Income Shifting along the Intensive vs. the Extensive Margin: Implications for Optimal Tax Design*

Håkan Selin† and Laurent Simula ‡

April 24, 2016

Abstract

The optimal tax literature has modelled income shifting as a decision along the intensive margin. However, empirical evidence suggests that income shifting involves significant fixed costs, which give rise to an important extensive margin. In this article, we show that the distinction between the intensive and extensive margins has crucial optimal policy implications as far as income shifting is concerned. We consider a population of agents differing both in terms of productivity and ability to shift income. We first investigate a model in which income shifting may only occur along an intensive margin. We demonstrate that the social planner should stop shifting when there is a negative dependency between skills and shifting costs, or if the two distributions are independent. In the extensive margin model the social planner should not in general combat shifting. In particular, numerical simulations suggest that the social planner should allow for income shifting if elasticities are heterogeneous in the population.

Keywords: Income Shifting, Optimal Taxation, Labor Income Tax, Capital Income Tax, Dual Taxation

JEL Classification: H21; H24

Preliminary and incomplete - please do not cite!

*We are grateful to Roger Gordon, Wojciech Kopczuk, Claus Thustrup Kreiner, Joel Slemrod and participants at the workshop on Individual and Small Business Taxation in Warsaw for comments and suggestions.

†Institute for Evaluation of Labour Market and Education Policy (IFAU), Uppsala, Sweden, and Uppsala Center for Fiscal Studies. Email: hakan.selin@ifau.uu.se;

‡University Grenoble Alpes, UMR GAEL, CS 40700, 38058 Grenoble Cedex 09, France. Email: Laurent.Simula@univ-grenoble-alpes.fr

1 Introduction

The possibility to shift income between different tax bases is often regarded as socially undesirable. To illustrate why, we may consider a representative-agent economy, in which the social planner aims at minimizing the deadweight loss. In addition, suppose that labor and capital incomes are taxed linearly at the rates τ_P and τ_C respectively, with $\tau_P \geq \tau_C$. For simplicity, there is no capital income other than the shifted income A, while labor earnings can be shifted at a convex cost $c(A)$. To maximize her utility, it is optimal for the agent to shift labor earnings until the marginal benefit of doing so, $\tau_P - \tau_C$, equates the marginal cost of shifting, $c'(A)$. Let us now assume that $\tau_P > \tau_C$. Increasing τ_C does not only raise tax revenues; it also reduces the agent's investment in socially wasteful tax planning activities. Therefore, in the social optimum, all possibilities of income shifting should be prevented, by setting τ_P equal to τ_C.

This *intensive margin* logic underlies the recent influential writings by Fuest and Huber (2001),Christiansen and Tuomala (2008), Piketty et al. (2013), Piketty and Saez (2013, section 3.3) and Hermle and Peichl (2015) Fuest and Huber (2001),Christiansen and Tuomala (2008) and Hermle and Peichl (2015). In the latter articles, individuals face a shifting cost which is smooth and convex. Given this specification, they both decide how much labor to supply and how much labor income to shift. More generally, convex cost functions are widely used to analyze the normative implications of tax avoidance (see e.g. Slemrod and Kopczuk (2002), Kopczuk (2001) and Chetty (2009)).

There is however practical evidence of a variety of fixed costs associated with income shifting, such as the cost of setting up a closely held corporation. These fixed costs give rise to an *extensive margin*, in addition to the previously emphasized intensive margin. The novelty of this article is to consider income shifting along *both* the intensive and extensive margins. For expositional purposes, we on the one hand re-examine the implications of a pure intensive margin logic and on the other hand contrast the former with results obtained when the emphasis is placed on a pure extensive margin. As a matter of fact, we show that the introduction of an extensive margin of income shifting radically modifies implications for policy making.

Income shifting for tax purposes may appear in many forms. In comprehensive tax systems, like in the United States, people may shift income between the personal and the corporate income tax bases, as emphasized in Gordon and Slemrod (1998). In dual income tax systems, with separate taxation of labor and capital incomes, a fraction of taxpayers shift income between the labor and the capital income tax bases, as recently shown by Pirttilä and Selin (2011) for Finland and Alstadsæter and Jacob

(2014) for Sweden. In these countries particularly owners of closely held corporations face opportunities to shift income because the government cannot distinguish the capital and labor income components of their business income. Recently, countries with dual income taxation have addressed the income shifting issue in very different ways. In Norway, dividends from closely held corporations in excess of a normal after-tax return to savings is taxed at a rate close to the top marginal tax rate on labor income. In Sweden, on the other hand, there is presently a large gap between the top marginal tax rate on labor income and the tax rate for dividends from closely held coporations.[1] Hence, Norway attempts to equalize tax rates, whereas the Swedish tax code allows taxpayers to buy themselves out from high marginal taxes on labor income. Presumably, there are pros and cons of these conflicting tax policies. A concern with tax rate equalization is that the dividend tax rate is set 'too high' and the labor income tax rate 'too low'. With tax rate differentiation, the typical concern is income shifting activities. In this article, we chose to neither model capital accumulation nor tax competition and, accordingly, abstract from standard motives in favor of taxing capital income lower than labor income. We rather focus on the income shifting mechanism *per se* and its interaction with labor supply.[2]

From a formal view point, we study an income tax problem in the spirit of Atkinson and Stiglitz (1980) and Slemrod (1994).[3] To focus on the income shifting mechanism from labor earnings into capital earnings, we let each individual's stock of capital be given exogenously and set it equal to zero. In this setting, a benevolent social planner implements a linear tax on labor incomes and a proportional tax on capital income, with a view of maximizing a weighted sum of individual utilities. Because the planner has two tax instruments at it disposal, it can always eliminate income shifting at no direct cost. Every agent has the possibility to engage in shifting activities between the two tax bases. She therefore faces a trade-off between reducing her total tax liability and paying the costs of shifting. Following Kopczuk (2001), we consider that agents differ

[1] Details here.

[2] As previous optimal tax literature on income shifting, we necessarily abstract from certain aspects that should also be accounted for when designing economic policy. In particular, we do not consider income splitting rules, which are present in the Nordic countries. The latter depend on a 'presumed rate of return', corresponding to an imputed 'normal' rate of return to business assets (see Sørensen (2005)).

[3] Our paper also relates to the literature on the taxation of entrepreneurial income, but has a different focus. Recently, Scheuer (2014) analyzes a model in which the production side is managed by entrepreneurs whilst both wages and the decision to become a worker or an entrepreneur are endogenous. Because Scheuer (2014) derives formulas for both the optimal marginal profit and income tax rates, it is conceptually related to our paper. In Scheuer's article, agents differ in two dimensions: with respect to their skill level and to their fixed cost of becoming an entrepreneur. Our paper has a different focus. We examine the role of the cost structure of shifting income into capital income, and we abstract from many of the issues that Scheuer analyzes, among which endogenous wages and non-linear tax functions.

with respect to two dimensions: the ability to transform effort into earnings ω as well as the capacity to shift income, that we captured through a cost parameter γ. The smaller the value of γ, the easier it is for an individual to shift income. This might be due to some intrinsic preferences or to access to specific shifting technologies. We allow for a joint distribution of the ω and γ, without making any restriction on their potential correlation. Individual utilities depend on both parameters, while social weights only depend on the skill level ω.

Considering a population differing with respect to two dimensions is both empirically relevant and conceptually important. First, there is a large heterogeneity in income shifting opportunities across individuals. Employees, on the one hand, are subject to third-party reporting of wage income and, hence, typically cannot convert labor income into capital income. Business owners, on the other hand, have larger opportunities to shift. When the gap between the two tax rates increases, one would expect some people to transit from employment to self-employment for tax purposes.[4] Second, from a conceptual viewpoint, there would be a concern that our results are contingent on the restriction to linear tax instruments if we only allowed for heterogeneity along the skill dimension.

Our key contribution in light of the earlier literature is to pay particular attention to the cost structure of income shifting. We first re-investigate the case in which income shifting operates along a pure intensive margin; specifically, we consider that for a given γ the cost of shifting is increasing and convex in the amount of shifted income. This is a generalization of the assumption considered in most of the previous literature. When ω and γ are either independent or negative quadrant dependent, we find that the government should equalize the marginal tax rates on labor income and capital income, therefore eliminating income shifting. Negative quadrant dependence, though not in the basic economic toolkit, is a widely used measure of dependence. It means that the joint probability that both ω and γ are larger than a given threshold is always smaller as compared to the situation when these variables are independent. It in particular implies negative correlation between ω and γ. Under positive quadrant dependence instead, the government should set the capital income tax rate at a lower value than the labor income tax rate. In this case, income shifting may be regarded as socially optimal.[5] We therefore see that the result of tax rate equalization, as formulated by Piketty and Saez (2013) in a model with a single dimension of heterogeneity (skills), does only generalize

[4] The tax structure is also an important determinant of the choice of organizational form within the population of self-employed, as shown by Edmark and Gordon (2013).

[5] This result is related to the point made by Kopczuk (2001) in the context of tax avoidance.

to a more realistic two-dimensional setting if the joint density exhibits certain properties.

We then investigate the case in which income shifting operates along a pure extensive margin, and thus involves a fixed cost. We show it is usually not socially optimal to equalize the marginal tax rates for labor and capital earnings. In the social optimum, the population is therefore typically partionned into "shifters" and "non-shifters". In the shifting sub-population, the marginal incentives to supply labor is determined by the capital income tax rate whilst, in the non-shifting group, by the labor income tax rate. Our optimal income tax formulae do not however inform us about the magnitude of the gap between the two marginal tax rates. We therefore provide numerical simulations to investigate this issue. Regarding preferences, we focus on the case in which the social planner aims at maximizing the well-being of the worst-off individuals (maximin) and abstract from income effects on taxable income. Using Swedish data for skill levels, we calibrate the joint distribution of skills and shifting costs so that, for the actual average values of the labour and capital tax rates, the amount of individuals deciding to shift income corresponds to the empirical estimate. In our benchmark scenario, we find that it is socially optimal to set the labor income tax rate 9 to 15 percentage points higher than the capital income tax, depending on the correlation structure of ω and γ. Our main conclusions extend to a mixed model, in which both intensive and extensive margins are simultaneously accounted for.

The article is organized as follows. Section 2 sets up a general model of labor supply and income shifting. Section 3 examines the optimal tax structure when income shifting operates along the intensive margin. Section 4 focuses on the extensive margin. Section 5 provides concluding comments.

2 The Model

We investigate the situation in which a benevolent policy-maker would like to redistribute income within its population. On the one hand, labor income is taxed in a linear manner, with lump-sum income G and marginal rate τ_P. On the other hand, capital income is taxed in a proportional manner, at the rate τ_C.

The population consists of individuals differing with respect to two dimensions: the ability to transform effort into earnings ω and the difficulty for them to shift labor earnings into capital earnings, captured by the parameter γ. The joint distribution of ω and γ is given by the probability density function $f(\omega, \gamma)$. We assume that its support is $\mathbb{R}^+ \times \mathbb{R}^+$ and do not make any restriction on the possible correlation between ω and γ. The policy-maker knows the distribution of types within the population, but is unable to observe nor recover the type of a specific individual.

Individual Choices

To model individual choices, we use the canonical labor-leisure model, augmented with a possibility of income shifting. We denote individual consumption (or net income) by Y and labor supplied by L. An individual of skill ω supplying L units of effort receives a gross income equal to ωL but incurs a utility loss $v\left(L;\omega\right)$, with $v'_L > 0$, $v'_\omega < 0$, $v''_{LL} > 0$ and $v''_{L\omega} \leq 0$. The individual utility function is given by:

$$U(Y, L) = Y - v(L; \omega). \tag{1}$$

Given this specification, there is no income effect on supplied labor. We allow for the possibility that the disutility of labor depends on the skill level. Every individual has the possibility to reduce her income subject to the income tax, from ωL to $Z = \omega L - A$ at a cost $\Gamma(A, \gamma)$. We refer to this possibility as *income shifting*. As emphasized in the introduction, this cost may correspond to a fixed cost and/or depend on how much earnings are shifted. A general specification is $\Gamma(A, \gamma) = k(\gamma) + C(A; \gamma)$. Most of the previous literature has focused on the case where $C(A; \gamma) = C(A)$. In the following, we investigate the implications of a more general cost structure. We proceed in steps and specifically look into the cases when $\Gamma(A, \gamma) = C(A; \gamma)$ and $\Gamma(A, \gamma) = k(\gamma)$.

Shifted income A is subject to the capital income tax. Overall, an individual pays a tax liability of $\tau_p Z + G - \tau_C A$ and thus receives a net income of:

$$Y = \omega L - (\tau_P Z + G - \tau_C A) - \Gamma(A, \gamma). \tag{2}$$

Individual choices proceed from the maximization of the utility $U(Y, L)$ subject to the budget constraint (2). The indirect utility is therefore defined as:

$$V(\omega, \gamma) = \max_{L,A}\{\omega L - \tau_P(\omega L - A) - \tau_C A - \Gamma(A, \gamma) + G - v(L; \omega)\}. \tag{3}$$

Given the additive separability of this specification, the optimal value of L is independent of γ. We call $L\left(\omega\right)$ the optimal supply of effort and $A\left(\omega, \gamma\right)$ the optimal amount of shifting for an individual of type (ω, γ). For later use, we also define:

$$\begin{aligned} V^P(\omega) &= \max_L\{(1 - \tau_P)\,\omega L + G - v(L; \omega)\}, && (4) \\ V^C(\omega, \gamma) &= \max_L\{(1 - \tau_C)\,\omega L + G - \Gamma(\omega L, \gamma) - v(L; \omega)\}. && (5) \end{aligned}$$

For an individual of type (ω, γ), the first optimization program provides the maximum utility $V^P(\omega, \gamma)$ that can be obtained in the absence of any income shifting ($A = 0$). We denote by $L^P(\omega)$ its solution in L. The second optimization program provides the maximum utility $V^C(\omega, \gamma)$ when the entire labor earnings are shifted ($A = \omega L$). We denote by $L^C(\omega, \gamma)$ its solution in L.

Policy-Maker's Choices

The policy-maker chooses the tax rates, τ_P and τ_C, and the lump-sum income G to maximize a weighted sum of individual utilities:

$$\int_\omega \int_\gamma g(\omega) V(\omega, \gamma) f(\omega, \gamma) d\gamma d\omega, \tag{6}$$

subject to:

$$G + R = \tau_P \int_\omega \int_\gamma [\omega L(\omega) - A(\omega, \gamma)] f(\omega, \gamma) d\gamma d\omega + \tau_C \int_\omega \int_\gamma A(\omega, \gamma) f(\omega, \gamma) d\gamma d\omega. \tag{7}$$

We consider that the social weights $g(\cdot)$ only depend on productivity levels. A more general specification would be to allow social weights to both depend on ω and γ. However, the reasons for which the policy-maker would like to redistribute incomes based on the second dimension of heterogeneity γ is not completely transparent to us. In addition, we would not be able to obtain clear-cut results without introducing additional, and rather *ad hoc*, sorting assumptions. R is a tax revenue requirement that does not enter the individuals' utility function. When it is set equal to zero, the tax policy is purely redistributive.

3 The Intensive Margin

In this section, we consider that the cost of income shifting is given by $\Gamma(A, \gamma) = C(A, \gamma)$ where $C > 0$, $C'_A > 0$ and $C''_{AA} > 0$. For a given γ, the cost is thus positive and convexly increasing in A. To avoid corner solutions, we further assume that an individual would face an arbitrarily large cost when shifting her entire earnings. In addition, given our interpretation of γ as an underlying difficulty to shift income (e.g., the difficulty to access shifting technologies), we consider that the marginal cost to shift income C'_A is increasing in γ, i.e., $C''_{A\gamma} > 0$.

3.1 Behavioral Responses

Differentiating (3), we obtain the following first-order conditions of the individual utility maximisation program:

$$v'(L;\omega) = \omega(1-\tau_P) \tag{8}$$

$$C'_A(A,\gamma) = \tau_P - \tau_C \tag{9}$$

The optimal effort level depends on the productivity ω and on the retention rate $1-\tau_P$. It is in particular increasing with respect to productivity ω. The optimal amount of shifting, from the individual perspective, is therefore independent of ω. It only depends on γ and on the tax differential $\tau_P - \tau_C$. By the implicit function theorem, we see that $\partial A/\partial\gamma = -C''_{A\gamma}/C''_{AA} < 0$. On this basis, it is useful to summarize behavioral responses to taxation in terms of elasticities. We call ε_Z the elasticity of Z with respect to the retention rate $1-\tau_P$, whilst keeping τ_C constant:

$$\varepsilon_Z(\omega,\gamma) = -\frac{1-\tau_P}{Z}\frac{\partial Z}{\partial \tau_P} = \frac{1-\tau_P}{Z}\left[\frac{\omega^2}{v''(L)} - \frac{1}{C''_{AA}}\right]. \tag{10}$$

As shown in the square bracket, a small increase in τ_P has two effects on the optimal individual choices. First, the usual substitution effect induces her to work less, which reduces labor earnings. Second, the incentive to shift income from the labor income tax to the capital income tax becomes larger. The part played by the second response in the total behavioral response is therefore:

$$\sigma(\omega,\gamma) = \frac{\partial A/\partial(\tau_P-\tau_C)}{\partial A/\partial(\tau_P-\tau_C) - \omega\,\partial L/\partial\tau_P} = \frac{\partial A/\partial(\tau_P-\tau_C)}{\partial Z/\partial(1-\tau_P)}. \tag{11}$$

Following Piketty et al. (2013), the shifting elasticity component corresponds to the share of the behavioral response due to income shifting. It is formally defined as:

$$\varepsilon_A(\omega,\gamma) = \sigma(\omega,\gamma)\;\varepsilon_Z(\omega,\gamma) = \frac{1-\tau_P}{Z}\frac{\partial A}{\partial(\tau_P-\tau_C)}. \tag{12}$$

3.2 Optimal Tax Policy

The social planner chooses the tax rates τ_P, τ_C and the lump-sum income G so as to maximize the welfare functional (6) subject to the tax revenue constraint (7). Denoting the shadow price of public funds by λ, the Lagrangian of the optimization problem is given by:

$$\int_\omega \int_\gamma \{g(\omega)V(\omega,\gamma) + \lambda\left[\tau_P \omega L(\omega) - (\tau_P - \tau_C)A(\omega,\gamma) - G - R\right]\} f(\omega,\gamma)d\gamma d\omega. \tag{13}$$

To simplify notations, we omit the arguments of the different functions. The first-order conditions with respect to τ_P, τ_C and G are respectively:

$$\int_\omega \int_\gamma \left(b\frac{\partial V}{\partial \tau_P} + \omega L - A + \tau_P \frac{\partial \omega L}{\partial \tau_P} - (\tau_P - \tau_C)\frac{\partial A}{\partial \tau_P} \right) f(\omega,\gamma)d\gamma d\omega = 0, \tag{14}$$

$$\int_\omega \int_\gamma \left(b\frac{\partial V}{\partial \tau_C} + A - (\tau_P - \tau_C)\frac{\partial A}{\partial \tau_C} \right) f(\omega,\gamma)d\gamma d\omega = 0, \tag{15}$$

$$\int_\omega \int_\gamma \left(b\frac{\partial V}{\partial G} - 1 \right) f(\omega,\gamma)d\gamma d\omega = 0. \tag{16}$$

The quantity $b(\omega) = g(\omega)/\lambda$ corresponds to the net social marginal valuation of income of an individual of skill ω. We call $\bar{b}$ its average value in the population. Using these weights and the behavioral elasticities ε_A and ε_Z introduced above, we can rearrange the optimality conditions in a more transparent manner:

$$\int_\omega \int_\gamma \left(-b \cdot Z + Z - \frac{\tau_P}{1-\tau_P}\varepsilon_Z Z + \frac{\tau_C}{1-\tau_P}\varepsilon_A Z \right) f(\omega,\gamma)d\gamma d\omega = 0, \tag{17}$$

$$\int_\omega \int_\gamma \left(-b \cdot A + A + \frac{\tau_P - \tau_C}{1-\tau_P}\varepsilon_A Z \right) f(\omega,\gamma)d\gamma d\omega = 0, \tag{18}$$

$$\int_\omega \int_\gamma (b-1)\, f(\omega,\gamma)d\gamma d\omega = 0 \quad i.e., \quad \bar{b} = 1. \tag{19}$$

Combining the first-order conditions with respect to τ_P and G , we obtain:

$$\int_\omega \int_\gamma Z \left(b - \bar{b} + \frac{\tau_P}{1-\tau_P}\varepsilon_Z \right) f(\omega,\gamma)d\gamma d\omega - \int_\omega \int_\gamma \frac{\tau_C}{1-\tau_P}\varepsilon_A Z f(\omega,\gamma)d\gamma d\omega = 0. \tag{20}$$

Introducing the covariance of Z and b, and rearranging, we can write:

$$\frac{\tau_P}{1-\tau_P} = -\frac{cov(Z,b)}{\int_\omega \int_\gamma \varepsilon_Z Z f(\omega,\gamma)d\gamma d\omega} + \frac{\tau_C}{1-\tau_P}\frac{\int_\omega \int_\gamma \sigma \varepsilon_Z Z f(\omega,\gamma)d\gamma d\omega}{\int_\omega \int_\gamma \varepsilon_Z Z f(\omega,\gamma)d\gamma d\omega}. \tag{21}$$

We see that the marginal tax rate on labor earnings depends on two terms. The first one looks the same as in the usual optimal linear income tax, with Z replacing ωL. It clearly reflects the trade-off between equity, on the numerator, and efficiency, on the denominator. The second term is new. It involves a weighted average of the amount of income shifted by the individuals, the weights being given by the elasticity ε_A. Rearranging the first-order condition with respect to τ_C, we obtain:

$$\frac{\tau_P - \tau_C}{1 - \tau_P} = \frac{cov\,(A, b)}{\int_\omega \int_\gamma \varepsilon_A Z f(\omega, \gamma) d\gamma d\omega}. \tag{22}$$

By definition, the denominator of the right-hand side is positive. This implies that τ_P is strictly larger than τ_C if and only if $cov(A, b) > 0$. We know that the function A is increasing in γ and independent of ω. Conversely, the social weights b are decreasing in ω and independent of γ. Therefore,

$$cov(A, b) = \int_\omega \int_\gamma [F(\omega, \gamma) - f(\omega)g(\gamma)] db(\omega) dA(\gamma) \tag{23}$$

where $F(\omega, \gamma)$ is the cumulative density function of the joint distribution of (ω, γ) whilst f and g are the conditional probability density functions. Given that $db(\omega) < 0$ and $dA(\gamma) < 0$, a sufficient condition for $cov(A, b)$ to be positive is that the square bracket inside the double integral be positive (see Cuadras (2002), Theorem 1). Conversely, a sufficient condition for $cov(A, b)$ to be negative is that the square bracket inside the double integral be negative. We therefore obtain the following proposition.

Proposition 1 *When $\Gamma(A, \gamma) = C(A, \gamma)$, it is socially optimal to set:*

1. *$\tau_P > \tau_C$ when ω and γ are positive quadrant dependent;*

2. *$\tau_P = \tau_C$ when ω and γ are independent or negative quadrant dependent.*

In words, positive quadrant dependent means that the joint probability that both ω and γ are larger than a pair ($\hat{\omega}$,$\hat{\gamma}$) is larger than the product of the two independent probabilities for all possible pairs ($\hat{\omega}$,$\hat{\gamma}$). Negative quadrant dependence means that this joint probability is smaller. Loosely speaking, quadrant dependence is a stronger version of the more often used concept of correlation. Positive quadrant dependence implies positive correlation, whilst negative quadrant dependence implies negative correlation. The converse statements are not necessarily true. For the bivariate normal distribution however, the sign of correlation coefficient and the sign of the quadratic dependence are always the same.

4 The Extensive Margin

We now investigate the case in which income shifting operates along the extensive margin. We therefore interpret the parameter of heterogeneity γ as a fixed cost.

4.1 Behavioral Responses

When income shifting is done against a fixed cost, a rational individual either shifts nothing ($A = 0$) or her entire labor earnings ($A = \omega L$). In the first case, her utility amounts to $V^P(\omega)$ and in the latter to $V^C(\omega, \gamma)$. Consequently, she chooses $A = 0$ when $V^P(\omega) \geq V^C(\omega, \gamma)$ and $A = \omega L$ when $V^C(\omega, \gamma) > V^P(\omega)$. Using (4) and (5), we see that $V^P(\omega) < V^C(\omega, \gamma)$ if and only if:

$$(1 - \tau_P)\omega L^P + G - v(L^P; \omega) < (1 - \tau_C)\omega L^C + G - \gamma - v(L^C; \omega), \tag{24}$$

which is equivalent to:

$$\gamma < [(1 - \tau_C)\omega L^C - (1 - \tau_P)\omega L^P] + [v(L^P; \omega) - v(L^C; \omega)]. \tag{25}$$

Given that the fixed costs enters the utility in an additively separable way, $L^C(\omega, \gamma)$ is equal to $L^C(\omega)$. Because $\tau_P \geq \tau_C$, the first square bracket on the right-hand side of (25) is increasing in ω. Moreover, $v''_{L\omega} \leq 0$ implies that the second square bracket is non-decreasing. Consequently, the right-hand side of (25) is monotonically increasing in ω. This implies the following:

Lemma 2 *There is a cut-off level $\hat{\gamma}(\omega)$, non-decreasing in ω, such that:*

1. *for $\gamma < \hat{\gamma}(\omega)$, $A(\omega, \gamma) = \omega L^C(\omega)$ and $L(\omega, \gamma) = L^C(\omega)$;*

2. *for $\gamma \geq \hat{\gamma}(\omega)$, $A(\omega, \gamma) = 0$ and $L(\omega, \gamma) = L^P(\omega)$.*

We can alternatively invert the $\hat{\gamma}$ function and formulate the above lemma in terms of ω. This will prove particularly useful when formulating the social planner's optimization problem.

Corollary 3 *There is a cut-off level $\hat{\omega}(\gamma)$, non-decreasing in γ such that:*

1. *for $\omega < \hat{\omega}(\gamma)$, $A(\omega, \gamma) = 0$ and $L(\omega, \gamma) = L^P(\omega)$;*

2. *for $\omega \geq \hat{\omega}(\gamma)$, $A(\omega, \gamma) = \omega L^C(\omega)$ and $L(\omega, \gamma) = L^C(\omega)$.*

4.2 Optimal Tax Policy

Given Corollary 3, we can separate the population in two groups, those facing the labor income tax and those facing the capital income tax. The social planner's problem can therefore be formulated as:

$$\max_{\tau_P,\tau_C,G}\int_0^\infty\int_0^{\hat{\omega}(\gamma)} g(\omega)V^P(\omega)\,f(\omega,\gamma)\,d\omega d\gamma+\int_0^\infty\int_{\hat{\omega}(\gamma)}^\infty g(\omega)V^C(\omega,\gamma)\,f(\omega,\gamma)\,d\omega d\gamma, \tag{26}$$

subject to the tax revenue constraint:

$$\tau_P\int_0^\infty\int_0^{\hat{\omega}(\gamma)}\omega L^P(\omega)f(\omega,\gamma)d\omega d\gamma+\tau_C\int_0^\infty \omega L^C(\omega,\gamma)f(\omega,\gamma)d\omega d\gamma-R-G=0. \tag{27}$$

Solving for the optimal tax rates, we obtain the following Proposition, the derivation of which is explained in the Appendix.

Proposition 4 *The optimal marginal tax rates τ_P and τ_C are given by:*

$$\frac{\tau_P}{1-\tau_P}=\frac{\int_0^\infty\int_0^{\hat{\omega}(\gamma)}\omega L^P(\omega)(1-b(\omega))f(\omega,\gamma)d\omega d\gamma+\int_0^\infty(\tau_P\hat{\omega}L^P(\hat{\omega})-\tau_C\hat{\omega}L^C(\hat{\omega}))\frac{\partial\hat{\omega}}{\partial\tau_P}f(\hat{\omega},\gamma)d\gamma}{\int_0^\infty\int_0^{\hat{\omega}(\gamma)}\omega L^P(\omega)\varepsilon_{\omega L}f(\omega,\gamma)d\omega d\gamma} \tag{28}$$

and

$$\frac{\tau_C}{1-\tau_C}=\frac{\int_0^\infty\int_{\hat{\omega}(\gamma)}^\infty\omega L^C(\omega)(1-b(\omega))f(\omega,\gamma)d\omega d\gamma+\int_0^\infty(\tau_P\hat{\omega}L^P(\hat{\omega})-\tau_C\hat{\omega}L^C(\hat{\omega}))\frac{\partial\hat{\omega}}{\partial\tau_C}f(\hat{\omega},\gamma)d\gamma}{\int_0^\infty\int_{\hat{\omega}(\gamma)}^\infty\omega L^P(\omega)\varepsilon_{\omega L}f(\omega,\gamma)d\omega d\gamma} \tag{29}$$

As explained above, the population is usually divided into two fractions in the social optimum. In the sub-population of individuals shifting income, the marginal incentive to supply labor is determined by the capital income tax rate τ_C. In the sub-population of agents who do not shift income at all, this marginal incentive depends on the labor income tax rate τ_P. However, we cannot rule out situations in which there would be no shifting in the optimum. In that case, the cut-off level $\hat{\omega}(\gamma)$ tends to $+\infty$ and the formulae of Proposition 4 collapse into the "usual" optimal income tax rules, with $\tau_P=\tau_C$.

4.3 Numerical Simulations

Numerical simulations are useful to gain further insights into Proposition 4. In particular, we are interested in knowing whether it is socially optimal to allow for income shifting for plausible calibrations and, if so, how large is the difference between the optimal marginal tax rates τ_P and τ_C.

We consider that the social objective is the maximin. In this case, the social planner chooses τ_P and τ_C such that tax revenues are maximized. It follows that the social planner would set τ_C lower than τ_P only if this results in larger collected taxes.

It is empirically well-known that the distribution of hourly wage rates can be well approximated by a log-normal distribution, if one abstracts from the top of the distribution. We have considerable less guidance regarding how to calibrate the distribution of shifting costs. Because we want to perform sensitivity analysis regarding the correlation of ω and γ, it is convenient for us to consider that these two parameters are described by a bivariate log-normal distribution. We use Swedish data to callibrate the mean and variance of the wage distribution. Regarding the shifting costs, we parametrize them so that the proportion of people deciding to shift incomes reproduces the actual figure for Sweden. We provide a more detailed discussion in the Appendix.

We consider that the utility function $U(Y, L)$ is given by:

$$U(Y, L) = Y - \alpha \frac{L^{1+1/g(\epsilon,\omega)}}{1 + 1/g(\epsilon, \omega)}, \tag{30}$$

where α and ϵ are taste parameters, constant across the population. Importantly, we allow for the possibility that the labor supply elasticity depends positively on the skill level.[6] Accordingly, the elasticity of earned income with respect to the net-of-tax rate $1 - \tau$ is given by $\varepsilon_{\omega L} = g(\epsilon, \omega)$. We let the elasticity be a linear function of the skill parameter, $g(\epsilon, \omega) = \epsilon + q\omega$ (where q indicates at which pace the elasticity varies with productivity).

A key issue in empirical public finance is how the earnings elasticity depends on the skill level. It should be emphasized that $\varepsilon_{\omega L}$ reflects the *real* labor supply response, and therefore does not include tax avoidance responses, as opposed to the commonly estimated 'taxable income elasticity'. As emphasized by Saez et al. (2012, p. 35), "there is compelling evidence of substantial responses of upper income taxpayers to changes in tax rates, at least in the short run. However, in all cases, the response is either due to short-term retiming or income shifting. There is no compelling evidence to date of real responses of upper income taxpayers to changes in tax rates." While this discussion supports the view that income shifting technologies are concentrated to high-skilled people, it also suggests that we actually know quite little on real labor supply behavior of highly skilled people. If we interpret the real labor supply elasticity more broadly to also accommodate for the migration margin, there is some recent evidence that migration decisions of top income earners may be very sensitive to taxes, as suggested by Kleven

[6] Alternatively, we could introduce a third dimension of heterogeneity. We have chosen a deterministic relationship between elasticities and skills to ease exposition.

et al. (2013) and Kleven et al. (2014). In the baseline simulations, we therefore assume an increasing elasticity, from 0.1 at the bottom of the income distribution to 0.5 for upper tail.

In Figure 1, we show the gap between τ_P and τ_C for 21 different values the correlation coefficient for $\log(\omega)$ and $\log(\gamma)$, denoted ρ. Additionally, this Figure reports the percentage of the population that chooses to pay the fixed cost and, thereby, shifts their entire labor income into the capital income tax base. The socially optimal allocation has the following features. First, there is a always a gap between τ_P and τ_P, which ranges from about 9,5 to 14,6 percentage points. The tax difference is actually increasing in the correlation coefficient ρ. Second, the share of shifters is declining in the correlation coefficient, from about 6% to 1%. These two opposite trends make sense. A negative correlation actually implies that highly skilled individuals (with large elasticities) face low shifting costs. Therefore, in the social optimum, number of shifters will be larger if there is a strong negative correlation.

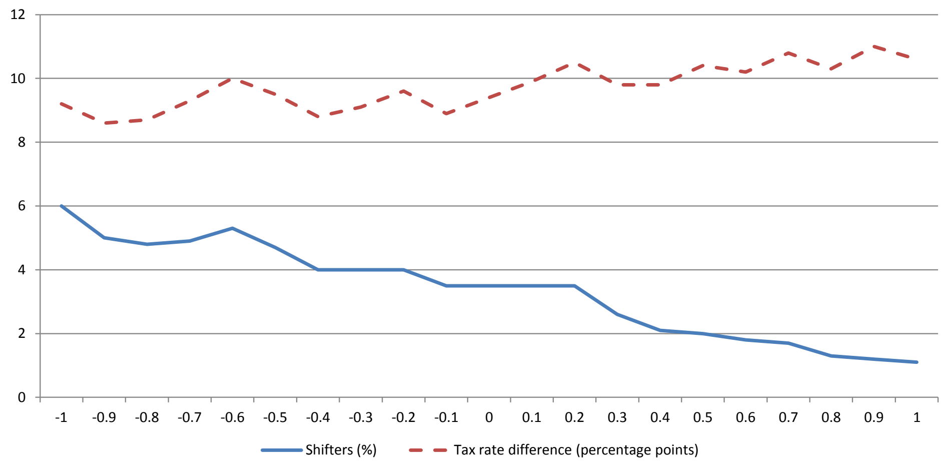

Figure 1: Features of the Optimal Allocation (Benchmark Case)

We now investigate to which extent our results are sensitive to the elasticity range. For three different values of the correlation coefficient $(-1, 0, 1)$ we examine four different elasticity ranges while keeping the average elasticity in the population constant (and equal to 0.23). The results are reported in Table 1. It appears that the variance of the elasticity is crucial for optimal tax policy.

First, when the elasticity is constant in the population, the social planner must set $\tau_P = \tau_C$. Let us assume that there elasticities are homegeneous and that there are two subpopulations in the social optimum, one reporting labor earnings and one capital

Table 1: Simulation results

Min elasticity	Max elasticity	ρ	τ_P^*	τ_C^*	$\tau_P^* - \tau_C^*$	Shifters %
0	0.725	-1	0.79	0.65	0.14	16.3
0	0.725	0	0.78	0.65	0.14	12.7
0	0.725	1	0.77	0.64	0.13	8.2
0.1	0.5	-1	0.80	0.71	0.09	6
0.1	0.5	0	0.79	0.70	0.09	3.5
0.1	0.5	1	0.79	0.68	0.11	1.1
0.15	0.4	-1	0.80	0.72	0.08	2.4
0.15	0.4	0	0.80	0.71	0.09	0.6
0.15	0.4	1	0.80	0.68	0.11	0.2
0.23	0.23	-1	0.81	0.81	0.00	0
0.23	0.23	0	0.81	0.81	0.00	0
0.23	0.23	1	0.81	0.81	0.00	0

earnings. Given the quasilinearity of individual preferences, the top of the Laffer curve would be obtained for the same marginal tax rate in the two subpopulations. Because the social objective we consider is the maximin, this implies that tax rates should not be differentiated. This is however a very particular case, which is usually looked at because it eases the calibration. Second, when the lowest ability individual exhibits an elasticity of 0 and the highest ability individual has an elasticity of 0.725, elasticities are more dispersed than in our baseline scenario. In this case, the fraction of shifters and the gap in marginal tax rates are much larger.

5 Concluding Discussion

This study highlights the distinction between the intensive and extensive margins of income shifting – a distinction neglected in the previous optimal tax literature which focused on the intensive margin. We introduce two dimensions of heterogeneity to capture the fact that income shifting possibilities may differ at a given skill level. In doing so, we concentrate on the income shifting and labor supply choices and we abstract from issues related to capital accumulation. Hence, we abstract from standard arguments in favor of setting the capital income tax rate lower than the labor income tax rate.

We first re-examine income shifting along the intensive margin, which mostly was examined in models with heterogeneity in skills only. At a given skill level, the cost of shifting is increasing and convex in the amount of shifted income. We find that the government should equalize the marginal tax rates on labor income and capital income when the joint distribution of the skills and shifting costs is either independent or exhibits negative quadrant dependence. In that case, possibilities of income shifting are eliminated. However, when there is a positive dependency between skills and shifting costs, with low skill individuals having lower shifting costs on average, the social planner can increase social welfare by allowing for income shifting, a point related to Kopczuk (2001).

We then investigate the extensive margin, i.e. the case when shifting involves a fixed cost but a zero marginal cost. In the social optimum, the population is divided into two fractions; "shifters" and "non-shifters". In the sub-population of "shifters", the marginal incentives to supply labor is determined by the capital income tax rate; and in the sub-population of "non-shifters" by the labor income tax rate. We derive inverse elasticity rules for each subpopulation which are valid for a wide class of social welfare objectives. We provide numerical simulations for a particular welfare objective, namely the maximin objective. The simulations clarify under which conditions the capital income tax rate should be set lower than the labor income tax rate. It turns out that the degree of heterogeneity in elasticites is crucial; if elasticities differ a lot in the population there is definitely a case for tax rate differentiation.

To our knowledge, no one has earlier made the point that income shifting done at a fixed cost potentially can increase social welfare. This does not mean, however, that related points have not been made in the past. Alesina and Weil (1994) show that Pareto improvements can be achieved by offering individuals a menu of linear tax schedules, where individuals of different skill levels self-select into different tax schemes.[7] While our "extensive margin model" also builds on the idea that individuals self-select into different tax schedules it nevertheless differs from the literature on "tax buyouts" in two important respects. First, in our extensive margin model the cost of choosing a lower tax rate is a pure waste from the society's point of view (it does not appear in the government's revenues). Second, we allow for heterogeneity in two dimensions.

As already emphasized, our model framework is very stylized, and the model can be developed further in several directions. One simplification is that the shifting costs are exogenous; both from the government's and the individuals' perspectives. In reality, e.g. the costs of starting up a closely held corporation, can be affected by the government.

[7] This idea of "tax buyouts" has recently been carried over to a dynamic overlapping generations economy by Del Negro et al. (2010). In an calibration exercise for the U.S. economy, they find that the introduction of the buyout benefits a significant fraction of the population.

This could be incorporated in the analysis.

Finally, we do not believe that the fact that we model two *linear* tax schedules is very important for the qualitative conclusions we make. With one skill dimension only the social planner would always be able to increase social welfare by introducing a non-linear tax on labor income. But with two dimensions of heterogeneity, however, two individuals with the same earnings in the absence of income shifting possibilities, may supply different amounts of earnings in the presence of income shifting if they face different fixed costs. Hence, introducing a non-linear labor income tax would not alter the basic mechanism at play. Note also that with some modifications, our model can be extended to a model for the 'top marginal tax rates', or the revenue maximizing applying to the highest-skilled individuals.

References

A. Alesina and Ph. Weil. Menus of linear income tax schedules. Technical Report WP No. 3968, NBER, 1994.

Annette Alstadsæter and Martin Jacob. Who Participates in Tax Avoidance? CESifo Working Paper Series 4219, CESifo Group Munich, 2013. URL `https://ideas.repec.org/p/ces/ceswps/`$_4219.html$.

Annette Alstadsæter and Martin Jacob. Dividend taxes and income shifting. arqus Discussion Papers in Quantitative Tax Research 154, arqus - Arbeitskreis Quantitative Steuerlehre, 2014. URL `http://ideas.repec.org/p/zbw/arqudp/154.html`.

A. B. Atkinson and J. E. Stiglitz. *Lectures on Public Economics*. McGraw-Hill Inc, US, 1980.

Raj Chetty. Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance. *American Economic Journal: Economic Policy*, 1(2):31–52, August 2009. URL `http://ideas.repec.org/a/aea/aejpol/v1y2009i2p31-52.html`.

Vidar Christiansen and Matti Tuomala. On taxing capital income with income shifting. *International Tax and Public Finance*, 15:527–545, 2008. ISSN 0927-5940. URL `http://dx.doi.org/10.1007/s10797-008-9076-x`. 10.1007/s10797-008-9076-x.

C.M. Cuadras. On the covariance between functions. *Journal of Multivariate Analysis*, 81(1):19 – 27, 2002. ISSN 0047-

259X. doi: http://dx.doi.org/10.1006/jmva.2001.2000. URL http://www.sciencedirect.com/science/article/pii/S0047259X01920007.

M. Del Negro, F. Perri, and F. Schivardi. Tax buyouts. *Journal of Monetary Economics*, 57(5):576 – 595, 2010. ISSN 0304-3932. doi: 10.1016/j.jmoneco.2010.06.005. URL http://www.sciencedirect.com/science/article/pii/S030439321000067X.

Karin Edmark and Roger H. Gordon. The choice of organizational form by closely-held firms in Sweden: tax versus non-tax determinants. *Industrial and Corporate Change*, 22(1):219–243, February 2013. URL http://ideas.repec.org/a/oup/indcch/v22y2013i1p219-243.html.

Clemens Fuest and Bernd Huber. Labor and capital income taxation, fiscal competition, and the distribution of wealth. *Journal of Public Economics*, 79(1):71–91, January 2001. URL http://ideas.repec.org/a/eee/pubeco/v79y2001i1p71-91.html.

Roger H. Gordon and Joel Slemrod. Are "Real" Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases? NBER Working Papers 6576, National Bureau of Economic Research, Inc, May 1998. URL http://ideas.repec.org/p/nbr/nberwo/6576.html.

Johannes Hermle and Andreas Peichl. Jointly optimal income taxes for different types of income. Technical report, 2015.

Henrik Jacobsen Kleven, Camille Landais, and Emmanuel Saez. Taxation and International Migration of Superstars: Evidence from the European Football Market. *American Economic Review*, 103(5):1892–1924, August 2013. URL http://ideas.repec.org/a/aea/aecrev/v103y2013i5p1892-1924.html.

Henrik Jacobsen Kleven, Camille Landais, and Esben Schultz. Migration and Wage Effects of Taxing Top Earners: Evidence from the Foreigners' Tax Scheme in Denmark. *The Quarterly Journal of Economics*, 129(1):333–378, 2014. URL http://ideas.repec.org/a/oup/qjecon/v129y2014i1p333-378.html.

Wojciech Kopczuk. Redistribution when avoidance behavior is heterogeneous. *Journal of Public Economics*, 81(1):51–71, July 2001. URL http://ideas.repec.org/a/eee/pubeco/v81y2001i1p51-71.html.

Thomas Piketty and Emmanuel Saez. Optimal labor income taxation. In Martin Feldstein Alan J. Auerbach, Raj Chetty and Emmanuel Saez, editors, *Handbook of public economics, vol. 5*, volume 5 of *Handbook of Public Economics*, pages 391 – 474.

Elsevier, 2013. doi: http://dx.doi.org/10.1016/B978-0-444-53759-1.00007-8. URL http://www.sciencedirect.com/science/article/pii/B9780444537591000078.

Thomas Piketty, Emmanuel Saez, and Stefanie Stantcheva. Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy*, (17616):forthcoming, November 2013. URL http://www.nber.org/papers/w17616.

Jukka Pirttilä and Håkan Selin. Income shifting within a dual income tax system: Evidence from the finnish tax reform of 1993*. *Scandinavian Journal of Economics*, 113(1):120–144, 2011. ISSN 1467-9442. doi: 10.1111/j.1467-9442.2010.01635.x. URL http://dx.doi.org/10.1111/j.1467-9442.2010.01635.x.

Emmanuel Saez, Joel Slemrod, and Seth H. Giertz. The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1):3–50, March 2012. doi: 10.1257/jel.50.1.3. URL http://www.aeaweb.org/articles.php?doi=10.1257/jel.50.1.3.

Florian Scheuer. Entrepreneurial Taxation with Endogenous Entry. *American Economic Journal: Economic Policy*, 6(2):126–63, May 2014. URL http://ideas.repec.org/a/aea/aejpol/v6y2014i2p126-63.html.

Joel Slemrod. Fixing the leak in okun's bucket optimal tax progressivity when avoidance can be controlled. *Journal of Public Economics*, 55(1):41 – 51, 1994. ISSN 0047-2727. doi: 10.1016/0047-2727(94)90079-5. URL http://www.sciencedirect.com/science/article/pii/0047272794900795.

Joel Slemrod and Wojciech Kopczuk. The optimal elasticity of taxable income. *Journal of Public Economics*, 84(1):91 – 112, 2002. ISSN 0047-2727. doi: 10.1016/S0047-2727(01)00095-0. URL http://www.sciencedirect.com/science/article/pii/S0047272701000950.

Peter Birch Sørensen. Dual income taxation: Why and how? Technical report, CESifo Working Paper Series No. 1551, 2005.

Appendix A : Proof of Proposition 4

We call λ the Lagrange multiplier of the budget constraint (27). The derivative of (26) with respect to τ_P is:

$$\int_0^{\infty}\int_0^{\hat{\omega}(\gamma)} g(\omega)\omega L^P(\omega)f(\omega,\gamma)d\omega d\gamma. \tag{31}$$

We used the fact that $V^P(\omega) = V^C(\omega,\gamma)$ for $\omega = \omega\hat{(\gamma)}$. We now compute the derivative of the budget constraint (27) with respect to τ_P. We obtain:

$$\int_0^{\infty}\int_0^{\hat{\omega}(\gamma)} \omega L^P(\omega)f(\omega,\gamma)d\omega d\gamma + \tau_P \int_0^{\infty}\int_0^{\hat{\omega}(\gamma)} \omega \frac{\partial L^P(\omega}{\partial \tau_P} f(\omega,\gamma)d\omega d\gamma$$
$$+ \int_0^{\infty} \left(\tau_P \hat{\omega} L^P(\hat{\omega}) - \tau_C \hat{\omega} L^C(\hat{\omega})\right) \frac{\partial \hat{\omega}}{\partial \tau_C} f(\hat{\omega},\gamma)d\gamma \tag{32}$$

We know write $(31) - \lambda(32) = 0$, rearrange and use the definition of $\varepsilon_{\omega L}$ to obtain (28).

To obtain (29), we compute the derivative of the social objective with respect to τ_C. Using the indifference condition at $\hat{\omega}$, we obtain:

$$\int_0^{\infty}\int_{\hat{\omega}(\gamma)}^{\infty} g(\omega)\omega L^C(\omega)f(\omega,\gamma)d\omega d\gamma. \tag{33}$$

We now compute the derivative of the budget constraint (27) with respect to τ_C:

$$\tau_P \int_0^{\infty} \hat{\omega} L^P(\hat{\omega}) \frac{\partial \hat{\omega}}{\partial \tau_C} f(\hat{\omega},\gamma)d\omega d\gamma + \int_0^{\infty}\int_{\hat{\omega}}^{\infty} \omega L^C(\omega,\gamma)d\omega d\gamma$$
$$+\tau_C \left(-\int_0^{\infty} \hat{\omega} L^C(\hat{\omega}) \frac{\partial \hat{\omega}}{\partial \tau_C} f(\hat{\omega},\gamma)d\omega d\gamma + \int_0^{\infty} \omega \frac{\partial L^C}{\partial \tau_C} f(\omega,\gamma)d\omega d\gamma\right). \tag{34}$$

We know write $(33) - \lambda(34) = 0$, rearrange and use the definition of $\varepsilon_{\omega L}$ to obtain (29).

A Appendix B: Calibration of the fixed cost model

As stated in Section X, we assume that skills, ω, and shifting costs, γ, follow a bivariate log normal distribution, i.e. $(\omega,\gamma) \sim \ln\mathcal{N}(\mu_\omega, \mu_\gamma, \sigma_\omega^2, \sigma_\gamma^2, \rho)$, where μ_x and σ_x are the mean and standard deviation of $\log(x)$. ρ is the correlation coefficient for the bivariate normal distribution of $\log(\omega)$ and $\log(\gamma)$. The skill distribution is typically approximated by the distribution of wage rates. We do so as well. We observe the mean and standard deviation on micro data (the LINDA data source) on monthly wages in Sweden (full time equivalents) as of 2009. We do not, however, observe the moments of the shifting cost distribution; they must be calibrated somehow.

The basic strategy is to calibrate the shifting cost distribution by choosing μ_γ and σ_γ in such a way that the actual share of 'shifters' is reproduced, conditional on the actual Swedish wage distribution, the actual Swedish tax system and a particular distribution of elasticities. Two parameters are unknown to us. For convenience, we assume that the variances of $\log(\omega)$ and $\log(\gamma)$ are the same.[8] Ultimately, we therefore solely calibrate μ_γ.

We set our target, i.e. the actual fraction of shifters, to be 5 %. Alstadsæter and Jacob (2013) report that 2.8% of Swedish individuals aged 18-70 are active shareholders in closely held corporations 2000-08. Considering the fact that the share has increased over time and that our wage data covers a younger sample (individuals aged 18-65) we think that 5 % is a reasonable number to use in the calibration.

We calculate marginal labor income tax rates and marginal dividend income tax rates for all individuals in the LINDA sample of 2009. We do not only consider the statutory tax rates, but also the payroll tax rate and the corporate tax rate.[9] In the LINDA wage sample the average marginal labor tax rate amounted to 0.505, whereas the average (constant) marginal capital tax rate amounted to 0.410. Hence, we set $\tau_P = 0.505$ and $\tau_C = 0.410$ when calibrating the model.

We impose our baseline assumption regarding the labor supply elasticities; the elasticity is 0.1 for the lowest-skilled individual and 0.5 for the highest-skilled individual, and the elasticity is linearly increasing in ω. Then we find that the fraction of shifters is 5% when $\mu_\gamma = 11.795$. Since our model is very stylized we want to emphasize that we in no way consider this to be a valid 'estimate' of the average shifting costs. The purpose of the calibration exercise is to get a reasonable figure that facilitates qualitative insights. The parameters used in the simulations are summarized in Table 2.

[8] The correlation coefficient for the transformed distributions is given by $\frac{e^{\rho\sigma_\omega\sigma_\gamma}-1}{\sqrt{[e^{\sigma_\omega^2}-1][e^{\sigma_\gamma^2}-1]}}$, where the natural exponential function is denoted by e. When $\sigma_\omega = \sigma_\gamma$ the correlation coefficient for the transformed distributions is always relatively close to ρ, and identical for $\rho = 0$ and $\rho = 1$.

[9] If an owner of a closely held corporation distributes profits as wage income her marginal tax rate is $\frac{\tau_{personal}+\tau_{payroll}}{1+\tau_{payroll}}$. If she distributes profits as dividend income her marginal tax rate is $\tau_{corporate} + \tau_{dividends} - \tau_{dividends} \times \tau_{corporate}$. In 2009 $\tau_{corporate} = 0.263$, $\tau_{dividends} = 0.2$ and $\tau_{payroll} = 0.3142$ were all proportional, whereas $\tau_{personal}$ varied between 0 and 0.565. When calculating $\tau_{personal}$ we accounted for the Swedish central government tax, local tax, basic allowance and the earned income tax credit.

Table 2: Parameter values used in the simulations

	$\log(\omega)$	$\log(\gamma)$
μ	10.194	11.795
σ	0.302	0.302

Note: Moments of $\log(\omega)$ have been picked from LINDA data as of 2009, whereas the moments of $\log(\gamma)$ have been calibrated.