30 December 2025

ISSN 1011-8888

INSTITUTE OF ECONOMIC STUDIES

WORKING PAPER SERIES

W25:04

December 2025

Optimal Taxation when access to Income Shifting is Heterogeneous

Arnaldur Solvi Kristjansson

Address:

Faculty of Economics

University of Iceland

Oddi, at Sturlugata,

101 Reykjavik, Iceland

Email: askristjansson@hi.is

Optimal Taxation when access to Income Shifting is Heterogeneous*

Arnaldur Sölvi Kristjánsson†

December 2025

Abstract

This paper characterizes optimal labour and capital income taxation when access to income shifting opportunities is heterogeneous. I develop a two-period model in which individuals can reclassify labour income as capital income by paying a fixed cost (an extensive margin, e.g. incorporation or organisational changes) and a variable shifting cost (an intensive margin). In the benchmark without shifting, redistribution is implemented through the labour income tax while capital income is left untaxed. When shifting is possible, the optimal policy compresses the labour–capital tax wedge: stronger shifting responses call for lower labour income tax rates and higher capital income tax rates. Allowing the capital income tax to depend on reported labour income, the optimal capital tax schedule is progressive in reported labour income when shifting elasticities increase with income. A key implication is that extensive-margin participation receives disproportionate weight for capital taxation, whereas intensive and extensive responses affect the optimal labour tax symmetrically.

Keywords: Optimal income taxation, income shifting

JEL Codes: H21, H24, H26

*I would like to thank Vidar Christiansen and Ray Rees for useful suggestions and comments.
†Department of Economics, University of Iceland. E-mail: askristjansson@hi.is

1 Introduction

Economists continue to debate how capital income should be taxed. Views range from exempting capital income altogether to taxing it at a flat rate, at a rate that varies with labour income, or under a comprehensive income tax that applies the same rate schedule to all income sources. A broad consensus, however, is that large tax-rate differentials between labour and capital income create strong incentives for income shifting, especially among owner-managers. This concern is often cited as a central argument for keeping labour and capital income tax rates relatively close (see e.g., Banks & Diamond, 2010; Bastani & Waldenström, 2020; Jacobs, 2013).

Banks & Diamond (2010) argue that when labour income is taxed progressively, capital income taxation should be designed in relation to the labour income tax schedule rather than treated as an independent flat tax. This paper provides a theoretical model to analyse under what conditions it is optimal to link the rate on capital income to reported labour income.

In this paper, I study optimal labour and capital income taxation when access to income shifting opportunities is heterogeneous across individuals. A central feature of income shifting in practice is that it often requires discrete organisational changes, such as incorporation, suggesting an extensive margin that has been largely absent from optimal tax models. The government chooses a nonlinear labour income tax schedule and a labour-income-contingent linear tax on reported capital income. In the absence of income shifting, the optimal policy features redistribution through the labour income tax while leaving capital income untaxed, consistent with the classic Atkinson–Stiglitz logic (Atkinson & Stiglitz, 1976). Once income shifting is possible, the optimal response is to compress the labour-capital income tax differential: higher shifting incentives call for lower labour income tax rates and higher capital income tax rates.

This paper analyses a two-period model where individuals face a variable as well as a fixed cost when shifting income between the labour and capital income tax base. Income shifting leads to an increase in the optimal capital income tax rate with labour income when the intensive and extensive shifting elasticities increase with reported labour income. However, the intensive shifting behaviour plays a less important role. Intuitively, labour taxes are pinned down by how responsive reported income is at a given bracket, regardless of whether the response comes from more shifting or more people entering shifting. Capital taxes, in contrast, are sensitive to entry and exit from shifting, because stopping shifting changes the entire shifted base.

Individuals who legally shift income between tax bases in order to minimise tax payments usually face some costs such as monetary outlays (e.g. tax consultant costs), opportunity cost of time spent on tax-minimising strategies or engagement in suboptimal behaviour and other non-pecuniary costs such as tax morale. However, to shift income, individuals must be able to alter their remuneration scheme, and such possibilities are heterogeneous. Employers usually report the income of their employees to tax authorities, making it almost impossible for employees to shift income unless they get paid as a contractor. This is something employers and employees are often unwilling to do (Barth & Ognedal, 2018). Owner-managers usually have more control over the way their work is remunerated, though they may also be constrained by tax laws. For example, in many countries, enterprises need to incorporate in order to be able to shift income. This means that there may be fixed costs involved in getting access to income-shifting opportunities, due to direct and indirect costs of organisational changes or other adjustment costs.

There is a growing empirical literature, which is reviewed in Section 2, that analyses how differences in capital and labour income taxation affect income shifting. The general result from this literature is that income shifting responds strongly to tax incentives, both along the intensive (shifts between tax bases) and extensive margin (organizational shifts). This implies that both behavioural margins are important when determining the effects of taxes on income shifting. Along the intensive margin, these effects are heterogeneous because owner-managers of firms have much easier access to income shifting than employees. Along the extensive margin, the effects seem to be stronger for small firms, as the organizational form for large firms is mostly determined by non-tax factors.

According to the study by Alstadsæter & Jacob (2017), three conditions are crucial in order to explain participation in income shifting: incentives, access and awareness. First, individuals need sufficient tax incentives to shift income (in Sweden this means having sufficiently high income). Second, individuals need access to control the combination of their income, i.e. whether income is paid as labour or capital income. Owner-managers of firms have generally greater possibilities to determine tax-minimizing combinations of their compensation, while employees need to negotiate with their employer. Third, individuals need to be aware of the tax rules and of the possibility of shifting.

There is a growing literature on optimal taxation with legal income shifting at the intensive margin (Christiansen & Tuomala, 2008; Fuest & Huber, 2001, 2005; Piketty et al., 2014; Reis, 2011; Selin & Simula, 2020). The typical assumption is that everyone has access to income shifting and faces the same increasing and convex cost of shifting. The general result of this literature is that income shifting is a rationale for taxing capital income. Most of the theoretical papers build on a baseline model where capital income should not be taxed in the absence of income shifting.

To the best of my knowledge, Selin & Simula (2020) is the only paper that considers income shifting at the extensive margin. They analyse a model where people differ in skills, taste for work effort and the fixed cost incurred when shifting income. They find that when people who shift easily along the extensive margin are also more elastic in labour supply, the government should not necessarily combat income shifting. They argue that this is similar to third-degree price discrimination and works as a form of endogenous tagging.

I present a model where people differ in labour market productivity and in fixed and variable shifting costs. The model is closely related to the model in Christiansen & Tuomala (2008). My main innovation compared to their paper is to add the fixed shifting cost. There are two periods and all individuals work in the labour market and allocate resources between the periods. In addition to this, individuals can decide to pay the fixed cost and gain access to income shifting opportunities. The government only observes reported labour and capital income.

This paper is organized as follows. Section 2 reviews the related empirical literature. In Section 3, I present the model. Thereafter, behavioural elasticities are discussed. Section 5 presents the optimal tax results. I conclude in the final section.

2 Related Empirical Literature

This section reviews empirical evidence on behavioural responses to tax rate differentials between labour and capital income, with a focus on income shifting. I emphasize two margins that are central for my framework: (i) *intensive-margin* shifting within a given legal form (timing and relabeling), and (ii) *extensive-margin* responses through incorporation and organizational-form choices.

Taxable income responses and the role of shifting. Following Feldstein (1995), there has emerged a large literature estimating the elasticity of taxable income. Feldstein finds a substantial elasticity for top income groups in response to a large cut in top marginal tax rates in 1986. Analysing the 1986 tax reform, Slemrod (1995) proposes a three-tier hierarchy of behavioural responses to taxation, with timing responses being most responsive, followed by avoidance responses and real responses at the bottom. Using tax records, Miller et al. (2024) show that large responses of UK company owner-managers to personal taxes are driven by intertemporal income shifting (profit retention and withdrawal timing), rather than responses to real business activity. Consistent with this interpretation, Slemrod (1996) and Saez (2004) find large responses among top income groups around the 1986 reform, with evidence pointing to substantial income shifting. The review by Saez et al. (2012) similarly highlights that high-income responses are frequently driven by avoidance opportunities, including deductions and income shifting.

Corporate versus personal taxation and incorporation. In the US and many other countries, corporate profits are taxed under a separate corporate income tax, while unincorporated business profits are taxed at the individual level. In the US, closely held corporations with few shareholders may be taxed solely at the individual level. Relatedly, closely held firms can also serve as a vehicle for tax sheltering through the accumulation and sheltering of income inside the firm, consistent with avoidance and income shifting rather than purely real responses (Alstadsæter et al., 2014). Gordon & Slemrod (2000) exploit time series from the US and find substantial income shifting for high-income taxpayers with respect to the difference between the corporate and the personal tax rates. Using data from the UK, Devereux et al. (2014) decompose corporate income responses into real responses and income shifting and find a rather modest income shifting elasticity at low levels of profits.

Taxes can also affect behaviour through the extensive margin by changing the incentives to incorporate. Earlier evidence using aggregate time-series variation typically finds modest effects of taxes on incorporation (Goolsbee, 1998; Gordon & MacKie-Mason, 1994; Mackie-Mason & Gordon, 1997). That may be because aggregate data are dominated by larger firms, for which non-tax factors may be more important. Using cross-sectional evidence from retail sales, where activity is concentrated among small or single-establishment firms, Goolsbee (2004) finds larger tax effects and argues that the time-series approach is limited by modest policy variation and confounding contemporaneous reforms. Using European panel data, De Mooij & Nicodème (2008) and Elschner (2013) also find sizeable tax effects on incorporation. Tazhitdinova (2020) shows that increases in the tax savings from incorporation raise entry into business ownership and are associated with income shifting. Studying the 2012 Kansas reform that exempted pass-through business income from state taxation, DeBacker et al. (2019) find responses on both the extensive and intensive margins that are largely consistent with income recharacterization and shifting of effort into pass-through form, with limited evidence of increased real economic activity.

Income shifting in dual income tax systems. Income shifting is a particular challenge in dual income tax systems, where capital income is taxed at a flat rate while labour income is taxed progressively. For sufficiently high-income individuals, this creates incentives to report labour income as more lightly taxed capital income. Empirical studies for Scandinavian countries typically find large responses concentrated among owner-managers, while employees respond little, consistent with more limited scope for shifting. For Finland, Pirttilä & Selin (2011) and Harju & Matikka (2016) study the 1993 and 2005 reforms, respectively; for Sweden, Alstadsæter & Jacob (2016) studies the 2006 reform.[1] There is also evidence that organizational form in dual income tax systems is affected by tax differentials, particularly among small firms (Alstadsæter & Jacob, 2017, 2016; Edmark & Gordon, 2013; Romanov, 2006).

Implications for the framework. Taken together, the evidence points to two empirically relevant margins of income shifting. Within-form responses (timing and relabelling) suggest modelling shifting as a continuous choice with increasing marginal costs, while incorporation and organizational-form changes motivate a fixed-cost participation margin. Finally, the fact that responses are concentrated among owner-managers but limited for employees supports modelling heterogeneous access to shifting opportunities, a central element of our theoretical analysis.

3 The Model

Individuals are heterogeneous along three dimensions: their labour market productivity $w \in [\underline{w}, \overline{w}]$, the fixed cost they face to shift income $k \in [\underline{k}, \overline{k}]$, and the variable cost of income shifting $m \in [\underline{m}, \overline{m}]$. The heterogeneity in k and m leads to heterogeneous shifting decisions at the extensive and intensive margins, respectively. The distribution of w, k and m is given by the joint probability density function $f(w, k, m)$. The size of the total population

[1] See (Selin, 2025) on income splitting rules in the Nordic countries.

is normalized to one.

To focus on the shifting margins, I abstract from labour supply by assuming that all individuals work full-time. Labour income is denoted by y. I follow Saez (2002) and Diamond & Spinnewijn (2011) by having a discrete distribution of income levels y_i, for $i = 1, ..., I$. Labour income y_i is increasing in i, i.e. $y_1 < ... < y_I$. To choose a given income level, the individual's labour productivity needs to be at least as high as income level y_i. Everyone will choose the highest possible income level. Individuals with $w \in [\underline{w}, w_2)$ choose y_1, individuals with $w \in [w_2, w_3)$ choose y_2, and so forth. Those with $w \in [w_{I-1}, \overline{w}]$ choose y_I. Note that $\underline{w} \geq y_1$ and $\overline{w} > y_I$. The difference between adjacent income levels is constant: $y_{i+1} - y_i = \delta$, $\forall i$.[2]

3.1 Individual choices

I consider a standard two-period model where individuals save in the first period and work in the second. Individuals can legally shift income between the labour income and capital income tax bases in order to reduce the total tax liability. In practice, it can be difficult for tax authorities to distinguish labour income from capital income. An owner-manager of a firm will by definition receive labour and capital income. However, the exact division of income may be unclear, even conceptually.[3] The tax rules will therefore likely be arbitrary to some degree. It is therefore not surprising that some people are able to influence the composition of income reported to the tax authorities.

Access to income shifting opportunities is heterogeneous. It is generally easier for owner-managers to influence how their work is remunerated than it is for employees. This element is incorporated into the model by assuming that income shifting is only available to those who pay a fixed cost k, which reflects various factors such as the costs of setting up and operating a corporation, and other necessary adjustments to enable income shifting. This

[2] Having discrete income levels simplifies conditioning the capital-income tax on reported labour income, as is done in the model.

[3] Imagine a self-employed dentist who owns their own equipment. It's not clear what share of the total income should be categorized as a return on the equipment and as remuneration for work effort.

cost varies across individuals. While the costs of incorporation may not vary significantly across individuals, other related costs can differ. For instance, individuals working in the public sector might need to switch to a different sector to take a job where income shifting is possible. Additionally, some individuals may have a preference for being self-employed, or they may differ in their knowledge of income shifting opportunities. For example, accountants are likely to be better informed about these opportunities than others. Finally, individuals may have heterogeneous preferences regarding tax avoidance behaviour.

Individuals who pay the fixed cost k are defined to be shifters, and those who do not pay the cost are non-shifters. This is the binary income shifting decision individuals face. Once individuals decide to become a shifter, they have to decide how much income to shift. This is the marginal income shifting decision individuals face, i.e. the choice of j. The benefits of becoming a shifter are to reduce reported labour income by $j\delta$ and simultaneously increase reported capital income by $j\delta$, where $j\delta$ denotes the amount that is shifted. Conditional on paying k, reported labour income can be adjusted by choosing $j \in \mathbb{Z}$; $j > 0$ corresponds to shifting labour income to capital income, and $j < 0$ corresponds to shifting capital income to labour income. However, under the optimal tax schedules that are characterised below, equilibrium choices satisfy $j \geq 0$, so the relevant margin is the extent of shifting from labour income to capital income.

Once individuals have chosen to become a shifter, income shifting involves some variable costs. The role of the variable cost is to determine the optimal choice of j. Without such costs, changes in the tax system might not affect j. I follow the standard approach from the literature, where shifting an amount of $j\delta$ entails a loss of consumption of $d(j\delta, m)$, which is increasing and convex in $j\delta$, increasing in m, and has a positive cross-derivative.[4] I interpret m as a parameter capturing heterogeneity in access to income shifting technologies, with higher m corresponding to higher marginal shifting costs, and incur a higher marginal cost of increased income shifting. This implies that those with the lowest m will shift the most

[4]See e.g. Christiansen & Tuomala (2008) and Selin & Simula (2020).

if k is sufficiently low for them to be willing to become a shifter. The variable cost $d(j\delta, m)$ entails several elements: monetary outlays (e.g., tax consultant costs), the opportunity cost of time spent on tax-minimizing strategies or engaging in suboptimal behaviour, and other non-pecuniary costs such as tax morale. Additionally, if there is more than one owner of the business, altering the composition of total income becomes more burdensome, as increasing profits at the expense of wages will raise the capital income of all owners.

Individuals live for two periods. In the first period, individuals start with an endowment e, which is fixed across the population. Individuals are free to allocate their endowment between consumption in the first period and savings, denoted by b and s, respectively. The economy is small and open, and the rate of return on savings is exogenously given by r.

I follow Christiansen & Tuomala (2008) by letting individuals receive labour income in the second period rather than the first, which is standard in such two-period models. The reason for this approach is that in standard two period models, individuals do not receive labour and capital income in the same period, which is convenient for modelling income shifting.

Reported capital income consists of true capital income, which is rs, plus the shifted income, i.e., $rs + j\delta$. I aim to analyse how the optimal marginal tax rate on capital income relates to labour income. If the government could implement a general tax function, this relationship could be analysed by examining the cross-derivative of the tax function. However, this would require solving a multidimensional screening problem. To make the problem more tractable, I follow the approach of Diamond & Spinnewijn (2011) and allow the tax rate on capital income to depend on reported labour income. Individuals reporting y_i pay a tax rate t_i on their reported capital income.

Following Mirrlees (1971), the government does not observe individual types (w, k, and m). The government can only condition income taxation on reported income. The government sets a non-linear tax on reported labour income. Individuals reporting y_i face an average labour income tax rate τ_i and thus pay $\tau_i y_i$ in labour income tax; τ_i is an average (not marginal) tax rate. Individuals who report y_i are either non-shifters earning y_i or shifters earning y_{i+j} who shift $y_{i+j} - y_i = j\delta$ from the labour income tax base to the capital income tax base, where j denotes the total number of income levels that a shifter shifts downwards. The choice of j constitutes the intensive shifting decision. The extensive shifting decision is choosing between $j = 0$ and $j \neq 0$.

Individuals earning y_i who do not shift have the following tax liability:

$$T_i = \tau_i y_i + t_i rs,$$

where τ_i is the average tax rate on labour income, t_i is a labour-income-contingent linear capital tax rate and is both the marginal as well as the average tax rate on capital income. The marginal tax rate on labour income is denoted by $MTR_i^L = \frac{\Delta(\tau_i y_i)}{\Delta y_i}$, with $\Delta y_i = \delta$.

Tax liabilities for a shifter individual earning y_{i+j} and reporting y_i as labour income are

$$T_i^{i+j} = \tau_i y_i + t_i(rs + j\delta) = T_i + t_i j\delta,$$

Individuals pay labour income tax and face the linear capital income tax as if they were working in job i. Note that $T_i^i = T_i$.

Second-period consumption for non-shifters and shifters is, respectively,

$$\begin{aligned} c_i =& y_i + s_i(1 + r) - T_i, \\ c_i^{i+j} =& y_{i+j} + s_i(1 + r) - T_i^{i+j} - k - d(j\delta, m). \end{aligned}$$

Individuals have identical, separable, and additively separable preferences, represented by the utility function

$$U = u(b) + c, \tag{2}$$

where $u(b)$ is increasing and concave.

Due to the quasi-linearity in c and since the fixed cost for shifting (k) is paid in the second-period, the behaviour of the shifters does not depend on k. This means that, conditional on w, shifters will make choices irrespective of k. Trivially, the behaviour of non-shifters does not depend on k as well. This means that k will only affect the discrete choice between being a shifter or a non-shifter, and does not affect other decision variables (y, s, j). All individuals choose their savings, while their labour income is simply given by their productivity (due to the full-time work assumption).

3.2 The Government

The government adopts a welfarist criterion that sums over all individuals, a transformation $W(U)$ of individual utility, where $W(U)$ is increasing and concave, which captures the government's redistributive preferences. The government's objective function is denoted by

$$\Omega = E[W(U)]. \tag{3}$$

The expectation operator averages over w, k and m.

The government's budget constraint can be expressed in the following three ways:

$$\begin{aligned} R &= \sum_{i=1}^{I} \hat{n}_i T_i + \sum_{i=1}^{I} \sum_{j=1}^{I-1} \alpha_i^{i+j} t_i j \delta && \text{(4a)} \\ &= \sum_{i=1}^{I} n_i T_i - \sum_{i=1}^{I} \sum_{j=1}^{I-1} \alpha_i^{i+j} \left[T_{i+j} - T_i^{i+j} \right] && \text{(4b)} \\ &= \sum_{i=1}^{I} n_i T_i - \sum_{i=1}^{I} \sum_{j=1}^{i-1} \alpha_{i-j}^{i} \left[T_i - T_{i-j}^{i} \right] \geq \bar{R}, && \text{(4c)} \end{aligned}$$

where n_i is the share of people who earn y_i, $\hat{n}_i$ is the share of people who report y_i, and α_i^{i+j} is the share of the population who earn y_{i+j} and report y_i. R denotes the actual revenue received by the government and $\bar{R}$ a given revenue requirement, and can be interpreted as required revenue for essential public goods. At the optimum, this constraint holds with equality.

The first term on the RHS of expression (4a) represents government revenue without income shifting, using the reported earnings density $\hat{n}_i$. The second term shows government revenue from the capital income tax system by taxing shifted income, i.e. labour income reported as capital income. The first term on the RHS of expression (4b) represents government revenue when no income shifting takes place, as the true earnings density is used in the summation. The second term shows the revenue loss due to income shifting. This term can also be summed differently, as demonstrated in expression (4c).

Let λ denote the Lagrange multiplier on the government's budget constraint, which is also the marginal social cost of public funds. I denote the social marginal value of income by β, which depends on the individual type. For non-shifters β only depends on y_i. The social marginal value of income for non-shifters earning y_i is $\beta_i = \frac{W'(U_i)}{\lambda}$, where $U_i = u(e - s_i) + c_i$. For shifters, β depends on y_i, k, and m. The social marginal value of income for shifters earning y_{i+j} and reporting y_i is denoted by $\beta_i^{i+j} = \frac{W'(U_i^{i+j})}{\lambda}$ (which is a decreasing function of k and m), where $U_i^{i+j} = u(e - s_i) + c_i^{i+j}$.

The average β for those reporting y_i is

$$\beta_i = E\left[\frac{W_i'}{\lambda}\right],$$

where W_i' is either $W'(U_i)$ or $W'(U_i^{i+j})$. The expectation operator averages over w, k, and m. The average social marginal value of income for those reporting y_i matters because the labour income tax only depends on reported labour income.

A variable in the optimality conditions is the average β for those reporting y_i and higher,

$$\beta_i^+ = \frac{\sum_{h=i}^{I} \beta_h \hat{n}_h}{\sum_{h=i}^{I} \hat{n}_h},$$

this captures the average social marginal value of increasing second-period consumption for individuals reporting y_i or higher.

The average β for those earning y_{i+j} and reporting y_i is denoted by

$$\beta_i^{i+j} = E\left[\frac{W'(U_i^{i+j})}{\lambda}\right],$$

where the expectation operator averages over k and m.

4 Behavioural Elasticities

This section characterises (i) savings responses to the labour-income-contingent capital tax t_i, (ii) the intensive shifting choice j via cutoffs in m, and (iii) the extensive shifting choice via cutoffs in k. These objects deliver the elasticities ζ, ε, and η that enter Proposition 1.

The necessary condition associated with savings is

$$u'(e - s_i) = 1 + r(1 - t_i), \tag{5}$$

where s_i is the optimal saving choice of anyone who ends up reporting y_i. From this condition, savings depend only on e and the after-tax rate of return. Since r and e are fixed, the after-tax return varies only with t_i. The optimal level of savings for an individual facing tax rate t_i is denoted by $s_i = s(1 - t_i)$. The behavioural response for savings is captured by the elasticity of savings with respect to $1 - t_i$,

$$\zeta_i = \frac{\partial s_i}{\partial(1 - t_i)} \frac{1 - t_i}{s_i} = -\frac{r(1 - t_i)}{u'' s_i} > 0, \tag{6}$$

where the final expression is found by implicitly differentiating (5). As there are no income effects, ζ_i is a compensated elasticity. An increase in t_i will reduce savings as individuals will substitute second-period consumption for first period consumption. From (6), it follows that ζ_i only depends on t_i.

The extensive shifting decision (i.e. whether to be a shifter or a non-shifter) depends on the intensive shifting decision. I therefore begin by describing the intensive shifting decision. Individuals who earn y_i, have type m, and face a sufficiently low fixed cost k to become shifters have to decide how much to shift, i.e. choose j. The level of shifting depends on w and m. Individuals with the lowest possible m will shift the most. Individuals with a slightly higher m may shift the same amount, but once m has exceeded a certain limit, individuals will shift one income level less. This means that there will be cutoff values between adjacent values of j. I define $\hat{m}_i^{i+j}$ as the value of m at which shifters earning y_{i+j} are indifferent between reporting y_i and y_{i+1}. Those with $m > \hat{m}_i^{i+j}$ optimally report y_{i+1}, whereas those with $m < \hat{m}_i^{i+j}$ optimally report y_i. The following condition defines $\hat{m}_i^{i+j}$,

$$\begin{aligned} &u(e-s_i)+s_i(1+r)-T_i^{i+j}-d(j\delta,\hat{m}_i^{i+j}) \\ &=u(e-s_{i+1})+s_{i+1}(1+r)-T_{i+1}^{i+j}-d((j-1)\delta,\hat{m}_i^{i+j}), \end{aligned} \tag{7}$$

which simply states that the utility of reporting y_i and y_{i+1} for those earning y_{i+j} is the same. Condition (7) implicitly defines $\hat{m}i^{i+j}$ as an increasing function of $Ti+1^{i+j}-T_i^{i+j}$. For $j=1$, condition (7) is slightly modified because the indifference condition includes the fixed cost k.[5]

An individual will decide to become a shifter if their k is sufficiently low. The threshold level of k depends on the maximum gain from shifting (evaluated at the optimal intensive choice of j). An individual who earns y_i and has type m will become a shifter if k is less than the following level:

$$\hat{k}_i^{i+j}=[u(e-s_i)-u(e-s_{i+j})]+(s_i-s_{i+j})(1+r)+(T_{i+j}-T_i^{i+j})-d(j\delta,m_i^{i+j}), \tag{8}$$

which is decreasing in m and increasing in the tax savings of shifting income, i.e. $T_{i+j}-T_i^{i+j}$. The range of m for which individuals earning y_{i+j} report y_i (as shifters) is $m_i^{i+j}\in[\hat{m}_i^{i+j},\hat{m}_{i+1}^{i+j}]$. Those with $k<\hat{k}_i^{i+j}$ will become shifters while those with $k>\hat{k}_i^{i+j}$ will become non-shifters.

[5]The condition is: $u(e-s_i)+s_i(1+r)-T_i^{i+1}-d(\delta,\hat{m}_i^{i+1})-k=u(e-s_{i+1})+s_{i+1}(1+r)-T_{i+1}$.

Having defined the optimal intensive and extensive shifting decisions for individuals, I next turn to the distribution of reported income. For any i and j, such that $i \geq 2$ and $i + j \leq I$, the share of people who earn y_{i+j} and report y_i is

$$\alpha_i^{i+j} = \int_{w_{i+j}}^{w_{i+j+1}} \int_{\underline{k}}^{\hat{k}_i^{i+j}} \int_{\hat{m}_{i-1}^{i+j}}^{\hat{m}_i^{i+j}} f(w,k,m)\, dw\, dk\, dm, \tag{9}$$

which defines α_i^{i+j} as an increasing function of $T_{i+j} - T_i^{i+j}$ and $T_{i+1}^{i+j} - T_i^{i+j}$ through $\hat{k}_i^{i+j}$ and $\hat{m}_i^{i+j}$, respectively, and a decreasing function of $T_i^{i+j} - T_{i-1}^{i+j}$ through $\hat{m}_{i-1}^{i+j}$. That is:[6]

$$\alpha_i^{i+j}(\underbrace{T_{i+j} - T_i^{i+j}}_{+}, \underbrace{T_{i+1}^{i+j} - T_i^{i+j}}_{+}, \underbrace{T_i^{i+j} - T_{i-1}^{i+j}}_{-}).$$

Using the above notation, the share of the population reporting y_i is:

$$\hat{n}_i = \left[n_i - \sum_j \alpha_{i-j}^i \right] + \sum_j \alpha_i^{i+j}, \tag{10}$$

where the first term represents the non-shifters earning y_i, while the second term represents shifters earning y_{i+j} and reporting y_i.

Changes in the tax schedules will have effects on the intensive margin through $\hat{m}$ and on the extensive margin through $\hat{k}$. The behavioural elasticity of income shifting along the intensive margin is defined by

$$\varepsilon_i^{i+j} = \frac{\partial \alpha_i^{i+j}}{\partial \hat{m}_i^{i+j}} \frac{\partial \hat{m}_i^{i+j}}{\partial (T_{i+1}^{i+j} - T_i^{i+j})} \left[T_{i+1}^{i+j} - T_i^{i+j} \right] \tag{11a}$$

$$= \frac{\int_{w_{i+j}}^{w_{i+j+1}} \int_{\underline{k}}^{\hat{k}_i^{i+j}} f(w,k,\hat{m}_i^{i+j}\, dw\, dk}{d_2\left(j\delta, \hat{m}_i^{i+j}\right) - d_2\left((j-1)\delta, \hat{m}_i^{i+j}\right)} \left[T_{i+1}^{i+j} - T_i^{i+j} \right] > 0, \tag{11b}$$

[6] Similarly, α_{i-j}^i is increasing in $T_i - T_{i-j}^i$ and $T_{i-j+1}^i - T_{i-j}^i$ through $\hat{k}_{i-j}^i$ and $\hat{m}_{i-j}^i$, respectively, and decreasing in $T_{i-j}^i - T_{i-j-1}^i$ through $\hat{m}_{i-j-1}^i$. That is, $\alpha_i^{i+j}(\underbrace{T_i - T_{i-j}^i}_{+}, \underbrace{T_{i-j+1}^i - T_{i-j}^i}_{+}, \underbrace{T_{i-j}^i - T_{i-j-1}^i}_{-})$.

where $d_2(\cdot)$ denotes the partial derivative of $d(\cdot)$ with respect to m. The denominator in (11b) is positive since the cross-derivative of $d(\cdot)$ is positive. This elasticity measures the percentage change in α_i^{i+j} (the share of the population who earn y_{i+j} and report y_i) induced by a one-percent increase in the marginal shifting incentive $T_{i+1}^{i+j} - T_i^{i+j}$. The tax savings from shifting down by one additional income level, $T_{i+1}^{i+j} - T_i^{i+j}$, can be expressed in terms of the marginal tax rate on labour income, MTR_i^L, and the change in capital income tax payments induced by reporting higher labour income:

$$\begin{aligned} T_{i+1}^{i+j} - T_i^{i+j} &= \delta MTR_{i+1}^L + a_{i+1} + [t_{i+1}(j-1) - t_i j]\delta, \\ MTR_i^L &= \frac{\tau_i y_i - \tau_{i-1} y_{i-1}}{\delta}, \quad a_i = t_i r s_i - t_{i-1} r s_{i-1}. \end{aligned} \tag{12}$$

The intensive shifting elasticity depends on three factors. First, the elasticity increases as the tax system becomes more progressive. This applies to both the labour income tax and the capital income tax. When τ_i and t_i increase rapidly with income, more individuals will report lower labour income in response to an increase in $T_{i+1}^{i+j} - T_i^{i+j}$.

Second, it depends on the cost structure $d(j\delta, m)$. The greater the convexity in the cost structure, the higher the cost of a marginal increase in income shifting. Conversely, as the cost function becomes less convex, more individuals will report lower labour income in response to an increase in $T_{i+1}^{i+j} - T_i^{i+j}$.

Third, it depends on the distribution of w, k and m. The elasticity is larger the greater the mass of shifters reporting y_i who are near the margin between shifting slightly less and slightly more. If income shifting is mostly present at the top of the income distribution, then intensive income shifting behaviour mostly calls for a reduction in labour tax rates at the top of the income distribution.

The behavioural elasticity of income shifting along the extensive margin is defined by

$$\eta_i^{i+j} = \frac{\partial \alpha_i^{i+j}}{\partial \hat{k}_i^{i+j}} \frac{\partial \hat{k}_i^{i+j}}{\partial (T_{i+j} - T_i^{i+j})} \left[T_{i+j} - T_i^{i+j}\right] \tag{13a}$$

$$= \int_{w_{i+j}}^{w_{i+j+1}} \int_{\hat{m}_{i-1}^{i+j}}^{\hat{m}_{i}^{i+j}} f(w, \hat{k}_i^{i+j}, m)\, dw\, dm \left[T_{i+j} - T_i^{i+j}\right] > 0, \tag{13b}$$

where the relevant range of m is $m \in [\hat{m}_{i-1}^{i+j}, \hat{m}_i^{i+j}]$. This elasticity measures the percentage change in α_i^{i+j} induced by a one-percent increase in the incentive to shift $(T_{i+j} - T_i^{i+j})$ increases by one percent.

The extensive shifting elasticity depends on two factors. First, like the intensive shifting elasticity, it increases as the tax system becomes more progressive. For the extensive shifting elasticity, what matters is the overall progressivity of the tax system, which refers to the tax difference between the earned and reported income of shifters, rather than the difference in tax payments for adjacent income levels, as is the case for the intensive shifting elasticity. The overall progressivity can be reduced either by lowering taxes for individuals reporting more than y_i or by raising taxes for individuals reporting y_i.

Second, it depends on the distribution of w, k, and m. The higher the mass of potential shifters earning more than y_i, the greater the extensive shifting elasticity at income level i. Potential shifters refer to individuals who are indifferent between being shifters and non-shifters, i.e. those with k close to $\hat{k}_i^{i+j}$, $w \in [w_{i+j}, w_{i+j+1}]$, and $m \in [\hat{m}_{i-1}^{i+j}, \hat{m}_i^{i+j}]$. What matters is how much potential shifters shift, i.e. which j is chosen by individuals who are indifferent between shifting and not shifting. If potential shifters earning above y_i have sufficiently low m, they will shift a lot.

The intensive and extensive shifting elasticities both depend on where, in the income distribution, the marginal individuals for intensive shifting and for entry into shifting are located, respectively. Since income shifting is predominantly observed among higher income groups, these elasticities are expected to be highest in these groups. However, it is less clear whether the intensive or the extensive elasticity should be larger at a given income level.

5 Optimal Taxation

This section characterises optimal labour and capital income taxation in the presence of heterogeneous access to income shifting. I first present a benchmark without income shifting, which highlights the role of the quasi-linear structure and provides a point of comparison. I then introduce income shifting and derive the optimal tax formulas.

5.1 No Income Shifting

As a baseline model, I start by solving the government's problem where income shifting is not possible. In this model, individuals earn y_i and only need to decide how much to save. Unlike the model described above, all individuals who report y_i are homogeneous.[7] This simplifies the model. The government solves:

$$\underset{\{\tau_i, t_i\}}{\text{Max}} \qquad \Omega = \sum_i^I n_i W(U_i), \quad \text{subject to} \quad \sum_i^I n_i[\tau_i y_i + t_i r s_i] \geq \overline{R} \tag{14}$$

The optimal tax system is characterised by

$$1 - \beta_i = 0, \quad \text{and} \quad t_i = 0, \quad \forall i, \tag{15}$$

which states that the government will equalise consumption of all individuals using the labour income tax system while capital income remains untaxed. As labour income is exogenous, the government has essentially access to a lump sum tax, and therefore it should not come as a surprise that the government will completely redistribute income. The capital income tax has the same effects as the labour income tax, except that the capital income tax also distorts the savings decision. Therefore, using τ_i is superior to t_i and capital income should not be taxed, as the result from the Atkinson & Stiglitz (1976) model.

Because labour income is exogenous, the nonlinear labour income tax can replicate type-

[7] In the model w is heterogeneous among individuals earning y_i, this does however not affect utility. Hence, this heterogeneity does not affect the model.

dependent lump-sum transfers across income levels, so the government can fully implement its redistributive objectives without inducing real behavioural distortions. Moreover, since capital income arises from saving, taxing capital income would create a pure distortion to saving with no additional redistributive benefits.

5.2 With Income Shifting

This subsection derives the main optimal tax results when income shifting is possible. The government chooses a nonlinear labour income tax schedule and a labour-income-contingent linear tax on capital income, $\{\tau_i, t_i\}_{i=1}^{I}$, subject to a revenue requirement. Behavioural responses enter the optimality conditions through the savings elasticity ζ_i, the intensive shifting elasticities ε_i^{i+j}, and the extensive shifting elasticities η_i^{i+j}. Proposition 1 characterises the optimal labour and capital income tax schedules in terms of these objects, and the subsequent discussion interprets the resulting wedges and the roles of the intensive and extensive margins.

Proposition 1 *The optimal labour income tax schedule is characterised by*

$$\frac{\partial \mathcal{L}/\hat{\lambda}}{\partial MTR_i^L} = (1-\beta_i^+) - \overline{\varepsilon}_{i-1} - \overline{\eta}_{i-1} = 0, \tag{16}$$

for $i \geq 2$, where $MTR_i^L = \frac{\Delta\tau_i y_i}{\delta}$ is the marginal tax rate on labour income, the average intensive shifting elasticity for those earning y_{i+j} or more is denoted by $\overline{\varepsilon}_{i-1} = \sum_{j=0}^{I-i} \varepsilon_{i-1}^{i+j}/\hat{N}_i > 0$, the average extensive shifting elasticity for those earning y_{i+j} or more is denoted by $\overline{\eta}_{i-1} = \sum_{j=0}^{I-i} \eta_{i-1}^{i+j}/\hat{N}_i > 0$, $\hat{\lambda} = \lambda \hat{N}_i$, λ is the Lagrange multiplier on the government's budget constraint, $\hat{N}_i$ denotes the share of individuals reporting more than y_i.

The optimal capital income tax schedule is characterised by

$$\frac{t_i}{1-t_i} = \sum_{j=1}^{I-i} \frac{j\hat{\delta}_i}{\hat{n}_i \zeta_i} \left[\alpha_i^{i+j}\left(1-\beta_i^{i+j}\right) + \eta_i^{i+j} + \left(\varepsilon_i^{i+j} - \varepsilon_{i-1}^{i+j}\right)\right], \tag{17}$$

where $\hat{\delta}_i = \frac{\delta}{rs_i}$.

The labour income tax condition (16) depends on income shifting only through the total responsiveness of reported income at the relevant boundary, which combines intensive- and extensive-margin shifting incentives. By contrast, the optimal capital income tax condition (17) reflects three components: (i) a redistributive motive that depends on how the social marginal value of income differs between shifters and non-shifters, (ii) an extensive-margin revenue effect operating through entry into shifting (weighted by the amount shifted), and (iii) an intensive-margin reallocation effect reflecting how a change in t_i shifts mass across adjacent reported income levels. The discussion below considers these components in turn.

Labour income tax rate In the literature, optimality conditions often have τ^i explicitly on the LHS, while τ^i implicitly appears on the RHS (see e.g. Diamond, 1998). In (16), however, τ^i appears only implicitly on both sides. The condition represents the welfare effect of raising the marginal tax rate (MTR_i^L) on y_i. This redistributes resources from those reporting y_i and above to those reporting less.

Condition (16) states that $\beta_i^+ < 1$ which means that the social benefit of increasing consumption of those reporting y_i and more by one unit is on average less than one, which is the direct cost from increasing their consumption by one unit, on average. This means that the government will not completely equalise income, as is the case when income shifting is not possible.[8]

The optimal marginal tax rate on labour income is positively associated with β_i^+. An increase in MTR_i^L will reduce the consumption of those reporting y_i and above, thereby raising their social marginal value of income, i.e., increasing their β_i. In the first-best scenario, marginal tax rates are sufficiently high such that everyone's social marginal value of income equals one, i.e. $\beta_i = 1\ \forall i$. The higher $1 - \beta_i^+$ is, the lower the MTR_i^L. This implies that the optimal marginal tax rate negatively depends on the income shifting elasticities. As the

[8]The average β for the whole population is 1, $\beta_1^+ = 1$, which states that the benefit of giving everybody one more unit of consumption is equal to its cost, which is a standard result from optimal tax theory.

share of individuals (shifters and non-shifters) who report y_i or more increasingly report y_{i-1} in response to a marginal increase in MTR_i^L, the lower is the optimal MTR_i^L. The intensive and extensive responses enter symmetrically the optimal labour income tax. That is, what matters is what share of those reporting y_i or more report y_{i-1} in response to an increase in taxes.

A change in the labour tax at y_i affects the mass of individuals who end up *reporting* y_i rather than y_{i-1}. This movement across reported income levels can occur in two ways: existing shifters may report y_{i-1} instead of y_i (the intensive margin), and additional individuals may choose to become shifters by paying the fixed cost k and report y_{i-1} instead of y_{i+j} (the extensive margin). Both mechanisms change the reported earnings density at the same boundary (y_{i-1}) and therefore have equivalent implications for labour tax revenue and for the incentive-compatibility constraints. As a result, the optimal labour tax schedule is governed by the total responsiveness of reported income at that boundary, which is captured by the sum of the intensive and extensive shifting elasticities.

An increase in τ_i will make it more attractive for non-shifters earning y_i to become shifters and report less than y_i, and for shifters who report y_i to shift slightly more and report y_{i-1} instead of y_i. However, it will also make it less attractive for shifters who report y_i to remain shifters, leading some to become non-shifters, and for shifters who report y_i to shift slightly less and report y_{i+1} instead of y_i. Nevertheless, when the labour tax rate on income at y_i and above increases, the latter effect cancels out. The only remaining effect is an increased incentive to report y_{i-1} instead of y_i.

Income shifting affects the optimal labour income tax schedule through where in the *reported* income distribution shifting behaviour is most prevalent. In particular, both the intensive and extensive shifting elasticities are largest at reported income levels where a substantial mass of individuals is close to the relevant cutoffs—either close to adjusting the amount shifted (the intensive margin) or close to entering shifting by paying the fixed cost (the extensive margin). Since empirical evidence suggests that shifting opportunities are concentrated among high-income owner-managers, one might expect these elasticities to be most pronounced at the upper end of the reported income distribution, which tends to push towards lower labour tax rates at the top. However, this prediction depends on how shifting maps true income into reported income and therefore on where shifters appear in the reported distribution.

If high-income individuals shift only modestly, they remain concentrated in high reported income brackets and the shifting elasticities will primarily matter near the top. By contrast, if high earners shift sufficiently to report substantially lower labour income, then the density of shifters may be sizeable at lower reported income levels, implying that labour tax rates at those levels also affect shifting behaviour and tax revenue. In that case, the labour tax schedule may need to be more progressive in reported income even if shifting is concentrated among high earners in terms of true income. This mechanism, however, requires large shifts in reported income (large j), and whether such substantial reclassification is empirically important is ultimately an empirical question. The condition (16) is a standard inverse elasticity rule stating that the optimal MTR_i^L is inversely related to income shifting elasticities. As more redistribution from those reporting y_i and more to those reporting less will cause increased income shifting, less redistribution should be carried out. These findings are not surprising and consistent with previous studies (see e.g., Piketty et al., 2014).

Capital income tax rate The optimal capital income tax depends on three factors. First, redistributive motive from shifters to non-shifters, represented by the first term in the bracket in (17). Imagine a tax reform where t_i is increased by $1/rs_i$ accompanied by a decrease in τ_i by $1/y_i$. This will not have a direct effect on non-shifters, as tax payments by non-shifters (T_i) remain unchanged. But the tax reform will make shifters worse off, as tax payments by shifters ($T_i^{i+j} = T_i + t_i\delta j$) increase. This is desirable when the social marginal value of income for shifters reporting y_i is lower than for the whole population (which is 1 at the optimum). The social marginal value of income decreases with income and is lower for shifters than for non-shifters, all else equal. That is, the social marginal value of income for shifters earning y_{i+j} and reporting y_i (β_i^{i+j}) is decreasing in i and j. If income shifting is concentrated in higher income groups, then $1 - \beta_i^{i+j} > 0$, which calls for a positive capital income tax rate. However, if the welfare function is sufficiently concave, the term will be small, as redistribution between shifters and non-shifters will have negligible welfare effects. In contrast, if they engage in so much income shifting such that their reported income is low, this term might be higher.

Second, the optimal t_i is positively associated with the weighted sum of the extensive shifting elasticities for individuals earning $y \geq y_i$ and reporting y_i in response to a marginal increase in t_i. This relationship explains the second term on the RHS of (17). The weight is j because it reflects the revenue effect: when shifters become non-shifters, government revenues increase by $j\delta$. Imagine again the reform discussed above, an increase in t_i accompanied by an increase in τ_i, which only has a direct effect on shifters. This reduces the incentive to participate in shifting and can induce some individuals to stop being shifters. The magnitude of this exit response is governed by the extensive shifting elasticity at reported income level y_i, i.e. η_i^{i+j} (aggregated over j in (17)).

Third, the optimal t_i depends positively on ε_i^{i+j} and negatively on ε_{i-1}^{i+j}. Imagine again the reform discussed above, an increase in t_i accompanied with an increase in τ_i, which only has direct effect on shifters. Such an increase discourages individuals from reporting y_i. This will lead some to shift less, i.e. report y_{i+1} instead of y_i. However, the tax reform will while encourage some to shift *more*, i.e., report y_{i-1} instead of y_i. Which effect is stronger is generally ambiguous. It depends on the progressivity of the tax system, the cost function $d(j\delta, m)$, and the joint distribution of w, k, and m. The discussion in Appendix A.2 shows that it is ambiguous whether $\varepsilon_i^{i+j} \lesseqgtr \varepsilon_{i-1}^{i+j}$. That is, as shifters shift more income, it is not clear whether their intensive shifting responsiveness tends to be smaller or larger.

The term $\varepsilon_i^{i+j} - \varepsilon_{i-1}^{i+j}$ captures the *net* intensive-margin reallocation across the reported-income boundary at y_i induced by a change in t_i. On the one hand, raising t_i increases the tax burden on individuals who report y_i, which tends to encourage additional shifting and pushes some shifters from reporting y_i down to y_{i-1} (an inflow into α_{i-1}^{i+j} from α_i^{i+j}). On the other hand, the same change also alters the wedge between reporting y_i and y_{i+1}, which can push some shifters who previously reported y_i to instead report y_{i+1} (an inflow into α_{i+1}^{i+j} from α_i^{i+j}). The difference $\varepsilon_i^{i+j} - \varepsilon_{i-1}^{i+j}$ therefore measures whether the intensive response to t_i leads, on net, to less income shifting (calling for a higher tax rate) or to more income shifting (calling for a lower tax rate). Which effect dominates depends on local progressivity in the labour and capital tax schedules, the curvature of the shifting cost function, and the density of shifters near the relevant cutoffs.

To clarify the role of intensive-margin responses, consider the special case in which shifting elasticities are constant across reported income levels, i.e. $\varepsilon_i^{i+j} = \varepsilon_{i-1}^{i+j}$. Then condition (17) becomes

$$\frac{t_i}{1-t_i} = \sum_{j=1}^{I-i} \frac{\hat{\delta}_i}{\hat{n}_i \zeta_i} \left[j\alpha_i^{i+j} \left(1 - \beta_i^{i+j}\right) + j\eta_i^{i+j} \right], \tag{17$'$}$$

for $i \geq 2$. That is, when the capital tax rate is contingent on reported labour income, the intensive-margin term in (17) enters through differences of elasticities across adjacent reported brackets, $\varepsilon_i^{i+j} - \varepsilon_{i-1}^{i+j}$. If intensive shifting elasticities are constant across reported income levels, $\varepsilon_i^{i+j} = \varepsilon_{i-1}^{i+j}$ for all $i \geq 2$, the intensive income shifting margin effect vanishes. In that case, the optimal capital tax depends only on the redistributive motive (how the social marginal value of income differs between shifters and non-shifters) and on the extensive-margin response captured by the η-elasticities.

In general, however, intensive responses need not cancel out in the aggregate. As shown in (24) in Appendix A.3, the average intensive-margin effect tends to push towards higher capital taxation. A key reason is the asymmetry at the bottom of the reported income distribution: at the lowest reported income level, an increase in t_1 can only induce shifters to report higher labour income (shift less, moving from y_1 to y_2).

Appendix A.3 derives the optimal *linear* capital income tax rate, i.e. the case in which $t_i = t$ is constant across reported labour income levels, see (23). In this setting, the optimality condition shows that the intensive shifting elasticity enters with a positive sign: a higher responsiveness of shifting at the margin raises the marginal fiscal gain from taxing reported capital income and therefore pushes towards a higher t. In the expression for the average optimal tax rate in (24), the intensive-margin term aggregates across groups without the $j\delta$ weights, whereas the extensive-margin term is weighted with how much income is shifted ($j\delta$). This highlights that extensive-margin responses have a more pronounced effect on the optimal capital income tax rate than the intensive-margin response.

The intensive and extensive shifting elasticities enter the optimal labour income tax condition symmetrically: what matters is the total responsiveness of reported income at a given boundary, captured by the sum of the intensive and extensive elasticities, as discussed above. This symmetry does not carry over to the optimal capital income tax. A change in t_i affects revenue through the shifted tax base. On the extensive margin, some individuals stop shifting altogether; the associated revenue effect is proportional to the total amount they previously reclassified, $j\delta$. On the intensive margin, existing shifters adjust the amount shifted at the margin, so the revenue effect is proportional to δ (a one-step change in reported income) rather than $j\delta$. Hence, when the capital income tax is linear in reported capital income, extensive-margin responses typically receive greater weight in the capital-tax optimality condition than in the labour-tax condition.

Comparing (16) and (17) shows that stronger shifting incentives push the government to reduce the tax wedge between labour and capital income: higher shifting elasticities tend to lower the optimal labour income tax rate and raise the optimal capital income tax rate at a given reported income level. This implication, that income shifting calls for narrowing the labour-capital tax differential, accords with the optimal-tax literature on shifting technologies, where shifting opportunities reduce the optimal gap between tax rates on labour and capital incomes (Christiansen & Tuomala, 2008; Pirttilä & Selin, 2011).

When the intensive and extensive shifting elasticities increase with reported income, income shifting causes the labour tax schedule to become less progressive and the capital income tax system more progressive with respect to labour income. Unlike the labour tax condition, however, the capital tax condition weights the extensive margin more heavily, because entry and exit from shifting change the shifted base by $j\delta$ rather than by δ at the margin. Hence, where shifting incentives and the density of shifters are larger, the government has stronger reasons to compress the labour-capital tax wedge, but the strength of this force depends especially on extensive-margin participation.

An important qualification is that the relevant distribution is the distribution of *reported* income among shifters. If high earners shift sufficiently to appear in relatively low reported labour income brackets (large j), then the strongest shifting pressures may arise lower down in the reported distribution, potentially raising (lowering) the tax rate on capital (labour) income at lower reported income levels.

6 Conclusion

This paper studies optimal labour and capital income taxation when individuals differ in their access to income shifting opportunities. Income shifting is modelled as a combination of an intensive margin—how much income is reclassified—and an extensive margin—whether individuals enter shifting by paying a fixed cost, such as through incorporation or organisational changes.

In the model, the capital income tax rate is labour-income contingent. This enables the government to target shifting incentives directly and to compress the labour–capital tax wedge where shifting pressures are strongest. When shifting elasticities increase with reported labour income, the optimal capital income tax schedule is progressive in reported labour income, even though capital income itself is taxed linearly within each bracket.

A key asymmetry emerges between the intensive and extensive margins of income shifting. While both margins affect the optimal labour income tax symmetrically, the extensive margin plays a disproportional role for capital income taxation. Entry into or exit from shifting changes the entire shifted tax base, whereas intensive adjustments affect only marginal reclassification. As a result, participation responses receive greater weight in the optimal capital tax schedule.

References

Alstadsæter, A., & Jacob, M. (2016). Dividend taxes and income shifting. *The Scandinavian Journal of Economics*, *118*(4), 693–717.

Alstadsæter, A., & Jacob, M. (2017). Who participates in tax avoidance? evidence from swedish microdata. *Applied Economics*, *49*(28), 2779–2796.

Alstadsæter, A., Kopczuk, W., & Telle, K. (2014). Are closely held firms tax shelters? *Tax Policy and the Economy*, *28*(1), 1–32.

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

Banks, J., & Diamond, P. (2010). The base for direct taxation. In *Dimensions of tax design: The mirrlees review* (pp. 548–648). Oxford University Press.

Barth, E., & Ognedal, T. (2018). Tax evasion in firms. *Labour*, *32*(1), 23–44.

Bastani, S., & Waldenström, D. (2020). How should capital be taxed? *Journal of Economic Surveys*, *34*(4), 812–846.

Christiansen, V., & Tuomala, M. (2008). On taxing capital income with income shifting. *International Tax and Public Finance*, *15*, 527–545.

DeBacker, J., Heim, B. T., Ramnath, S. P., & Ross, J. M. (2019). The impact of state taxes on pass-through businesses: Evidence from the 2012 kansas income tax reform. *Journal of Public Economics*, *174*, 53–75.

De Mooij, R. A., & Nicodème, G. (2008). Corporate tax policy and incorporation in the eu. *International Tax and Public Finance*, *15*, 478–498.

Devereux, M. P., Liu, L., & Loretz, S. (2014). The elasticity of corporate taxable income: New evidence from uk tax records. *American Economic Journal: Economic Policy*, *6*(2), 19–53.

Diamond, P. (1998). Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review*, 83–95.

Diamond, P., & Spinnewijn, J. (2011). Capital income taxes with heterogeneous discount rates. *American Economic Journal: Economic Policy*, *3*(4), 52–76.

Edmark, K., & Gordon, R. H. (2013). The choice of organizational form by closely-held firms in sweden: tax versus non-tax determinants. *Industrial and corporate Change*, *22*(1), 219–243.

Elschner, C. (2013). Special tax regimes and the choice of organizational form: Evidence from the european tonnage taxes. *Journal of Public Economics*, *97*, 206–216.

Feldstein, M. (1995). The effect of marginal tax rates on taxable income: a panel study of the 1986 tax reform act. *Journal of Political Economy*, *103*(3), 551–572.

Fuest, C., & Huber, B. (2001). Labor and capital income taxation, fiscal competition, and the distribution of wealth. *Journal of Public Economics*, *79*(1), 71–91.

Fuest, C., & Huber, B. (2005). The effect of income shifting on the efficiency properties of consumption-tax systems. *FinanzArchiv/Public Finance Analysis*, 139–153.

Goolsbee, A. (1998). Taxes, organizational form, and the deadweight loss of the corporate income tax. *Journal of Public Economics*, *69*(1), 143–152.

Goolsbee, A. (2004). The impact of the corporate income tax: evidence from state organizational form data. *Journal of Public Economics*, *88*(11), 2283–2299.

Gordon, R., & MacKie-Mason, J. K. (1994). Tax distortions to the choice of organizational

form. *Journal of Public Economics*, *55*(2), 279–306.
Gordon, R., & Slemrod, J. (2000). Are "real" responses to taxes simply income shifting between corporate and personal tax bases? In (chap. 8). Harvard University Press.
Harju, J., & Matikka, T. (2016). The elasticity of taxable income and income-shifting: what is "real" and what is not? *International Tax and Public Finance*, *23*, 640–669.
Jacobs, B. (2013). From optimal tax theory to applied tax policy. *FinanzArchiv/Public Finance Analysis*, 338–389.
Mackie-Mason, J. K., & Gordon, R. H. (1997). How much do taxes discourage incorporation? *The Journal of Finance*, *52*(2), 477–506.
Miller, H., Pope, T., & Smith, K. (2024). Intertemporal income shifting and the taxation of business owner-managers. *Review of Economics and Statistics*, *106*(1), 184–201.
Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The review of economic studies*, *38*(2), 175–208.
Piketty, T., Saez, E., & Stantcheva, S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American economic journal: economic policy*, *6*(1), 230–271.
Pirttilä, J., & Selin, H. (2011). Income shifting within a dual income tax system: Evidence from the finnish tax reform of 1993. *Scandinavian Journal of Economics*, *113*(1), 120–144.
Reis, C. (2011). Entrepreneurial labor and capital taxation. *Macroeconomic Dynamics*, *15*(3), 326–335.
Romanov, D. (2006). The corporation as a tax shelter: Evidence from recent israeli tax changes. *Journal of Public Economics*, *90*(10-11), 1939–1954.
Saez, E. (2002). Optimal income transfer programs: intensive versus extensive labor supply responses. *The quarterly journal of economics*, *117*(3), 1039–1073.
Saez, E. (2004). Reported incomes and marginal tax rates, 1960-2000: evidence and policy implications. *Tax policy and the economy*, *18*, 117–173.
Saez, E., Slemrod, J., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of economic literature*, *50*(1), 3–50.
Selin, H. (2025). Taxing dividends in a dual income tax system: The nordic experience with the income splitting rules. *Nordic Tax Journal*, *2024*(s1), 83–94.
Selin, H., & Simula, L. (2020). Income shifting as income creation? *Journal of Public Economics*, *182*, 104081.
Slemrod, J. (1995). Income creation or income shifting? behavioral responses to the tax reform act of 1986. *The American Economic Review*, *85*(2), 175–180.
Slemrod, J. (1996). High-income families and the tax changes of the 1980s: the anatomy of behavioral response. In *Empirical foundations of household taxation* (pp. 169–192). University of Chicago Press.
Tazhitdinova, A. (2020). Are changes of organizational form costly? income shifting and business entry responses to taxes. *Journal of Public Economics*, *186*, 104187.

A Appendix

A.1 Proof of proposition 1

The necessary condition for τ_i is

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \tau_i}\frac{1}{y_i\lambda} = \hat{n}_i(1-\beta_i) &- \sum_{j=1}^{I-i+1} \frac{\partial \alpha_{i-1}^{i-1+j}}{\partial \hat{m}_{i-1}^{i-1+j}} \frac{\partial \hat{m}_{i-1}^{i-1+j}}{\partial (T_i^{i-1+j} - T_{i-1}^{i-1+j})}(T_{i-1+j} - T_{i-1}^{i-1+j}) \\
&- \sum_{j=1}^{I-i} \frac{\partial \alpha_i^{i+j}}{\partial \hat{m}_{i-1}^{i+j}} \frac{\partial \hat{m}_{i-1}^{i+j}}{\partial (T_i^{i+j} - T_{i-1}^{i+j})}(T_{i+j} - T_i^{i+j}) \\
&+ \sum_{j=1}^{I-i} \frac{\partial \alpha_i^{i+j}}{\partial \hat{m}_i^{i+j}} \frac{\partial \hat{m}_{i-1}^{i+j}}{\partial (T_{i+1}^{i+j} - T_i^{i+j})}(T_{i+j} - T_i^{i+j}) \\
&+ \sum_{j=1}^{I-i-1} \frac{\partial \alpha_{i+1}^{i+1+j}}{\partial \hat{m}_i^{i+j}} \frac{\partial \hat{m}_i^{i+1+j}}{\partial (T_{i+1}^{i+1+j} - T_i^{i+1+j})}(T_{i+1+j} - T_{i+1}^{i+1+j}) \\
&+ \sum_{j=1}^{I-i} \frac{\partial \alpha_i^{i+j}}{\partial \hat{k}_i^{i+j}} \frac{\partial \hat{k}_i^{i+j}}{\partial (T_{i+j} - T_i^{i+j})}(T_{i+j} - T_i^{i+j}) \\
&- \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-j}^{i}}{\partial \hat{k}_{i-j}^{i}} \frac{\partial \hat{k}_{i-j}^{i}}{\partial (T_i - T_{i-j}^{i})}(T_i - T_i^{i-j}).
\end{aligned}$$

Next, I use the fact that $\frac{\partial \alpha_i^{i+j}}{\partial \hat{m}_{i-1}^{i+j}} = -\frac{\partial \alpha_{i-1}^{i+j}}{\partial \hat{m}_{i-1}^{i+j}}$, that $T_i^i = T_i$, the second to the fifth term on the right-hand side become

$$\begin{aligned}
-\varepsilon_{i-1}^{i} \quad &- \sum_{j=2}^{I-i+1} \frac{\partial \alpha_{i-1}^{i-1+j}}{\partial \hat{m}_{i-1}^{i-1+j}} \frac{\partial \hat{m}_{i-1}^{i-1+j}}{\partial (T_i^{i-1+j} - T_{i-1}^{i-1+j})}(T_{i-1+j} - T_{i-1}^{i-1+j}) \\
&+ \sum_{j=2}^{I-1-i} \frac{\partial \alpha_{i-1}^{i-1+j}}{\partial \hat{m}_{i-1}^{i-1+j}} \frac{\partial \hat{m}_{i-1}^{i-1+j}}{\partial (T_i^{i-1+j} - T_{i-1}^{i-1+j})}(T_{i-1+j} - T_i^{i-1+j}) \\
\varepsilon_i^{i+1} \quad &+ \sum_{j=2}^{I-i} \frac{\partial \alpha_i^{i+j}}{\partial \hat{m}_i^{i+j}} \frac{\partial \hat{m}_i^{i+j}}{\partial (T_{i+1}^{i+j} - T_i^{i+j})}(T_{i+j} - T_{i+1}^{i+j}) \\
&- \sum_{j=2}^{I+1} \frac{\partial \alpha_i^{i+j}}{\partial \hat{m}_i^{i+j}} \frac{\partial \hat{m}_i^{i+j}}{\partial (T_{i+1}^{i+j} - T_i^{i+j})}(T_{i+j} - T_i^{i+j}) \\
&= \sum_{j=1}^{I-i} \varepsilon_i^{i+j} - \sum_{j=1}^{I-i+1} \varepsilon_{i-1}^{i-1+j} = \sum_{j=1}^{I-i} \varepsilon_i^{i+j} - \sum_{j=0}^{I-i} \varepsilon_{i-1}^{i+j}
\end{aligned}$$

The first order condition now is

$$\frac{\partial \mathcal{L}}{\partial \tau_i}\frac{1}{y_i\lambda} = \hat{n}_i(1-\beta_i) + \sum_{j=1}^{I-i}\varepsilon_i^{i+j} - \sum_{j=0}^{I-i}\varepsilon_{i-1}^{i+j} + \sum_{j=1}^{I-i}\eta_i^{i+j} - \sum_{j=1}^{i-1}\eta_{i-j}^{i}.$$

Next, summing from i' to I,

$$\sum_{i=i'}^{I}\frac{\partial \mathcal{L}}{\partial \tau_i}\frac{1}{y_i\lambda} = \sum_{j=1}^{I-i}\hat{n}_{i+j}(1-\beta_{i+j}) - \sum_{j=0}^{I-i}\left[\varepsilon_{i-1}^{i+j} + \eta_{i-1}^{i+j}\right].$$

Finally, I define $\hat{N}_i = \sum_{j=1}^{I-i}\hat{n}_i$ as the share of individuals who report more than y_i in labour income. Next, dividing both sides by $\hat{N}_i$, and equation (16) follows.

The necessary condition for t_i is

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial t_i}\frac{1}{rs_i\lambda} = {} & \hat{n}_i(1-\beta_i) + \sum_{j=1}^{I-i}\alpha_i^{i+j}(1-\beta_i^{i+j})j\hat{\delta}_i \\
& - \sum_{j=1}^{I-(i-1)}\frac{\partial \alpha_{i-1}^{i-1+j}}{\partial \hat{m}_{i-1}^{i-1+j}}\frac{\partial \hat{m}_{i-1}^{i+j}}{\partial (T_i^{i-1+j} - T_{i-1}^{i-1+j})}(T_{i-1+j} - T_{i-1}^{i-1+j})(1+(j-1)\hat{\delta}_i) \\
& + \sum_{j=1}^{I-i}\frac{\partial \alpha_{i-1}^{i+j}}{\partial \hat{m}_{i-1}^{i+j}}\frac{\partial \hat{m}_{i-1}^{i+j}}{\partial (T_i^{i+j} - T_{i-1}^{i+j})}(T_{i+j} - T_i^{i+j})(1+j\hat{\delta}_i) \\
& + \sum_{j=1}^{I-i}\frac{\partial \alpha_{i}^{i+j}}{\partial \hat{m}_{i}^{i+j}}\frac{\partial \hat{m}_{i}^{i+j}}{\partial (T_{i+1}^{i+j} - T_{i}^{i+j})}(T_{i+j} - T_i^{i+j})(1+j\hat{\delta}_i) \\
& + \sum_{j=2}^{I-(i+1)}\frac{\partial \alpha_{i}^{i+1+j}}{\partial \hat{m}_{i}^{i+1+j}}\frac{\partial \hat{m}_{i}^{i+1+j}}{\partial (T_{i+1}^{i+1+j} - T_{i}^{i+1+j})}(T_{i+1+j} - T_{i+1}^{i+1+j})(1+(j+1)\hat{\delta}_i) \\
& + \sum_{j=1}^{I-i}\frac{\partial \alpha_{i}^{i+j}}{\partial \hat{k}_{i}^{i+j}}\frac{\partial \hat{k}_{i}^{i+j}}{\partial (T_{i+j} - T_{i}^{i+j})}(T_{i+j} - T_i^{i+j})(1+j\hat{\delta}_i) \\
& - \sum_{j=1}^{i-1}\frac{\partial \alpha_{i-j}^{i}}{\partial \hat{k}_{i-j}^{i}}\frac{\partial \hat{k}_{i-j}^{i}}{\partial (T_{i} - T_{i-j}^{i})}(T_{i} - T_i^{i-j}) + \hat{n}_i\frac{t_i}{s_i}\frac{\partial s_i}{\partial t_i} \\
= {} & \underbrace{\hat{n}_i(1-\beta_i) + \sum_{j=1}^{I-i}\varepsilon_i^{i+j} - \sum_{j=0}^{I-i}\varepsilon_{i-1}^{i+j} + \sum_{j=1}^{I-i}\eta_i^{i+j} - \sum_{j=0}^{i-1}\eta_{i-j}^{i}}_{=0} \\
& + \sum_{j=1}^{I-i}j\hat{\delta}_i\left[\alpha_i^{i+j}(1-\beta_i^{i+j}) + (\varepsilon_i^{i+j} - \varepsilon_{i-1}^{i+j}) + \eta_i^{i+j}\right] - \hat{n}_i\zeta_i\frac{t_i}{1-t_i} = 0.
\end{aligned}$$

In the final step, I used similar steps as in deriving the necessary condition for τ_i.

A.2 Intensive and extensive elasticities

Here, I analyse whether $\varepsilon_i^{i+j} \lesseqgtr \varepsilon_{i-1}^{i+j}$, from (11b),

$$\varepsilon_i^{i+j} = \frac{\int_{w_{i+j}}^{w_{i+j+1}} \int_{\underline{k}}^{\hat{k}_i^{i+j}} f(w,k,\hat{m}_i^{i+j}) dwdk}{d_2\left(j\delta, \hat{m}_i^{i+j}\right) - d_2\left((j-1)\delta, \hat{m}_i^{i+j}\right)} \left[T_{i+1}^{i+j} - T_i^{i+j}\right] > 0, \tag{21a}$$

$$\varepsilon_{i-1}^{i+j} = \frac{\int_{w_{i+j}}^{w_{i+j+1}} \int_{\underline{k}}^{\hat{k}_{i-1}^{i+j}} f(w,k,\hat{m}_{i-1}^{i+j}) dwdk}{d_2\left(j\delta, \hat{m}_{i-1}^{i+j}\right) - d_2\left((j-1)\delta, \hat{m}_{i-1}^{i+j}\right)} \left[T_i^{i+j} - T_{i-1}^{i+j}\right] > 0. \tag{21b}$$

From these conditions, one cannot claim which is larger as it depends on these factors:

- If the tax system is progressive, i.e. if $T_{i+1}^{i+j} - T_i^{i+j} > T_i^{i+j} - T_{i-1}^{i+j}$, then this implies $\varepsilon_i^{i+j} > \varepsilon_{i-1}^{i+j}$ due to the higher bracket on the RHS in (21a).
- If $d(j\delta, m) = (j\delta)^2 m^a$, with $a > 0$, satisfying the conditions stated in Section 3, then the intensive shifting cost structure implies $\varepsilon_i^{i+j} > \varepsilon_{i-1}^{i+j}$, as the denominator in (21a) is larger than in (21b).
- If the tax threshold for extensive shifting costs increases (decreases) with the amount shifted, i.e. if $\hat{k}_{i-1}^{i+j} > \hat{k}_i^{i+j}$ ($\hat{k}_{i-1}^{i+j} < \hat{k}_i^{i+j}$), then this implies $\varepsilon_i^{i+j} < \varepsilon_{i-1}^{i+j}$ ($\varepsilon_i^{i+j} > \varepsilon_{i-1}^{i+j}$). From condition (8), it is not clear whether $\hat{k}_{i-1}^{i+j}$ or $\hat{k}_i^{i+j}$ should be higher.
- If $\frac{\partial f(w,k,m)}{\partial m} > 0$, then this implies $\varepsilon_i^{i+j} > \varepsilon_{i-1}^{i+j}$, as $\hat{m}_i^{i+j} > \hat{m}_{i-1}^{i+j}$.

The above implies that $\varepsilon_i^{i+j} > \varepsilon_{i-1}^{i+j}$, unless $\hat{k}_{i-1}^{i+j} - \hat{k}_i^{i+j}$ is sufficiently large and/or $\frac{\partial f(w,k,m)}{\partial m}$ sufficiently negative.

A.3 Average capital income tax rate

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial t}\frac{1}{rs\lambda} = & \underbrace{\sum_i \hat{n}_i(1-\beta_i)}_{=0} + \sum_{i,j} \frac{j\delta}{rs}\alpha_i^{i+j}(1-\beta_i^{i+j}) - \sum_i n_i \zeta \frac{t}{1-t} \\
& + \sum_{i,j} \frac{j\delta}{rs} \frac{\partial \alpha_i^{i+j}}{\partial \hat{k}_i^{i+j}} \frac{\partial \hat{k}_i^{i+j}}{\partial (T_{i+j} - T_i^{i+j})} (T_{i+j} - T_i^{i+j}) \\
& + \sum_{i,j} \frac{\delta}{rs} \frac{\partial \alpha_i^{i+j}}{\partial \hat{m}_i^{i+j}} \frac{\partial \hat{m}_i^{i+j}}{\partial (T_{i+1}^{i+j} - T_{i+j}^{i})} (T_{i+j} - T_{i+j}^{i}) \\
& - \sum_{i,j} \frac{\delta}{rs} \frac{\partial \alpha_{i-1}^{i+j}}{\partial \hat{m}_{i-1}^{i+j}} \frac{\partial \hat{m}_{i-1}^{i+j}}{\partial (T_i^{i+j} - T_{i+j}^{i-1})} (T_{i+j} - T_{i+j}^{i}) \\
= & \sum_{i,j} j\hat{\delta}\alpha_i^{i+j}(1-\beta_i^{i+j}) - \zeta\frac{t}{1-t} + \sum_{i,j} j\hat{\delta}\eta_i^{i+j} + \sum_{i,j} \hat{\delta}\varepsilon_i^{i+j} = 0,
\end{aligned}$$

where $\hat{\delta} = \delta/(rs)$. The optimal capital income tax rates are characterised by the following condition,

$$\frac{t}{1-t} = \sum_{i,j} \frac{\hat{\delta}}{\zeta} \left[j\hat{\delta}\alpha_i^{i+j}(1-\beta_i^{i+j}) + j\eta_i^{i+j} + \varepsilon_i^{i+j} \right], \tag{23}$$

Assuming that $\frac{\hat{\delta}_i}{\zeta_i} = \frac{\hat{\delta}_{i+1}}{\zeta_{i+1}}$, then the average capital income tax rate is:[9]

$$\frac{\bar{t}}{1-\bar{t}} = \sum_i \hat{n}_i \frac{t_i}{1-t_i} = \sum_{i,j} \frac{\hat{\delta}_i}{\zeta_i} \left[j\alpha_i^{i+j}(1-\beta_i^{i+j}) + j\eta_i^{i+j} \right] + \sum_{j=1}^{I-1} \varepsilon_j^{1+j}. \tag{24}$$

The optimal top capital income tax rates are characterised by the following condition,

$$\frac{t_n}{1-t_n} = \sum_{\substack{i=n,\\ j=1}} \frac{j\hat{\delta}_n}{\hat{\zeta}_n} \left[\alpha_i^{i+j}\left(1-\beta_i^{i+j}\right) + \left(\varepsilon_i^{i+j} - \varepsilon_{i-1}^{i+j}\right) + \eta_i^{i+j} \right], \tag{25}$$

for $i \geq n$.

[9] This holds when $t_i = t_{i+1}$. This will only hold under special conditions. However, when the tax rate only slightly varies with income, then (24) might be a good approximation.