to see this:

$$\frac{p}{f} = \left(\alpha_p(p) - r^b\right) \cdot \frac{\mathrm{var}(\tilde{r}^m)}{\mathrm{var}(\tilde{\varepsilon}) \cdot \mathrm{E}\left[\tilde{r}^m - r^b\right]}. \tag{38}$$

With linear returns to scale, $\alpha_{pp} = 0$, such that p/f is constant. This leads to the following lemma.

---

[6]In fact we only use the assumption that the distribution is fully determined by its first two moments. We are thus interested in the impact of the tax reform on these moments.

Lemma 4. *In the mean-variance framework, the quantity $p/\left(f \cdot \left(\alpha_p - r^b\right)\right)$ is unaffected by skill, taxes or the size of the portfolio. If there are constant returns to scale, the relative proportions p/f invested in the risky assets are constant.*

Denote the compensated semi-elasticity of the investment $\gamma = f, p$ with respect to some tax reform parameter ν as $\xi_\nu^\gamma \equiv -\partial \log \gamma / \partial \nu$.

Lemma 5. *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:*

$$\forall \nu = \sigma^z, t^n, t^e, \rho : \; \xi_\nu^f = \xi_\nu^p \cdot \left(1 - \frac{\alpha_{pp} p}{\alpha_p - r^b}\right).$$

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

Proof. When there is a tax reform, individuals update their portfolio composition in such a way that (38) remains valid. Take logarithms on both sides and derivatives with respect to ν. □

4. Government Optimization

The first-order conditions for the government optimization problem (1)–(2) require that any marginal reform of the tax functions, of the provision of the public good or of the stochastic lump sum leaves the value of the objective function unchanged. We argued that any reform of the labour income tax can be decomposed into a linear combination of reforms of the tax intercept and local reforms of the marginal tax rates over arbitrarily small intervals. To find optimal tax equations, it thus suffices to find optimal values of the reform parameters ρ and σ^z for all labour incomes Z and of the capital income tax parameters t^n and t^e. Similarly, to find the optimal provision of the public good and the lump sum, it suffices to find the optimal values of the reform parameters π^m and κ^m for any possible value R^m of the market shock.

To each realization of the market shock r^m corresponds a budget multiplier λ^{r^m}. To facilitate notation, we introduce a *stochastic budget multiplier* $\tilde\lambda \equiv \lambda^{\tilde r^m}$. Write the Lagrangian for the government optimization problem as:

$$
\begin{aligned}
\Lambda&\left(\rho, \sigma^z, t^n, t^e, \pi^m, \kappa^m, \tilde\lambda \mid Z, R^m, \delta\right) \\
&= \int_{\underline{w}}^{\overline{w}} \mathrm{E}\left[\tilde V^w\left(\rho, \sigma^z, t^n, t^e, \pi^m, \kappa^m\right)\right] \mathrm{d}G^{\mathrm{w}}(w) \\
&\quad - \mathrm{E}\left[\tilde\lambda \cdot \left\{P\left(\tilde r^m\right) + K\left(\tilde r^m\right) + \left(\pi^m + \kappa^m\right) \cdot h\left(\tilde r^m, R^m, \delta\right)\right\}\right] \\
&\quad + \left(1 + r^b\right) \mathrm{E}\left[\tilde\lambda\right] \int_{\underline{w}}^{\overline{w}} \left[t^\ell\left(z^w\right) + \rho + \sigma^z H\left(z^w, Z, \delta\right)\right] \mathrm{d}G^{\mathrm{w}}(w) \\
&\quad + t^n \mathrm{E}\left[\tilde\lambda\right] \int_{\underline{w}}^{\overline{w}} y^{n,w} \mathrm{d}G^{\mathrm{w}}(w) + t^e \int_{\underline{w}}^{\overline{w}} \mathrm{E}\left[\tilde\lambda y^{e,w}\left(\tilde r^m, \tilde\varepsilon\right)\right] \mathrm{d}G^{\mathrm{w}}(w).
\end{aligned}
\tag{39}
$$

Let the term $G^z(z^w)$ denote the cumulative distribution function for labour incomes in the tax optimum, and let the term $g^z(z^w)$ be the corresponding density function. To keep things tractable, we assume that z^w is monotonous in w, such that $G^w(w) = G^z(z^w)$ and $g^w(w) = g^z(z^w)\mathrm{d}z^w/\mathrm{d}w$.[7] The first-order conditions are derived in appendix D. Rearranging them yields:

$$\sigma^z: \quad \forall w \in [\underline{w}, \overline{w}]: \mathscr{W}^w = -\frac{1}{g^z(z^w)}\int_w^{\overline{w}}\left(1 + r^b - \beta^{\hat{w}}\right)\mathrm{d}G^w(\hat{w}), \tag{40}$$

$$t^e: \quad \int_{\underline{w}}^{\overline{w}}\frac{\mathrm{cov}(\tilde{\lambda}, \tilde{y}^e)}{\mathrm{E}[\tilde{\lambda}]}\mathrm{d}G^w(w) = \int_{\underline{w}}^{\overline{w}}\frac{\mathrm{cov}(\tilde{U}_2, \tilde{y}^e)}{\mathrm{E}[\tilde{U}_2^w]}\mathrm{d}G^w(w), \tag{41}$$

$$t^n: \quad t^n\int_{\underline{w}}^{\overline{w}}y_{t^n}^{nc**}\mathrm{d}G^w(w) = -t^e\int_{\underline{w}}^{\overline{w}}\frac{\mathrm{E}[\tilde{\lambda}\tilde{y}_{t^n}^{ec**}]}{\mathrm{E}[\tilde{\lambda}]}\mathrm{d}G^w(w), \tag{42}$$

$$\rho: \quad \int_{\underline{w}}^{\overline{w}}\beta^w\mathrm{d}G^w(w) = 1 + r^b, \tag{43}$$

$$\pi^m: \quad \lambda^{r^m} = \int_{\underline{w}}^{\overline{w}}\mathrm{E}_{\mathscr{E}}[U_P^{r^m}]\mathrm{d}G^w(w), \tag{44}$$

$$\kappa^m: \quad \frac{\lambda^{r^m}}{\mathrm{E}[\tilde{\lambda}]} = \int_{\underline{w}}^{\overline{w}}\left(\frac{\mathrm{E}_{\mathscr{E}}[\tilde{U}_2^{r^m}]}{\mathrm{E}[\tilde{U}_2]} + t^n y_{\kappa^m}^{nc**} + t^e\frac{\mathrm{E}[\tilde{\lambda}\tilde{y}_{\kappa^m}^{ec**}]}{\mathrm{E}[\tilde{\lambda}]}\right)\mathrm{d}G^w(w), \tag{45}$$

with definitions, at any skill level $w \in [\underline{w}, \overline{w}]$:

$$\beta^w \equiv \frac{\mathrm{E}[\tilde{U}_1]}{\mathrm{E}[\tilde{\lambda}]} - (1 + r^b)t_z^\ell \cdot z_\rho - t^n y_\rho^n - t^e\frac{\mathrm{E}[\tilde{\lambda}\tilde{y}_\rho^e]}{\mathrm{E}[\tilde{\lambda}]}, \tag{46}$$

$$\mathscr{W}^w \equiv (1 + r^b)t_z^\ell \cdot z_{\sigma^z}^* + t^n y_{\sigma^z}^{n*} + t^e\frac{\mathrm{E}[\tilde{\lambda}\tilde{y}_{\sigma^z}^{e*}]}{\mathrm{E}[\tilde{\lambda}]},$$

where a superscript r^m refers to the value of a variable in a particular realization of the market return. The term β^w is the marginal social utility of income for an individual of skill w, following Diamond [1975]. The term $\mathscr{W}^w$ contains the effects on the government budget, in terms of social welfare, of a change to the marginal tax rate on labour income.

The different marginal policy reforms may have effects on any of the tax bases. This explains the additional terms in the first-order conditions. Some of these additional effects

---

[7]The monotonicity assumption is not strictly needed here, neither for the individual second-order conditions to be fulfilled, nor for our derivations to be correct. In the standard Mirrlees model, since the individual first-order conditions are equivalent to the Spence-Mirrlees conditions, monotonicity $\mathrm{d}z^w/\mathrm{d}w > 0$ would indeed be required. In the present model though, where the individual has more degrees of choice, it suffices that the bordered Hessian of his constrained optimization problem is negative semi-definite, which can be the case even when z^w is not monotonically increasing. Furthermore, the reader can verify in the appendix that we only change variables when we determine averages over the entire population. This is in contrast e.g. to the more heuristic derivations of Saez [2001]. To avoid the notational complexities that would occur though when individuals of different abilities pool at the same level of labour income, we stick to the more traditional monotonicity assumption.

have been cancelled out by substituting the optimal-tax equation for labour income, with the remaining terms being conditional on labour income. This is shown in more detail in appendix D. The first-order conditions also are also impacted by the uncertainty of government revenues, as is evident e.g. from the balancing between private consumption risk and public revenue risk in equation (41) (discussed further in subsection 4.3), and from the presence of terms of the form $\mathrm{E}[\tilde{\lambda}\tilde{y}_\nu^e]/\mathrm{E}[\tilde{\lambda}]$. The stochastic budget multiplier $\tilde{\lambda}$ cannot be divided away in those terms, because it is correlated with excess capital incomes. We discuss this in more detail in the next subsection.

4.1. Marginal Social Risk Premiums

To interpret the government first-order conditions, it is useful to notice first that for any perturbation to a policy reform parameter ν, the term indicating the social welfare effects of the resulting change in tax revenues from excess returns can be decomposed into an expectation term and a risk term, as is shown in the following lemma.

Lemma 6. *The social-welfare effect of a reform to a policy instrument ν can be split into an effect on expected tax revenues, and an effect on the uncertainty of government revenues:*

$$\forall \nu = \sigma^z, t^n, t^e, \rho, \kappa^m : \ t^e \frac{\mathrm{E}[\tilde{\lambda}\tilde{y}_\nu^e]}{\mathrm{E}[\tilde{\lambda}]} = t^e \cdot \left(\mathrm{E}[\tilde{y}_\nu^e] + \frac{cov(\tilde{\lambda}, \tilde{y}_\nu^e)}{\mathrm{E}[\tilde{\lambda}]} \right).$$

When evaluating the effect on tax revenues from excess returns, of an increased reliance on a policy instrument, one should evaluate not only the expected change. Because this change is uncertain, an optimizing government also demands a risk premium. The *marginal social risk premium* associated to the instrument ν is captured in the covariance term. It is similar to the private marginal risk premium terms in individual first-order conditions (12)-(13).

We will encounter similar marginal social risk premium terms throughout this section. The following lemma will be useful.

Lemma 7. *The effect of any policy reform on government revenue from excess returns, net of the social welfare effect of its uncertainty, is determined by the relative shift towards or away from private investment:*

$$\forall \nu = \sigma^z, t^n, t^e, \rho, \kappa^m :$$
$$t^e \int_{\underline{w}}^{\overline{w}} \left(\mathrm{E}[\tilde{y}_\nu^e] + \frac{cov(\tilde{\lambda}, \tilde{y}_\nu^e)}{\mathrm{E}[\tilde{\lambda}]} \right) \mathrm{d}G^\mathrm{w}(w)$$
$$= -t^e \int_{\underline{w}}^{\overline{w}} (\alpha_p - r^b) p \, \mathrm{d}G^\mathrm{w}(w)$$
$$\cdot \left(\frac{\int_{\underline{w}}^{\overline{w}} (\alpha_p - r^b) p \xi_\nu^p \mathrm{d}G^\mathrm{w}(w)}{\int_{\underline{w}}^{\overline{w}} (\alpha_p - r^b) p \, \mathrm{d}G^\mathrm{w}(w)} - \frac{\int_{\underline{w}}^{\overline{w}} \mathrm{E}[\tilde{r}^m - r^b] f \xi_\nu^f \mathrm{d}G^\mathrm{w}(w)}{\int_{\underline{w}}^{\overline{w}} \mathrm{E}[\tilde{r}^m - r^b] f \, \mathrm{d}G^\mathrm{w}(w)} \right).$$

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

Proof. Use definition (15) and the fact that the private shocks are i.i.d. to find:

$$\text{cov}\left(\tilde{\lambda}, \tilde{y}^e\right) = \text{cov}\left(\tilde{\lambda}, \tilde{r}^m\right) \cdot f.$$

Substitute this and integrate over the population:

$$\forall \, \nu = \sigma^z, t^n, t^e, \rho, \kappa^m :$$
$$\int_{\underline{w}}^{\overline{w}} \frac{\text{cov}\left(\tilde{\lambda}, \tilde{y}_\nu^e\right)}{\text{E}\left[\tilde{\lambda}\right]} dG^{\text{w}}(w) = \int_{\underline{w}}^{\overline{w}} \frac{\text{cov}\left(\tilde{\lambda}, \tilde{y}^e\right)}{\text{E}\left[\tilde{\lambda}\right]} dG^{\text{w}}(w) \cdot \frac{\partial}{\partial \nu}\left(\log \int_{\underline{w}}^{\overline{w}} f \, dG^{\text{w}}(w)\right).$$

Substitute condition (41) and use first-order conditions (12)-(13) to find the result. $\qquad\square$

In the mean-variance framework, applying lemma 5, it follows that the risk-compensated effect is zero when returns to private investment are linear.

Lemma 8. *When the mean-variance framework applies and returns to private investment are linear, the budget effect of a policy reform on tax revenue from excess returns is exactly nullified by the social-welfare effect its uncertainty:*

$$\alpha_{pp} = 0 \Longleftrightarrow t^e \frac{\text{E}\left[\tilde{\lambda}\tilde{y}_\nu^e\right]}{\text{E}\left[\tilde{\lambda}\right]} = 0.$$

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

We will see in subsection 4.3 that condition (41) implies that the uncertainty of government revenues is to be balanced against private consumption risk. If a reform to policy instrument ν does not affect the risk properties of the individual portfolios, thus the proportions invested in each asset remain constant, then in our model the budget effect of a policy reform on tax revenue from excess returns is exactly nullified by the social-welfare effect of its uncertainty. On the contrary, a shift towards investment in the market asset, increasing aggregate risk, negatively affects the revenue effect taken into account by the government. The semi-elasticities of investment in each asset are weighted more heavily for individuals whose excess returns stem to a greater extent from that particular asset. The direction of the change, towards private investment or market investment, depends on the investment technology, on preferences, and on the policy instrument in question.

4.2. Optimal Linear Tax on Risk-Free Capital Income

Condition (42) is a multi-person Ramsey-style proportional-reduction condition. Using lemma 7 we can rewrite it:

$$t^n \int_{\underline{w}}^{\overline{w}} y_{t^n}^{nc**} dG^{\text{w}}(w)$$
$$= t^e \int_{\underline{w}}^{\overline{w}} \left(\alpha_p - r^b\right) p \, dG^{\text{w}}(w)$$
$$\cdot \left(\frac{\int_{\underline{w}}^{\overline{w}} \left(\alpha_p - r^b\right) p \xi_{t^n}^{pc**} dG^{\text{w}}(w)}{\int_{\underline{w}}^{\overline{w}} \left(\alpha_p - r^b\right) p \, dG^{\text{w}}(w)} - \frac{\int_{\underline{w}}^{\overline{w}} \text{E}\left[\tilde{r}^m - r^b\right] f \xi_{t^n}^{fc**} dG^{\text{w}}(w)}{\int_{\underline{w}}^{\overline{w}} \text{E}\left[\tilde{r}^m - r^b\right] f \, dG^{\text{w}}(w)}\right). \tag{47}$$

The sign of the optimal tax on risk-free returns depends on how the relative shares of aggregate risk and idiosyncratic risk respond to a reform to the tax on risk-free returns, with the responses weighted more heavily for individuals whose excess returns rely more heavily on the respective assets.

To learn more about this equation, let us assume that the mean-variance framework applies. We then see in the following lemma that the sign of the optimal tax on risk-free returns depends on whether those with more wealth have higher or lower semi-elasticities with respect to the tax rate on risk-free returns. Given that those with more wealth have more avoidance opportunities, we will assume that they have larger elasticities: $\mathrm{d}\xi_{t^n}^{pc**}/\mathrm{d}p > 0$. The following lemma then shows how the sign of the first-order condition is affected by the investment technology.

Lemma 9. *When the mean-variance framework applies, and individuals with higher wealth have higher semi-elasticities of investment with respect to the tax on risk-free returns, a tax increase on risk-free returns shifts investment towards private assets when returns to scale are increasing, and vice versa:*

$$\alpha_{pp} \gtreqless 0 \Longleftrightarrow \frac{\int_{\underline{w}}^{\overline{w}} (\alpha_p - r^b) p \xi_{t^n}^{pc**} \mathrm{d}G^{\mathrm{w}}(w)}{\int_{\underline{w}}^{\overline{w}} (\alpha_p - r^b) p \,\mathrm{d}G^{\mathrm{w}}(w)} \gtreqless \frac{\int_{\underline{w}}^{\overline{w}} \mathrm{E}\big[\tilde{r}^m - r^b\big] f \xi_{t^n}^{fc**} \mathrm{d}G^{\mathrm{w}}(w)}{\int_{\underline{w}}^{\overline{w}} \mathrm{E}\big[\tilde{r}^m - r^b\big] f \,\mathrm{d}G^{\mathrm{w}}(w)}.$$

Proof. Use lemmas 4 and 5 to find:

$$\frac{\int_{\underline{w}}^{\overline{w}} f \xi_{t^n}^{fc**} \mathrm{d}G^{\mathrm{w}}(w)}{\int_{\underline{w}}^{\overline{w}} f \,\mathrm{d}G^{\mathrm{w}}(w)} = \frac{\int_{\underline{w}}^{\overline{w}} \frac{p}{\alpha_p - r^b} \xi_{t^n}^{pc**} \left(1 - \frac{\alpha_{pp} p}{\alpha_p - r^b}\right) \mathrm{d}G^{\mathrm{w}}(w)}{\int_{\underline{w}}^{\overline{w}} \frac{p}{\alpha_p - r^b} \,\mathrm{d}G^{\mathrm{w}}(w)}.$$

It then follows:

$$\alpha_{pp} \gtreqless 0 \Longleftrightarrow \frac{\int_{\underline{w}}^{\overline{w}} f \xi_{t^n}^{fc**} \mathrm{d}G^{\mathrm{w}}(w)}{\int_{\underline{w}}^{\overline{w}} f \,\mathrm{d}G^{\mathrm{w}}(w)} \lesseqgtr \frac{\int_{\underline{w}}^{\overline{w}} \frac{p}{\alpha_p - r^b} \xi_{t^n}^{pc**} \mathrm{d}G^{\mathrm{w}}(w)}{\int_{\underline{w}}^{\overline{w}} \frac{p}{\alpha_p - r^b} \,\mathrm{d}G^{\mathrm{w}}(w)} \lesseqgtr \frac{\int_{\underline{w}}^{\overline{w}} (\alpha_p - r^b) p \xi_{t^n}^{pc**} \mathrm{d}G^{\mathrm{w}}(w)}{\int_{\underline{w}}^{\overline{w}} (\alpha_p - r^b) p \,\mathrm{d}G^{\mathrm{w}}(w)},$$

where we used the assumption $\mathrm{d}\xi_{t^n}^{pc**}/\mathrm{d}p > 0$ in the last step. $\square$

It then follows from condition (47) and corollary 8 that in the mean-variance framework with linear returns to private investment, the optimal tax on risk-free returns to capital is zero. When returns to private investment are non-linear, the sign of the optimal tax on risk-free returns depends on the sign of t^e, which we will see in the next subsection should never be negative, and on whether returns to private investment are increasing or decreasing. These findings are summarized in the following proposition.

Proposition 1. *The optimal tax on risk-free capital returns is zero when the mean-variance framework applies and marginal rates of return to private investment are constant:*

$$\alpha_{pp} = 0 \Longrightarrow t^n = 0.$$

The risk-free returns should be subsidized when returns are increasing, and taxed when returns are decreasing:

$$\alpha_{pp} \gtreqless 0 \Longleftrightarrow t^n \lesseqgtr 0.$$

This proposition applies regardless of whether the risk is returned to individuals by the public good or a lump sum.

Note finally that without private assets, $p = 0$, the optimal tax rate is $t^n = 0$. This extends the conclusion of Schindler [2008] who derived this result for the case of a representative household model. Without aggregate risk, $f = 0$, the optimal tax rate is also $t^n = 0$. When the Mean-Variance framework does not apply and there are both aggregate and idiosyncratic risks, the sign of the optimal tax on risk-free returns is unclear without further information; it is determined by equation (47). These results are summarized in table 1.

Table 1: Optimal tax on risk-free returns

Risk type	Mean-Variance framework applies?	
	yes	no
Aggregate only	$t^n = 0$	
Mixed	$\alpha_{pp} \gtreqless 0 \Longleftrightarrow t^n \lesseqgtr 0$	unclear (eq. 47)
Idiosyncratic only	$t^n = 0$	

4.3. Optimal Linear Tax on Excess Capital Income

The government itself cannot bear any of the unpooled risks in the economy: the uncertainty of its revenues somehow needs to be borne by individuals. Condition (41) on the choice of t^e shows how the government balances the risk of private consumption against the uncertainty of public revenues. When there is an increase in the tax rate t^e on excess returns, this reduces the riskiness of private consumption, but there is an increase in the riskiness of public revenues. How this is done depends on how the uncertainty is returned to the individuals. We focus on two channels by which this occurs. After the realization of the market rate of return, the government balances the budget using either an uncertain provision of a public good as in Christiansen [1993], or using a stochastic lump sum as in Gordon [1985].

The fact that our model has only two periods affects our results. Neither the government, nor the individuals are able to smooth the impact of the aggregate shocks towards the future. Our results are relevant since it is often impossible to fully pool the risks over time, e.g. due to liquidity constraints or due to the permanent nature of the aggregate shocks.

We analyze the consequences for the optimal tax on excess capital income of both possibilities in turn, and then we study the optimal choice between the two instruments.

4.3.1. Returning Risk through Uncertain Provision of a Public Good

Suppose that the government can only return the risk of its revenues using uncertain provision of the public good. The lump sums are thus zero, $K = 0$, and we disregard condition (45). If individual preferences are concave in the public good, uncertainty in its provision will have a welfare cost. In choosing the tax rate on excess capital income, the government

balances the risk in second-period consumption, the right-hand side of (41), against the riskiness of its revenues, the left-hand side. When the government increases the tax on excess capital income, it reduces the variance of second-period disposable income, but it increases the variance of tax revenues.[8] We can use this result to show that t^e should be strictly between zero and unity.

Consider first the case where $t^e = 0$. If there is a high realization of the market rate of return, excess returns $y^{e,w}$ will be high, and thus second-period consumption $c^{2,w}$ will be high. The marginal utility of second-period consumption U_2^w will be low. Conversely, when there is a low realization of the market rate of return, the marginal utility of second-period consumption will be high. The covariance on the right-hand side of (41) will thus be negative. The left-hand side though will be zero, since government revenues will not be stochastic without a tax on excess returns. Therefore we can rule out $t^e = 0$.

Similarly, we can exclude the case $t^e < 0$, because in this case the covariance on the left-hand side will be positive: a high realization of the market rate of return will lead to low government revenues and to a higher value of λ. Similar reasoning shows that the only remaining possibility, a positive tax on excess returns $t^e > 0$, yields the correct signs for both covariances.

We can confirm that condition (41) will be fulfilled for a value of t^e strictly below 100%. Suppose thus that t^e approaches unity from below. The covariance on the right-hand side of (41) will go to zero, as all excess returns are taxed away, and in the present context the risk is not returned by the government as a lump sum. The covariance on the left-hand side though will increase in absolute value: the government budget responds more to market shocks when taxes on excess returns are high. As long as individuals are averse towards risk in the provision of the public good, the optimum will be strictly smaller then 100%. We summarize these findings in the following proposition.

Proposition 2. *If all risk is returned using uncertain provision of a public good, then the optimal tax rate on excess capital returns is strictly positive, and strictly smaller than 100%:*

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

Not all risk in individual excess returns increases the variance of public good provision. Returns from private investment are risky but idiosyncratic, so the variance of individual returns evens out over the population. Returns from investments in market funds on the other hand are subject to aggregate risk, thus impacting the variance of government revenues. The relative importance of either risk type has an impact on the covariance on the left-hand side of equation (41). To see this, we study the limit cases with only idiosyncratic or only aggregate risk.

First suppose that there are no private investment opportunities available, such that $\forall w : p^w = 0$. Condition (41) can then be rewritten as in the following proposition.

---

[8]Mintz [1981] shows how with decreasing returns to scale, a tax increase on excess returns has a negative impact on second-period consumption, and with decreasing absolute risk aversion this has a negative impact on the amount invested in risky assets. We assume that this effect is insufficiently strong to outweigh both Domar-Musgrave effects and the mechanical effect of the tax reform. The tax increase thus increases the variance of tax revenues.

Proposition 3. *When there is no private investment opportunity so there is only aggregate risk, the optimality condition for taxes on excess returns becomes:*

$$\forall w \in \left[\underline{w}, \overline{w} \right] : \frac{\int_{\underline{w}}^{\overline{w}} \mathrm{E}\left[\tilde{U}_P \right] \cdot \frac{\mathrm{cov}\left(\tilde{U}_P, \tilde{r}^m \right)}{\mathrm{E}\left[\tilde{U}_P \right]} \mathrm{d}G^{\mathsf{w}}(\hat{w})}{\int_{\underline{w}}^{\overline{w}} \mathrm{E}\left[\tilde{U}_P \right] \mathrm{d}G^{\mathsf{w}}(\hat{w})} = \frac{\mathrm{cov}\left(\tilde{U}_2^w, \tilde{r}^m \right)}{\mathrm{E}\left[\tilde{U}_2^w \right]}.$$

Proof. In the case that $p = 0$, condition (41) reduces to $\mathrm{E}\left[\tilde{\lambda} \cdot \left(\tilde{r}^m - r^b \right) \right] = 0$ using lemma 7. Substitute into this result (44) to find:

$$\int_{\underline{w}}^{\overline{w}} \mathrm{E}\left[\tilde{U}_P \cdot \left(\tilde{r}^m - r^b \right) \right] \mathrm{d}G^{\mathsf{w}}(w) = 0.$$

Rewrite this equation as a covariance and combine with the relevant terms in the individual portfolio conditions (12).[9] □

This shows that with only aggregate risk, the government offers a kind of insurance. When an adverse shock occurs, individual consumption will go down, but this is attenuated by the tax on excess returns. The provision of the public good will decrease instead. Optimal government policies are such that decreases in second-period consumption are balanced against decreases in the provision of the public good, yielding a situation superior to the case where all risk falls on second-period consumption.

The insurance condition in Proposition 3 is equity sensitive. If individuals with a higher marginal utility of public good provision are also the ones who are more risk averse with respect to public good provision, then an inequality-averse government will want to levy a *lower* tax on excess returns, reducing the uncertainty of its revenues. This argument only applies when utility is not additive in the public good. If it were, the marginal utility of public good provision would be equal for all individuals, and equity considerations would become irrelevant.

Instead suppose that the left hand side of (41) goes to zero. This could occur e.g. because the variance of the return to market funds decreases, because the individuals become less risk averse with respect to public good provision (that is, individual utilities approach linearity in the public good), or individuals invest more in private investments. For the optimal tax condition to remain valid, the right-hand side should also go to zero. This is the case when the tax rate on excess returns approaches unity: $t^e \to 1$. The more investments with idiosyncratic risk dominate, or the less risk-averse individuals are with respect to the public good, the more optimal government policies approach full insurance.

We exclude the possibility $t^e = 1$ because conditions (12)–(13) would imply that individuals are indifferent with respect to the composition of their portfolios. Still social welfare improves as the tax on excess returns is brought closer to 100%. In a more complex model the 100% result could be attenuated e.g. by the addition of investment effort or moral hazard. We do not consider these extensions here since they would significantly complicate our derivations. Instead we write $t^e \to 1$ in the optimum.

---

[9]Note that the term $-\mathrm{cov}\left(\tilde{U}_2^w, \tilde{r}^m \right) / \mathrm{E}\left[\tilde{U}_2^w \right] = \tilde{r}^m - r^b$ is the same for all individuals.

When there is both aggregate and idiosyncratic risk, the government strikes a balance between the principle of insurance against aggregate risk, through fluctuations in the provision of the public good, and the principle of insurance against idiosyncratic risk, by taxing away a higher share of the excess returns.

4.3.2. Returning the Risk as Stochastic Lump Sums

Suppose now that there is no public good provision, $P = 0$, and all government risk is returned using stochastic lump sums. To avoid the complexities introduced by the compensated effects of reforms to the stochastic lump sum, we limit ourselves to the mean-variance framework with linear returns to private investment. In this context, the optimality condition for excess returns (41) then implies that the tax on excess returns should approach 100%, unless there is no idiosyncratic risk. This is shown in the following proposition.

Proposition 4. *When the government returns all risk to the individuals as second-period lump sum transfers, in the mean-variance framework with linear returns to capital, the optimal tax rate on excess returns approaches 100%:*

$$t^e \to 1.$$

If there is no private investment risk, either because $p = 0$ or because $\mathrm{var}(\tilde{\varepsilon}) = 0$, then the tax on excess returns becomes irrelevant.

Proof. Suppose $t^e < 1$. Since we assume linearity of returns to private investment within the mean-variance framework, corollary 8 implies $\mathrm{E}\left[\tilde{\lambda}\,\tilde{y}_{\kappa\cdot m}^{ec**}\right] = 0$, and proposition 1 implies $t^n = 0$. The optimality condition for the lump sum (45) then becomes:

$$\frac{\lambda^{R^m}}{\mathrm{E}\left[\tilde{\lambda}\right]} = \int_{\underline{w}}^{\overline{w}} \frac{\mathrm{E}_{\mathscr{E}}\left[\tilde{U}_2^{R^m}\right]}{\mathrm{E}\left[\tilde{U}_2\right]} \mathrm{d}G^{\mathrm{w}}(w). \tag{48}$$

Substituting this into (41) yields the condition:

$$\int_{\underline{w}}^{\overline{w}} \frac{\mathrm{cov}\left(\tilde{U}_2, \tilde{\varepsilon}\right)}{\mathrm{E}\left[\tilde{U}_2\right]} p^w \mathrm{d}G^{\mathrm{w}}(w) = 0. \tag{49}$$

This can only be true if $p = 0$ for all individuals, or if $\mathrm{var}(\tilde{\varepsilon}) = 0$. In all other cases it follows that $t^e \to 1$. □

Let us interpret condition (49) in the situation without idiosyncratic risk ($p^w = 0$ or $\mathrm{var}(\tilde{\varepsilon}) = 0$). The optimality condition then is fulfilled for any value of t^e: the tax on excess returns becomes irrelevant. Since any revenue raised through a change to the tax on excess returns is returned through the lump sum, any change in the variance of private consumption is counteracted as the government balances its budget. The insurance mechanism that drove our results when private risk was balanced against the uncertainty of public good provision no longer applies in the present context. This result remains valid in the presence of risk-free private investments.

The fact that the lump sums are not matched to the individuals does not undo the irrelevance of the tax on excess returns. Any redistributional effects of the reform are absorbed as the other policy instruments are re-optimized, and the final effect on social welfare is zero. This extends the results of Atkinson and Stiglitz [1980], who assume the presence of a representative consumer for whom the tax on excess returns is exactly undone by the stochastic lump sum, and of Gordon [1985], who assumes that the lump sums are tailored to the individuals.

Note that even though the tax on excess returns does not affect social welfare, there is no reason to prefer it as an instrument. The presence of non-zero marginal rates on labour income remains optimal for redistributional purposes, and since the government can meet its budgetary requirements with no compensated effects by changing the intercept of the labour income tax, it is best in the current context not to introduce the additional compliance cost implied by a tax on excess returns. We thus find that when aggregate risk is returned to individuals using lump sum transfers, and there is no idiosyncratic risk, capital taxes are optimally zero.

The situation changes with the introduction of idiosyncratic risk. It becomes optimal to provide full insurance by taxing excess returns at 100%, even when the relative importance of the idiosyncratic risk in the portfolio returns is minor. This full insurance condition is no longer mitigated by a balancing act between public risk and private risk, as was the case when all risk of government revenues was returned through the public good. This extends the results of Gordon [1985], who finds in a comparative statics exercise that the tax on excess returns has no behavioural effects when the lump sums are tailored to the individuals in the non-corporate sector. Despite the absence of behavioural effects at any tax rate on excess returns, from a welfare perspective it would be better to fully insure the idiosyncratic risk. In presence of a non-linear tax on earnings, this is the case even when lump sums are not tailored to the individuals.

4.3.3. Choosing between Stochastic Lump Sums and Public Good Provision

Finally, consider what the government should do if it can choose how to return the risk: using uncertain public good provision, using a stochastic lump sum, or a combination of both. Note first that proposition 4 remains valid. As in the case when there was only a stochastic lump sum and no public good provision, the tax on excess returns is irrelevant when there is no idiosyncratic risk, and it should be 100% when there is idiosyncratic risk. On the other hand, condition (44) remains valid. Substituting this into optimal tax condition (41), and substituting condition (48):

$$\frac{\int_{\underline{w}}^{\overline{w}} E\big[\tilde{U}_P^w\big] \frac{\text{cov}\big(\tilde{U}_P^w, \tilde{r}^m\big)}{E[\tilde{U}_P^w]} dG^w(\hat{w})}{\int_{\underline{w}}^{\overline{w}} E\big[\tilde{U}_P^w\big] dG^w(w)} = \int_{\underline{w}}^{\overline{w}} \frac{\text{cov}\big(\tilde{U}_2^w, \tilde{y}^e\big)}{E\big[\tilde{U}_2^w\big]} dG^w(w) \tag{50}$$

This condition is similar to the one found in proposition 3, only here it is valid even in presence of idiosyncratic risk.

Summarizing these results, we find that when there is no idiosyncratic risk, the tax on excess returns is irrelevant and should be set to zero to avoid the compliance costs of doing

otherwise. In this case government tax revenues become deterministic. It would still deploy a stochastic lump sum though, in order to create stochastic public goods, converting private risk into public risk, satisfying condition (50). As soon as there is some idiosyncratic risk, the tax on excess returns should be used to provide full insurance. The aggregate risk is then returned to the individuals using both the stochastic lump sum and the public good provision, again in such a way that public risk is balanced against private risk as in condition (50). When there is no aggregate risk, public good provision becomes deterministic and follows a Samuelson rule. The lump sum becomes redundant.

4.3.4. Summary

We have found that the optimal tax on excess returns depends on the type of investment risk, and on how the uncertainty of government revenues is returned to the individuals. The different possibilities are summarized in table 2. When there is only idiosyncratic risk, government revenue becomes deterministic, and the tax on excess returns is used to offer full insurance. When there is aggregate risk and the uncertainty of government revenues is returned to the individuals using only a stochastic public good, then the tax on excess returns is strictly positive but less than 100%. The tax on excess returns is set such that the risk of public good provision is balanced against the risk of private consumption, in an equity-sensitive manner.

For the case where the government relies on a stochastic lump sum to return some or all of the risk, we considered only the mean-variance framework with linear returns to private investment. In the case with aggregate risk only, the tax on excess returns becomes irrelevant, and as soon as there is some idiosyncratic risk, it should be used to offer full insurance. If the government can choose between the lump sum and the public good, then these instruments should be set such that private risk is balanced against public risk in an equity sensitive manner.

Note that the choice of instrument to return the risk to the individuals does not affect the optimal tax on risk-free returns, summarized in proposition 1, nor does it affect in a qualitative way the optimal tax on labour income or the optimal public good provision, discussed in the next sections.

Table 2: Optimal tax on excess returns

Risk type	Instrument used to return risk		
	Public good only	Lump sum only**	Both**
Aggregate only	$0 < t^e < 1$*	t^e irrelevant	t^e irrelevant*
Mixed		$t^e \to 1$	$t^e \to 1$*
Idiosyncratic only	$t^e \to 1$		

*Use available instruments to balance private against public risk in an equity-sensitive manner; $t^e \to 1$ if neutral w.r.t. public risk.
** Mean-variance framework with linear returns to private investment

4.4. Optimal Non-Linear Labour Income Tax

Changes in the intercept ρ are equivalent to imposing an equal lump-sum tax or credit on all individuals. The optimality condition for the intercept (43) is quite standard. Note that when the conditions apply for a zero tax on risk-free returns to be optimal, $t^n = 0$, then using lemma 7, corollary 8 and table 1 we find that the marginal social utility of income reduces to the following:

$$\beta^w \equiv \left(1 + r^b\right)\left(\frac{E\left[\tilde{U}_2^w\right]}{E\left[\tilde{\lambda}\right]} - t_z^{\ell,w} \cdot z_\rho^w\right). \tag{51}$$

The optimal marginal labour income tax (40) resembles the standard Mirrlees result, amended to take into account induced capital income tax revenue effects of a labour income tax reform, and the uncertainty involved therein. The standard Mirrlees result can be recovered when $t^n = 0$ is optimal. Again using lemma 7, corollary 8 and table 1, changing variables from skill to labour income, and rearranging, we find the following corollary.

Proposition 5. *When it is optimal not to tax risk-free capital income, $t^n = 0$, then the marginal optimal tax rates on labour income are determined by the classic Mirrlees equation:*

$$\forall z: \frac{t_z^\ell(z)}{1 - t_z^\ell(z)} = \frac{1}{\xi^z} \cdot \frac{1 - G^z(z)}{z g^z(z)} \cdot \frac{\int_z^{\overline{z}}\left(1 + r - \beta^{\hat{w}}\right)dG^z\left(z^{\hat{w}}\right)}{1 - G^z(z^z)},$$

with present value compensated elasticity of labour income:

$$\xi^z \equiv -\frac{1}{1 + r^b}\frac{1 - t_z^\ell(z)}{z}z_{\sigma^z}^*,$$

where β^w is given by (51).

With increasing returns to scale, it is not clear a priori whether the optimal tax rates on labour income will be lower or higher. Since individuals with higher skills have more savings, in expectation they will obtain higher rates of return, yielding a higher life-cycle level of consumption and thus a reason for additional redistribution. Remember though from Proposition 3 that the tax on excess returns was positive for reasons of insurance, rather than for reasons of redistribution. Similarly, the optimality condition for the tax on risk-free returns (42) is not directly impacted by redistributional concerns. Even so, these taxes do have a distributional impact. It is not a priori clear whether, for each specific skill level, the additional distributional advantage from levying a higher marginal labour tax, is exhausted by the introduction of non-zero tax rates on capital income. Similar reasoning shows that the same is true with decreasing returns to scale.

Furthermore, still with non-constant returns to scale, as the tax rates on capital income become non-zero, a reform in the labour income tax causes additional distortions through the capital income tax bases. The additional terms in the optimality condition for the tax on labour supply (40) show that the net impact is not a priori clear. A change in the amount saved has an impact on the subsidy for risk-free capital income, but also an impact of opposing sign via the tax levied on excess capital income. It is not clear which effect prevails.

4.5. Optimal Public Good Provision

From condition (44) it follows that for each outcome r^m, the level of public good is chosen such the sum of social marginal benefits to all individuals equals the shadow price of government revenue. Note that if individual utility were linear in the public good, U_P would be the same for all individuals for any value of public good provision. The government budget multiplier λ would in this case cease to stochastic.

Take the expected value of this expression:

$$\mathrm{E}\left[\tilde{\lambda}\right] = \int_{\underline{w}}^{\overline{w}} \mathrm{E}\left[\tilde{U}_P\right] \mathrm{d}G^{\mathrm{w}}(w) > 0, \tag{52}$$

To relate this to a Samuelson-type rule for public goods provision, we adopt the approach of Christiansen [1981]. Note that equation (52) is the average over all possible market returns r^m of the effect of marginally increasing the public good provision at such realizations. The expected marginal valuation of the public good by an individual in terms of first-period consumption is $\mathrm{E}\left[\tilde{U}_P^w\right]/\mathrm{E}\left[\tilde{U}_1^w\right]$. The expected effect (52) of the reforms to the public good provision can be split into the effect of the same reform if the social weights of all individuals, $\mathrm{E}\left[\tilde{U}_1^w\right]/\mathrm{E}\left[\tilde{\lambda}\right]$, were equal to $1 + r^b$, and the distributional effects caused by the fact that the true social weights are different:

$$1 = \int_{\underline{w}}^{\overline{w}} \left(1 + r^b\right) \frac{\mathrm{E}\left[\tilde{U}_P\right]}{\mathrm{E}\left[\tilde{U}_1\right]} \mathrm{d}G^{\mathrm{w}}(w) + \int_{\underline{w}}^{\overline{w}} \left(\frac{\mathrm{E}\left[\tilde{U}_1\right]}{\mathrm{E}\left[\tilde{\lambda}\right]} - \left(1 + r^b\right)\right) \cdot \frac{\mathrm{E}\left[\tilde{U}_P\right]}{\mathrm{E}\left[\tilde{U}_1\right]} \mathrm{d}G^{\mathrm{w}}(w), \tag{53}$$

where we wrote the welfare effects in monetary terms by dividing by $\mathrm{E}\left[\tilde{\lambda}\right]$. The point now is that any reform with the particular distributional effects captured in the second term on the right-hand side could be replicated by an appropriate reform of the labour income tax. Since the labour income tax is assumed to be optimal, these distributional gains should be exactly equal to the costs of achieving them. In appendix D.5 we show how this leads to the following expression for the optimal public provision of the public good.

Proposition 6. *The optimal provision of the public good is determined by an equity-sensitive version of the Samuelson condition:*

$$1 = \left(1 + r^b\right) \int_{\underline{w}}^{\overline{w}} \frac{\mathrm{E}\left[\tilde{U}_P\right]}{\mathrm{E}\left[\tilde{U}_1\right]} \mathrm{d}G^{\mathrm{w}}(w)$$

$$+ \int_{\underline{z}}^{\overline{z}} \frac{\mathrm{E}\left[\tilde{U}_P\right]}{\mathrm{E}\left[\tilde{U}_1\right]} \cdot \left(\left(1 + r^b\right) t_z^\ell z_\rho + t^n y_\rho^n + t^e \frac{\mathrm{E}\left[\tilde{\lambda} \tilde{y}_\rho^e\right]}{\mathrm{E}\left[\tilde{\lambda}\right]}\right) \mathrm{d}G^{\mathrm{z}}(z)$$

$$+ \int_{\underline{z}}^{\overline{z}} \frac{\mathrm{d}}{\mathrm{d}z} \left(\frac{\mathrm{E}\left[\tilde{U}_P\right]}{\mathrm{E}\left[\tilde{U}_1\right]}\right) \left(\left(1 + r^b\right) t_z^\ell z_{\sigma^z}^* + t^n y_{\sigma^z}^{n*} + t^e \frac{\mathrm{E}\left[\tilde{\lambda} \tilde{y}_{\sigma^z}^{e*}\right]}{\mathrm{E}\left[\tilde{\lambda}\right]}\right) \mathrm{d}G^{\mathrm{z}}(z).$$

This extends the equity-sensitive version of the Samuelson condition found by Christiansen [1981] to our model with stochastic shocks. The second line contains the expected income effects associated with an increased labour tax liability of $\mathrm{E}\left[\tilde{U}_P\right]/\mathrm{E}\left[\tilde{U}_1\right]$ at each skill

level, amended to take into account the uncertainty involved. Note that the term $\mathrm{E}\big[\tilde{U}_P\big]/\mathrm{E}\big[\tilde{U}_1\big]$ is not the same at all income levels; the labour tax reform thus also causes compensated effects. These are captured by the third line.

When it is optimal not to tax risk-free capital income, we can find a reduced expression by using lemma 7, corollary 8 and table 1.

Corollary 1. *When it is optimal not to tax risk-free capital income, $t^n = 0$, then the optimality condition for public expenditure reduces to:*

$$1 = \int_{\underline{w}}^{\overline{w}} \frac{\mathrm{E}\big[\tilde{U}_P\big]}{\mathrm{E}\big[\tilde{U}_2\big]} \mathrm{d}G^{\mathrm{w}}(w) + \int_{\underline{z}}^{\overline{z}} t_z^{\ell} \cdot \left(\frac{\mathrm{E}\big[\tilde{U}_P\big]}{\mathrm{E}\big[\tilde{U}_2\big]} \cdot z_{\rho} + \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{\mathrm{E}\big[\tilde{U}_P\big]}{\mathrm{E}\big[\tilde{U}_2\big]} \right) \cdot z_{\sigma^z}^* \right) \mathrm{d}G^{\mathrm{z}}(z).$$

As before, the first integral corresponds to the traditional Samuelson condition, while the second integral corrects for the fact that different individuals have different social welfare weights. The latter redistributional effects of the marginal reform of the public good provision are again replicated by local reforms of the labour income tax of size $\mathrm{E}\big[\tilde{U}_P\big]/\mathrm{E}\big[\tilde{U}_2\big]$.

4.6. Optimality of the RRA

An RRA is defined as a zero-tax on risk-free capital income combined with a positive tax on excess returns. Table 3 summarizes the cases in our model where such policy is optimal.

With only aggregate risk or only idiosyncratic risk we found a zero-tax on risk-free capital income to be optimal. When there is only idiosyncratic risk, the optimal tax on excess returns is positive, providing insurance, and thus the RRA is optimal. When there is only aggregate risk, and all uncertainty of government revenues is returned using a stochastic public good, the optimal tax on excess returns again is positive, now spreading the risk between private and public consumption. Again the RRA is optimal. However as soon as the government disposes over a stochastic lump sum to return the aggregate risk, the tax on excess returns becomes irrelevant and the RRA is no longer optimal.

In the case with mixed risk types the optimality of a zero-tax on risk-free returns depends on the applicability of the mean-variance framework and on private investment technology being linear. If that is case, since the tax on excess returns is positive in order to insure against the idiosyncratic risk, the RRA is also optimal.

Table 3: Optimality of the RRA

Risk type	Instrument used to return risk		
	Public good only	Lump sum only	Both
Aggregate only	yes	no	
Mixed	yes in mean-variance framework with linear returns		
Idiosyncratic only	yes		

5. Conclusion

We have studied optimal nonlinear income labour taxation and optimal linear capital income taxation 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 assets with idiosyncratic risk. Aggregate risk is reflected in the government's revenue stream, which in turn leads to uncertainty in the quantity of public good provided.

In the special case where the mean-variance framework applies and returns to private investment satisfy constant returns to scale (so do not include rents), RRA taxation as proposed by the Mirrlees Review [2011] is optimal. That is, the tax rate applied to the risky components of capital income — both aggregate and idiosyncratic — is positive, while risk-free capital income should not be taxed. Nonlinear taxation of labour income takes the standard form. If there are returns to scale in private investments, this simple RRA system no longer applies. Returns on risk-free capital income will tend to be taxed with increasing returns to scale, and subsidized with decreasing returns. Marginal labour income tax rates will take into account capital income tax revenue changes induced by changes in the marginal income tax rate. Public goods decision rules are standard. If redistribution terms are separated out though, as in Christiansen [1981], they will also take into consideration changes in capital and labour income tax revenues induced by changes in public good supply.

The optimality rules for the tax on excess returns depend on how aggregate risk is returned to the tax payers. In absence of aggregate risk, when there is only idiosyncratic risk, optimal government policy approaches pure insurance – although in more complex models incorporating investment effort or moral hazard this result should likely be mitigated. If aggregate risk is returned using only a stochastic public good, the risk of private consumption is balanced against the risk of public consumption, in an equity-sensitive fashion, providing a different kind of insurance. As the importance of idiosyncratic risk increases, emphasis on pure insurance will grow. As soon as the government has a stochastic lump sum at its disposal to return the risk, either using it on its own or in combination with the public good, the tax on excess returns becomes irrelevant if there is only aggregate risk, and its optimal value approaches 100% as soon as there is some idiosyncratic risk. When the government has both instruments at its disposal, they are chosen in such a way that the risk of private consumption is balanced against the risk of public consumption, again sensitive to equity considerations.

6. Acknowledgements

References

Alberto Alesina, Omar Barbiero, Carlo Favero, Francesco Giavazzi, and Matteo Paradisi. Austerity in 2009–13. *Economic Policy*, 30(83):383–437, 2015.

A.B. Atkinson and J.E. Stiglitz. *Lectures on public economics*. McGraw-Hill College, 1980.

J. Banks and P. Diamond. The base for direct taxation. *Institute for Fiscal Studies, ed., Dimensions of Tax Design, New York: Oxford University Press*, 2010.

Wolfgang Buchholz and Kai Konrad. Taxes on risky returns – an update. Technical report, Max Planck Institute for Tax Law and Public Finance, 2014.

Howard Chernick, Adam Langley, and Andrew Reschovsky. The impact of the great recession and the housing crisis on the financing of america's largest cities. *Regional Science and Urban Economics*, 41(4):372–381, 2011.

Vidar Christiansen. Evaluation of public projects under optimal taxation. *The Review of Economic Studies*, 48(3):447–457, 1981.

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.

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. NBER Working Paper w22822, National Bureau of Economic Research, November 2016.

Irwin Friend and Marshall E Blume. The demand for risky assets. *The American Economic Review*, pages 900–922, 1975.

Aart Gerritsen, Bas Jacobs, Alexandra Rusu, and Kevin Spiritus. Optimal taxation of capital income when individuals have different returns. mimeo, Max Planck Institute Munich/Erasmus University Rotterdam/KU Leuven, 2016.

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

Marcin Kacperczyk, Stijn Van Nieuwerburgh, and Lauraar Veldkamp. A rational theory of mutual funds' attention allocation. *Econometrica*, 84(2):571–626, 2016.

Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.

Jack M Mintz. Some additional results on investment, risk taking, and full loss offset corporate taxation with interest deductibility. *The Quarterly Journal of Economics*, pages 631–642, 1981.

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

James Mirrlees, Stuart Adam, Tim Besley, Richard Blundell, Steve Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba. *Tax By Design: The Mirrlees Review*. Oxford University Press, Sep 2011.

Jan Mossin. Taxation and risk-taking: an expected utility approach. *Economica*, 35(137): 74–82, 1968.

Thomas Piketty. *Capital in the twenty-first century*. Belknap Press of Harvard University Press Cambridge, MA, 2014.

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

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68(1):205–29, January 2001.

Dirk Schindler. Taxing risky capital income—a commodity taxation approach. *FinanzArchiv: Public Finance Analysis*, 64(3):311–333, 2008.

Mark Skidmore and Eric Scorsone. Causes and consequences of fiscal stress in michigan cities. *Regional Science and Urban Economics*, 41(4):360–371, 2011.

Joseph E Stiglitz. The effects of income, wealth, and capital gains taxation on risk-taking. *The Quarterly Journal of Economics*, pages 263–283, 1969.

A. Slutsky Equations with Second-Period Compensation

A.1. Second-Period Compensated Effects of a First-Period Intercept Reform

Rewrite the utility function in terms of labour income z, savings a, investment in market funds f and investment in private investment opportunities p, conditional on a reform of size ρ to the first-period tax intercept:

$$\tilde{\mathcal{U}}\left(z, a, f, p | \rho\right) \equiv U\left(\hat{c}^1\left(z, a | \rho\right), \hat{\tilde{c}}^2\left(a, f, p | \rho\right), \frac{z}{w}, P\right),$$

with:

$$\hat{c}^1\left(z,a|\rho\right) \equiv z - t^\ell\left(z\right) - \left(a+\rho\right),$$
$$\hat{c}^2\left(a,f,p|\rho\right) \equiv \left\{1 + r^b \cdot \left(1-t^n\right)\right\} \cdot \left(a+\rho\right) + \left(1-t^e\right) \cdot \left[\tilde{y}^e\left(f,p\right)\right],$$
$$\tilde{y}^{e,w}\left(f,p\right) = f \cdot \tilde{r}^m + \tilde{F}\left(p\right) - r^b \cdot \left(f+p\right).$$

Note that the reform to the first-period intercept ρ is compensated in the second period. To see that the utility effect of the reform is indeed exactly compensated, first note the first-order conditions:

$$0 = \mathrm{E}\left[\tilde{\mathscr{U}}_z\right] = \mathrm{E}\left[\tilde{U}_1\right] \cdot \left(1-t_z^\ell\right) + \frac{\mathrm{E}\left[\tilde{U}_\ell\right]}{w},$$
$$0 = \mathrm{E}\left[\tilde{\mathscr{U}}_a\right] = -\mathrm{E}\left[\tilde{U}_1\right] + \mathrm{E}\left[\tilde{U}_2\right] \cdot \left\{1 + r^b \cdot \left(1-t^n\right)\right\},$$
$$0 = \mathrm{E}\left[\tilde{\mathscr{U}}_f\right] = \left(1-t^e\right) \cdot \mathrm{E}\left[\tilde{U}_2 \cdot \left(\tilde{r}^m - r^b\right)\right],$$
$$0 = \mathrm{E}\left[\tilde{\mathscr{U}}_p\right] = \left(1-t^e\right) \cdot \mathrm{E}\left[\tilde{U}_2 \cdot \left(\tilde{F}_p - r^b\right)\right].$$

Apply the envelope theorem and use the first-order condition for a to find $E\left[\tilde{\mathscr{U}}_\rho\right] = 0$. The reform to the first-period tax intercept is indeed fully compensated in the second period.

When there is an exogenous change in the intercept ρ, individuals will alter their behaviour such that the first-order conditions remain satisfied:

$$\forall \gamma = z,a,f,p : \frac{\mathrm{d}}{\mathrm{d}\rho}\left(\mathrm{E}\left[\tilde{\mathscr{U}}_\gamma\right]\right) = 0.$$

Expand this total differential to find a system of equations for the second-period compensated effects of a reform to the first-period tax intercept:

$$\forall \gamma = z,a,f,p : \sum_{\eta \in \{z,a,f,p\}} \mathrm{E}\left[\tilde{\mathscr{U}}_{\eta\gamma}\right] \cdot \gamma_\rho^{**} = -\mathrm{E}\left[\tilde{\mathscr{U}}_{\rho\gamma}\right]. \tag{54}$$

Note that the following partial derivatives are equal:

$$\forall \gamma = z,a,f,p : \mathrm{E}\left[\tilde{\mathscr{U}}_{a\gamma}\right] = \mathrm{E}\left[\tilde{\mathscr{U}}_{\rho\gamma}\right].$$

The system (54) is thus solved by the following second-period compensated effects:

$$a_\rho^{**} = -1; \tag{55}$$
$$z_\rho^{**} = f_\rho^{**} = p_\rho^{**} = 0. \tag{56}$$

A.2. Slutsky Equations

We first show a useful intermediate result. Note that in the individual's optimum the uncompensated demand functions should be equal to demand functions with compensation in the second period:

$$\forall b = a,\ell,p,f : \quad b\left(\overrightarrow{\sigma},\kappa^m\right) = b^{**}\left(\overrightarrow{\sigma},\pi^m,\kappa^m,\mathrm{E}\left[\tilde{V}\left(\overrightarrow{\sigma},\pi^m\right)\right]\right). \tag{57}$$

35

Take partial derivatives on both sides with respect to ρ and use the property $\mathrm{E}\left[\tilde{V}_\rho\right]=-\mathrm{E}\left[\tilde{U}_1\right]$ from (16):

$$\forall b=a,\ell,p,f: \quad b_\rho = b_\rho^{**} - \mathrm{E}\left[\tilde{U}_1\right] \cdot b_{\overline{V}}^{w**}.$$

Use results (55)-(56) and rearrange:

$$a_{\overline{V}}^{**} = -\frac{a_\rho + 1}{\mathrm{E}\left[\tilde{U}_1\right]}; \quad \forall b=\ell,p,f: \quad b_{\overline{V}}^{**} = -\frac{b_\rho}{\mathrm{E}\left[\tilde{U}_1\right]}. \tag{58}$$

Take partial derivatives of both sides of identity (57) to find Slutsky equations:

$$\forall b=a,\ell,p,f: \quad \forall \nu=t^n,t^e,\kappa^m: \quad b_\nu = b_\nu^{**} + b_{\overline{V}}^{**} \cdot \mathrm{E}\left[\tilde{V}_\nu\right].$$

Substituting results (58), the definitions for capital incomes y^n and $\tilde{y}^e$ and the definition for gross labour income z, taking into account envelope properties (17)-(18), leads to the Slutsky equations.

B. Compensated Effects of Capital Income Tax Reforms

We introduce the reform parameters σ^e and σ^n, respectively for the tax rates on excess and on risk-free capital income. Rewrite the utility function in terms of labour income z, savings a, investment in assets with excess returns f and p:

$$\tilde{\mathscr{U}}\left(z,a,f,p|\sigma^e,\sigma^n\right) \equiv U\left(\hat{c}^1(z,a),\hat{c}^2\left(a,f,p|\sigma^e,\sigma^n\right),\frac{z}{w},\mathscr{P}\right), \tag{59}$$

with:

$$\begin{aligned}
\hat{c}^1(z,a) &\equiv z-T(z)-a, \\
\hat{c}^2\left(a,f,p|\sigma^e,\sigma^n\right) &\equiv a+\left(1-t^n\right)\cdot r^b \cdot a - \sigma^n \cdot r^b \cdot \left(a-a^0\right) \\
&\quad + \left(1-t^e\right)\cdot \tilde{y}^e(f,p) - \sigma^e \cdot \left[\tilde{y}^e(f,p) - \tilde{y}^e\left(f^0,p^0\right)\right], \\
\tilde{y}^e(f,p) &= \left(\tilde{r}^m - r^b\right)\cdot f + \tilde{F}(p) - r^b \cdot p,
\end{aligned}$$

where superscripts $\cdot^0$ denote values of the respective equilibrium variables in the situation with no tax reform. The reform parameters σ^e and σ^n apply only to behavioural changes in the tax bases with respect to the situation with no tax reform. The effects of marginal reforms of these parameters are the compensated effects of reforms in the tax rates t^e and t^n, since the income effects have been removed. This can be verified by applying the envelope theorem to the Lagrangian of the optimization problem: the welfare effect of a marginal change in either σ^e or σ^n, evaluated in the situation without reform, is zero.

The first-order conditions of the utility optimization problem are the following:

$$\forall \gamma=z,a,f,p: \quad \mathrm{E}\left[\tilde{\mathscr{U}}_\gamma\right]=0.$$

When there is a reform to the parameters σ^e and σ^n, individuals update their behaviour such that these conditions continue to hold:

$$\forall \nu=\sigma^e,\sigma^n: \forall \gamma=z,a,f,p: \quad \frac{\mathrm{d}\mathrm{E}\left[\tilde{\mathscr{U}}_\gamma\right]}{\mathrm{d}\nu}=0.$$

Expand this total differential:

$$\forall \nu = \sigma^e, \sigma^n : \forall \gamma = z, a, f, p : \quad \sum_{\eta = z, a, f, p} E\big[\tilde{\mathscr{U}}_{\eta\gamma}\big] \cdot \eta_{\nu}^{**} = -E\big[\tilde{\mathscr{U}}_{\nu\gamma}\big]. \tag{60}$$

Solving this system for η_{ν}^{**} will yield compensated effects of reforms to the capital tax parameters.

We are also interested in compensated effects conditional on labour income. Substituting $z_{\nu}^{**} = 0$ yields the following system of equations for the compensated effects of capital income tax reforms:

$$\forall \nu = t^e, t^n : \forall \gamma = a, f, p : \sum_{\eta = a, f, p} E\big[\tilde{\mathscr{U}}_{\eta\gamma}\big] \cdot \eta_{\nu}^{c**} = -E\big[\tilde{\mathscr{U}}_{\nu\gamma}\big].$$

Again, compensated effects are found by solving this system.

Work out the following partial derivatives of $E\big[\tilde{\mathscr{U}}\big]$ by applying the envelope theorem to the optimization problem (59):

$$E\big[\tilde{\mathscr{U}}_z\big] = E\big[\tilde{U}_1\big] \cdot \big(1 - t_z^{\ell}\big) + \frac{E\big[\tilde{U}_{\ell}\big]}{w},$$
$$E\big[\tilde{\mathscr{U}}_a\big] = E\big[\tilde{U}_2\big] \cdot \big[1 + \big(1 - t^n - \sigma^n\big) \cdot r^b\big] - E\big[\tilde{U}_1\big],$$
$$E\big[\tilde{\mathscr{U}}_f\big] = E\big[\tilde{U}_2 \cdot \big(\tilde{r}^m - r^b\big)\big] \cdot \big(1 - t^e - \sigma^e\big),$$
$$E\big[\tilde{\mathscr{U}}_p\big] = E\big[\tilde{U}_2 \cdot \big(\tilde{F}_p - r^b\big)\big] \cdot \big(1 - t^e - \sigma^e\big).$$

B.1. Reform of Tax Rate on Excess Returns

Let us first consider the compensated effects of a reform in the tax on excess capital income. Assuming that the tax on excess capital income is not 100%, the following second-order derivatives with respect to σ^e, evaluated in the situation with no reform, are zero:

$$E\big[\tilde{\mathscr{U}}_{\sigma^e z}|\sigma^e = 0\big] = E\big[\tilde{\mathscr{U}}_{\sigma^e a}|\sigma^e = 0\big] = E\big[\tilde{\mathscr{U}}_{\sigma^e f}|\sigma^e = 0\big] = E\big[\tilde{\mathscr{U}}_{\sigma^e p}|\sigma^e = 0\big] = 0.$$

Then the system of equations (60) implies that the compensated effects of a reform in the tax rate on excess capital income are zero:

$$z_{t^e}^{**} = a_{t^e}^{**} = f_{t^e}^{**} = p_{t^e}^{**} = 0.$$

B.2. Reform of Tax Rate on Risk-Free Returns

Let us now consider the compensated effects of a reform of the tax on risk-free returns to capital, conditional on labour income. Calculate first the second-order derivatives with respect to σ^n, evaluating in the situation without reform:

$$E\big[\tilde{\mathscr{U}}_{\sigma^n a}|\sigma^n = 0\big] = -r^b E\big[\tilde{U}_2\big]; \qquad E\big[\tilde{\mathscr{U}}_{\sigma^n f}|\sigma^n = 0\big] = 0; \qquad E\big[\tilde{\mathscr{U}}_{\sigma^n p}|\sigma^n = 0\big] = 0;$$

so that we obtain the following system for the compensated effects of a reform in the tax rate on risk-free capital income:

$$\mathrm{E}\big[\tilde{\mathcal{U}}_{aa}\big]\cdot a_{t^n}^{c**} + \mathrm{E}\big[\tilde{\mathcal{U}}_{af}\big]\cdot f_{t^n}^{c**} + \mathrm{E}\big[\tilde{\mathcal{U}}_{ap}\big]\cdot p_{t^n}^{c**} = r^b\cdot\mathrm{E}\big[\tilde{U}_2\big],$$
$$\mathrm{E}\big[\tilde{\mathcal{U}}_{af}\big]\cdot a_{t^n}^{c**} + \mathrm{E}\big[\tilde{\mathcal{U}}_{ff}\big]\cdot f_{t^n}^{c**} + \mathrm{E}\big[\tilde{\mathcal{U}}_{fp}\big]\cdot p_{t^n}^{c**} = 0,$$
$$\mathrm{E}\big[\tilde{\mathcal{U}}_{ap}\big]\cdot a_{t^n}^{c**} + \mathrm{E}\big[\tilde{\mathcal{U}}_{fp}\big]\cdot f_{t^n}^{c**} + \mathrm{E}\big[\tilde{\mathcal{U}}_{pp}\big]\cdot p_{t^n}^{c**} = 0.$$

It immediately follows that when there are no conditional compensated effects on the amount of savings, then $f_{t^n}^{c**} = p_{t^n}^{c**} = 0$. All compensated conditional effects of the reform on portfolio composition, work through changes in the amount of savings.

C. Conditional Effects of Policy Reforms: Proof of Lemma 2

The assumption of weak separability between consumption and labour supply implies for all $\gamma = y^n, \tilde{y}^e, p, f$: $\frac{\mathrm{d}\gamma}{\mathrm{d}w} = \gamma_z^c \frac{\mathrm{d}z}{\mathrm{d}w}$. Combining this property with condition (31) yields:

$$\forall \gamma = y^n, \tilde{y}^e, p, f: \quad \frac{\mathrm{d}\gamma}{\mathrm{d}w} = \frac{\gamma_{\sigma^z}^{w*}}{z_{\sigma^z}^*}\frac{\mathrm{d}z}{\mathrm{d}w}. \tag{61}$$

Property (33) immediately follows by substituting Slutsky symmetry (28) for $z_{t^n}^{w**}$.

Similarly, using constancy of $\mathrm{E}\big[\tilde{U}_2\big]/\mathrm{E}\big[\tilde{U}_1\big]$ from Euler equation (10) and rearranging:

$$\frac{\mathrm{d}}{\mathrm{d}w}\left(\frac{\mathrm{E}\big[\tilde{U}_2\tilde{y}^e\big]}{\mathrm{E}\big[\tilde{U}_1\big]}\right) = \frac{\mathrm{E}\big[\tilde{U}_2\big]}{\mathrm{E}\big[\tilde{U}_1\big]}\left\{\frac{\mathrm{d}}{\mathrm{d}w}\left(\frac{\mathrm{cov}\big(\tilde{U}_2,\tilde{y}^e\big)}{\mathrm{E}\big[\tilde{U}_2\big]}\right) + \mathrm{E}\left[\frac{\mathrm{d}\tilde{y}^e}{\mathrm{d}w}\right]\right\}.$$

Use property (61) and substitute result (29) for $\mathrm{E}\big[\tilde{y}_{\sigma^z}^{e,w*}\big]$:

$$\frac{\mathrm{d}}{\mathrm{d}w}\left(\frac{\mathrm{E}\big[\tilde{U}_2\tilde{y}^e\big]}{\mathrm{E}\big[\tilde{U}_1\big]}\right) = \frac{\mathrm{E}\big[\tilde{U}_2\big]}{\mathrm{E}\big[\tilde{U}_1\big]}\left\{\frac{\mathrm{d}}{\mathrm{d}w}\left(\frac{\mathrm{cov}\big(\tilde{U}_2,\tilde{y}^e\big)}{\mathrm{E}\big[\tilde{U}_2\big]}\right) - \frac{\partial}{\partial\sigma^z}\left(\frac{\mathrm{cov}\big(\tilde{U}_2,\tilde{y}^{e*}\big)}{\mathrm{E}\big[\tilde{U}_2\big]}\right)\frac{1}{z_{\sigma^z}^*}\frac{\mathrm{d}z}{\mathrm{d}w}\right\}.$$

If $t^e = 1$, second-period marginal utility U_2^w is no longer stochastic and property (34) is true. If $t^e \neq 1$, we can substitute definition (15) for $\tilde{y}^e$ into first-order conditions (12)–(13) to find:

$$\frac{\mathrm{cov}\big(\tilde{U}_2,\tilde{y}^{e,}\big)}{\mathrm{E}\big[\tilde{U}_2\big]} = -\mathrm{E}\big[\tilde{r}^m - r^b\big]f - \mathrm{E}\big[\tilde{F}_p - r^b\big]p.$$

Property (61) can then be used to prove (34).

Lastly, twice using weak separability we find:

$$\frac{\mathrm{d}}{\mathrm{d}w}\left(\frac{\mathrm{E}\big[\tilde{U}_2 h(\tilde{r}^m, R^m, \delta)\big]}{\mathrm{E}\big[\tilde{U}_1\big]}\right) = \frac{\partial}{\partial z}\left(\frac{\mathrm{E}\big[\tilde{U}_2^c h(\tilde{r}^m, R^m, \delta)\big]}{\mathrm{E}\big[\tilde{U}_1^c\big]}\right)\frac{\mathrm{d}z}{\mathrm{d}w}$$
$$= \frac{\partial}{\partial\sigma^z}\left(\frac{\mathrm{E}\big[\tilde{U}_2^* h(\tilde{r}^m, R^m, \delta)\big]}{\mathrm{E}\big[\tilde{U}_1^*\big]}\right)\frac{1}{\cdot z_{\sigma^z}^*}\frac{\mathrm{d}z}{\mathrm{d}w}.$$

Substituting Slutsky symmetry (30) yields property (35).

D. Government First-Order Conditions

D.1. Optimal Intercept

Take the first-order derivative of Lagrangian (39) with respect to ρ, evaluate in the situation without reform, and require that the result is zero:

$$0 = \int_{\underline{w}}^{\overline{w}} \left(1 + r^b - \frac{\mathrm{E}\big[\tilde{U}_1\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} + \big(1 + r^b\big) t_z^\ell z_\rho + t^n y_\rho^n + t^e \mathrm{E}\big[\tilde{y}_\rho^e\big] + t^e \frac{\mathrm{cov}\big(\tilde{\lambda}, \tilde{y}_\rho^e\big)}{\mathrm{E}\big[\tilde{\lambda}\big]} \right) \mathrm{d}G^{\mathsf{w}}(w),$$

where we used envelope result (16).

D.2. Optimal Labour Income Tax

Take the first-order derivative of Lagrangian (39) with respect to σ^z, evaluate in the situation without reform, and require the result to be zero:

$$0 = \int_{\underline{w}}^{\overline{w}} \left(1 + r^b - \frac{\mathrm{E}\big[\tilde{U}_1\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} \right) \cdot H(z, Z, \delta) \mathrm{d}G^{\mathsf{w}}(w)$$
$$+ \int_{\underline{w}}^{\overline{w}} \left[\big(1 + r^b\big) t_z^{\ell,w} z_{\sigma^z}^w + t^n y_{\sigma^z}^{n,w} + t^e \frac{\mathrm{E}\big[\tilde{\lambda} \cdot \tilde{y}_{\sigma^z}^{e,w}\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} \right] \mathrm{d}G^{\mathsf{w}}(w).$$

Apply Slutsky equations (21) and substitute β^w from definition (46):

$$\int_{\underline{w}}^{\overline{w}} \big(1 + r^b - \beta^w\big) \cdot H(z, Z, \delta) \mathrm{d}G^{\mathsf{w}}(w)$$
$$= -\int_{\underline{w}}^{\overline{w}} \left(\big(1 + r^b\big) t_z^\ell z_{\sigma^z}^* + t^n y_{\sigma^z}^{n*} + t^e \frac{\mathrm{E}\big[\tilde{\lambda} \cdot \tilde{y}_{\sigma^z}^{e*}\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} \right) \cdot h(z, Z, \delta) \mathrm{d}G^{\mathsf{w}}(w).$$

Change bases in the integrals and substitute definitions (3)-(4) for $H(z^w, Z, \delta)$ and $h(z, Z, \delta)$:

$$\int_{Z-\frac{\delta}{2}}^{Z+\frac{\delta}{2}} \big(1 + r^b - \beta^w\big) \cdot \left(z^w - Z + \frac{\delta}{2} \right) \mathrm{d}G^{\mathsf{z}}\big(z^w\big) + \delta \cdot \int_{Z+\frac{\delta}{2}}^{\overline{z}} \big(1 + r^b - \beta^w\big) \mathrm{d}G^{\mathsf{z}}\big(z^w\big)$$
$$= -\int_{Z-\frac{\delta}{2}}^{Z+\frac{\delta}{2}} \left(\big(1 + r^b\big) t_z^\ell z_{\sigma^z}^{w*} + t^n y_{\sigma^z}^{n,w*} + t^e \frac{\mathrm{E}\big[\tilde{\lambda} \tilde{y}_{\sigma^z}^{e,w*}\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} \right) \mathrm{d}G^{\mathsf{z}}\big(z^w\big).$$

Divide the entire equation by the right-hand side and apply L'Hôpital's rule to find the limit $\delta \to 0$. Write up the result for a marginal labour income tax reform at skill level w to find equation (40).

D.3. Optimal Linear Tax on Excess Capital Income

Take the first-order derivative of Lagrangian (39) with respect to t^e, evaluate in the situation with no reform, and require that the result is zero:

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

Apply Slutsky equations (22)-(23), use the fact that a tax reform on excess returns has no compensated effects (lemma 1), and substitute definition (46) for β^w:

$$\int_{\underline{w}}^{\overline{w}} \frac{\mathrm{E}\left[\tilde{U}_2\tilde{y}^e\right]}{\mathrm{E}\left[\tilde{U}_1^w\right]} \cdot \left(1+r-\beta^w\right)\mathrm{d}G^w(w)$$
$$= \int_{\underline{w}}^{\overline{w}} \left(\frac{\mathrm{E}\left[\tilde{U}_2\tilde{y}^e\right]}{\mathrm{E}\left[\tilde{U}_2^w\right]} - \frac{\mathrm{E}\left[\tilde{\lambda}\tilde{y}^e\right]}{\mathrm{E}\left[\tilde{\lambda}\right]} \right)\mathrm{d}G^w(w).$$

Use integration by parts to rewrite the left-hand side (using the optimal intercept (43)),

$$\int_{\underline{w}}^{\overline{w}} \frac{\mathrm{d}}{\mathrm{d}w} \left(\frac{\mathrm{E}\left[\tilde{U}_2^w \tilde{y}^{e,w}\right]}{\mathrm{E}\left[\tilde{U}_1^w\right]} \right) \cdot \left(\int_w^{\overline{w}} \left(1+r-\beta^{\hat{w}}\right)\mathrm{d}G^w(\hat{w}) \right)\mathrm{d}w$$
$$= \int_{\underline{w}}^{\overline{w}} \left(\frac{\mathrm{E}\left[\tilde{U}_2^w \tilde{y}^e\right]}{\mathrm{E}\left[\tilde{U}_2^w\right]} - \frac{\mathrm{E}\left[\tilde{\lambda}\tilde{y}^e\right]}{\mathrm{E}\left[\tilde{\lambda}\right]} \right)\mathrm{d}G^w(w).$$

Use result (34) to find that the left-hand side is zero. Rearrange to find result (41).

D.4. Optimal Linear Tax on Risk-Free Capital Income

D.4.1. Optimality Condition

Take the first-order derivative of Lagrangian (39) with respect to t^n, evaluate in the situation with no reform, and require that the result is zero:

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

Apply Slutsky equations (22)-(23) to find:

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

where we substituted the Euler equation from (10) and definition (46) for β^w. Use integration by parts to rewrite the left-hand side (using the optimal intercept (43)):

$$\int_{\underline{w}}^{\overline{w}} \left\{ \int_{w}^{\overline{w}} \left(1 + r - \beta^{\hat{w}}\right) dG^{\mathsf{w}}(\hat{w}) \right\} \cdot \frac{d}{dw} \left(\frac{E[\tilde{U}_2]}{E[\tilde{U}_1]} y^n \right) dw$$
$$= -\int_{\underline{w}}^{\overline{w}} \left\{ \left(1 + r^b\right) t_z^{\ell} z_{t^n}^{**} + t^n y_{t^n}^{n**} + t^e \frac{E[\tilde{\lambda} \tilde{y}_{t^n}^{e**}]}{E[\tilde{\lambda}]} \right\} dG^{\mathsf{w}}(w).$$

Substitute the optimal labour tax equation (40) for $\left(1 + r^b\right) t_z^{\ell,w}$:

$$\int_{\underline{w}}^{\overline{w}} \left\{ \int_{w}^{\overline{w}} \left(1 + r - \beta^{\hat{w}}\right) dG^{\mathsf{w}}(\hat{w}) \right\} \cdot \left(\frac{E[\tilde{U}_2]}{E[\tilde{U}_1]} \frac{dy^n}{dw} - \frac{z_{t^n}^{**}}{z_{\sigma^z}^{*}} \frac{dz}{dw} \right) dw$$
$$= -\int_{\underline{w}}^{\overline{w}} \left\{ t^n \left(y_{t^n}^{n**} - \frac{z_{t^n}^{**}}{z_{\sigma^z}^{*}} y_{\sigma^z}^{n*} \right) + t^e \frac{E\left[\tilde{\lambda} \cdot \left(\tilde{y}_{t^n}^{e**} - \frac{z_{t^n}^{**}}{z_{\sigma^z}^{*}} \tilde{y}_{\sigma^z}^{e*} \right)\right]}{E[\tilde{\lambda}]} \right\} dG^{\mathsf{w}}(w).$$

Use result (32) and lemma (2) to simplify and find equation (42).

D.5. Optimal Public Good Provision

D.5.1. Optimality Condition

Take the first-order derivative of Lagrangian (39) with respect to π^m, evaluate in the situation without reform, and require that the result be zero, to find the optimality condition for provision of the public good:

$$E\left[\tilde{\lambda} \cdot h\left(\tilde{r}^m, R^m, \delta\right)\right] = \int_{\underline{w}}^{\overline{w}} E\left[\tilde{U}_P \cdot h\left(\tilde{r}^m, R^m, \delta\right)\right] dG^{\mathsf{w}}(w).$$

Here we used our assumption that there are no behavioural responses to changes in the public goods. Apply definition (3) for $h\left(\tilde{r}^m, R^m, \delta\right)$ and take the limit $\delta \to 0$ to find condition (44).

D.5.2. Proof of Proposition 6

Start from optimal tax expression (40) for labour income. Multiply both sides by $\frac{d}{dz^w}\left(E[\tilde{U}_P]/E[\tilde{U}_1]\right)$ and integrate over the entire population:

$$\int_{\underline{z}}^{\overline{z}} \frac{d}{dz} \left(\frac{E[\tilde{U}_P]}{E[\tilde{U}_1]} \right) \left\{ \left(1 + r^b\right) t_z^{\ell} z_{\sigma^z}^{*} + t^n y_{\sigma^z}^{n*} + t^e \frac{E[\tilde{\lambda} \cdot \tilde{y}_{\sigma^z}^{e*}]}{E[\tilde{\lambda}]} \right\} dG^{\mathsf{z}}(z)$$
$$= -\int_{\underline{z}}^{\overline{z}} \frac{d}{dz} \left(\frac{E[\tilde{U}_P]}{E[\tilde{U}_1]} \right) \cdot \int_{z}^{\overline{z}} \left(1 + r^b - \beta^{\hat{w}}\right) dG^{\mathsf{z}}\left(z^{\hat{w}}\right) dz.$$

Use integration by parts on the second line and use the optimality condition for the intercept (43):

$$\int_{\underline{z}}^{\overline{z}} \frac{E[\tilde{U}_P]}{E[\tilde{U}_1]} \left(\frac{E[\tilde{U}_1]}{E[\tilde{\lambda}]} - 1 - r^b \right) dG^z(z)$$

$$= \int_{\underline{z}}^{\overline{z}} \frac{E[\tilde{U}_P]}{E[\tilde{U}_1]} \left(\left(1 + r^b\right) t_z^\ell z_\rho + t^n y_\rho^n + t^e \frac{E[\tilde{\lambda} \tilde{y}_\rho^e]}{E[\tilde{\lambda}]} \right) dG^z(z)$$

$$+ \int_{\underline{z}}^{\overline{z}} \frac{d}{dz} \left(\frac{E[\tilde{U}_P]}{E[\tilde{U}_1]} \right) \left\{ \left(1 + r^b\right) t_z^\ell z_{\sigma^z}^* + t^n y_{\sigma^z}^{n*} + t^e \frac{E[\tilde{\lambda} \cdot \tilde{y}_{\sigma^z}^{e*}]}{E[\tilde{\lambda}]} \right\} dG^z(z).$$

The left-hand side is the distributional effect of a labour income tax reform where the liability at each income level z^w is increased by $E[\tilde{U}_P^w]/E[\tilde{U}_1^w]$. The second line contains the budgetary cost of this reform caused by income effects, while the third line contains compensated effects caused by the fact that the change in the tax liability is not equal for all income levels. Substitute this last result into condition (53).

D.6. Optimal Stochastic Lump Sum

Let us now derive the optimality condition for the stochastic lump sum. The first-order condition of Lagrangian 39 with respect to κ^m, evaluated in the situation without reforms, is as follows:

$$\frac{E[\tilde{\lambda} \cdot h(\tilde{r}^m, R^m, \delta)]}{\tilde{\lambda}} - \int_{\underline{w}}^{\overline{w}} \frac{E[\tilde{U}_2 h(\tilde{r}^m, R^m, \delta)]}{E[\tilde{\lambda}]} dG^w(w)$$

$$= \int_{\underline{w}}^{\overline{w}} \left(\left(1 + r^b\right) t_z^\ell z_{\kappa^m} + t^n y_{\kappa^m}^n + t^e \frac{E[\tilde{\lambda} y_{\kappa^m}^e(\tilde{r}^m, \tilde{\varepsilon})]}{E[\tilde{\lambda}]} \right) dG^w(w),$$

where we used envelope property (18). Apply the Slutsky equations of subsection 3.1 and substitute definition (46) of β^w:

$$\frac{E[\tilde{\lambda} \cdot h(\tilde{r}^m, R^m, \delta)]}{E[\tilde{\lambda}]} - \int_{\underline{w}}^{\overline{w}} \frac{E[\tilde{U}_2 h(\tilde{r}^m, R^m, \delta)]}{E[\tilde{U}_1]} \left(\beta^w - t^n r^b \right) dG^w(w)$$

$$= \int_{\underline{w}}^{\overline{w}} \left[\left(1 + r^b\right) t_z^\ell z_{\kappa^m}^{**} + t^n y_{\kappa^m}^{n**} + t^e \frac{E[\tilde{\lambda} \tilde{y}_{\kappa^m}^{e**}]}{E[\tilde{\lambda}]} \right] dG^w(w).$$

Use integration by parts and optimality condition (43) for the intercept:

$$\frac{E[\tilde{\lambda} \cdot h(\tilde{r}^m, R^m, \delta)]}{E[\tilde{\lambda}]} - \int_{\underline{w}}^{\overline{w}} \frac{E[\tilde{U}_2 h(\tilde{r}^m, R^m, \delta)]}{E[\tilde{U}_2]} dG^w(w)$$

$$= \int_{\underline{w}}^{\overline{w}} \frac{d}{dw} \left(\frac{E[\tilde{U}_2 h(\tilde{r}^m, R^m, \delta)]}{E[\tilde{U}_1^w]} \right) \int_w^{\overline{w}} \left(\beta^{\hat{w}} - 1 - r^b \right) dG^w(\hat{w}) dw$$

$$+ \int_{\underline{w}}^{\overline{w}} \left[\left(1 + r^b\right) t_z^\ell z_{\kappa^m}^{**} + t^n y_{\kappa^m}^{n**} + t^e \frac{E[\tilde{\lambda} \tilde{y}_{\kappa^m}^{e**}]}{E[\tilde{\lambda}]} \right] dG^w(w).$$

Insert optimality condition (40) for the labour income tax:

$$\frac{\mathrm{E}\big[\tilde{\lambda}\cdot h(\tilde{r}^m,R^m,\delta)\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} - \int_{\underline{w}}^{\overline{w}} \frac{\mathrm{E}\big[\tilde{U}_2 h(\tilde{r}^m,R^m,\delta)\big]}{\mathrm{E}\big[\tilde{U}_2\big]} \mathrm{d}G^{\mathrm{w}}(w)$$

$$= \int_{\underline{w}}^{\overline{w}} \left\{ \frac{\mathrm{d}}{\mathrm{d}w}\left(\frac{\mathrm{E}\big[\tilde{U}_2 h(\tilde{r}^m,R^m,\delta)\big]}{\mathrm{E}\big[\tilde{U}_1\big]}\right) - \frac{z_{\kappa^m}^{**}}{z_{\sigma^z}^*}\frac{\mathrm{d}z}{\mathrm{d}w} \right\} \cdot \int_w^{\overline{w}} \big(\beta^{\hat{w}} - 1 - r^b\big)\mathrm{d}G^{\mathrm{w}}(\hat{w})\mathrm{d}w$$

$$+ \int_{\underline{w}}^{\overline{w}} \left[t^n\left(y_{\kappa^m}^{n**} - \frac{y_{\sigma^z}^{n*}}{z_{\sigma^z}^*} z_{\kappa^m}^{**}\right) + t^e \frac{\mathrm{E}\big[\tilde{\lambda}\big(\tilde{y}_{\kappa^m}^{e**} - \frac{\tilde{y}_{\sigma^z}^{e*}}{z_{\sigma^z}^*} z_{\kappa^m}^{**}\big)\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} \right] \mathrm{d}G^{\mathrm{w}}(w).$$

Finally substituting conditional demands from equation (32) and property (35):

$$\frac{\mathrm{E}\big[\tilde{\lambda}\cdot h(\tilde{r}^m,R^m,\delta)\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} = \int_{\underline{w}}^{\overline{w}} \left(\frac{\mathrm{E}\big[\tilde{U}_2 h(\tilde{r}^m,R^m,\delta)\big]}{\mathrm{E}\big[\tilde{U}_2\big]} + t^n y_{\kappa^m}^{nc**} + t^e \frac{\mathrm{E}\big[\tilde{\lambda}\tilde{y}_{\kappa^m}^{ec**}\big]}{\mathrm{E}\big[\tilde{\lambda}\big]} \right) \mathrm{d}G^{\mathrm{w}}(w).$$