WID.world WORKING PAPER N° 2019/11

How Much are the Poor Losing from Tax Competition: The Welfare Effects of Fiscal Dumping in Europe

Mathilde Muñoz

August 2019

World Inequality Lab

How Much are the Poor Losing from Tax Competition: The Welfare Effects of Fiscal Dumping in Europe*

Mathilde Muñoz, Paris School of Economics

August 2019

Abstract

This paper quantifies the welfare effects of tax competition in an union where individuals can respond to taxation through migration. I derive the optimal linear and non-linear tax and transfer schedules in a free mobility union composed by symmetric countries that can either compete or set a federal tax rate. I show how in the competition union, the mobility-responses to taxation affect the redistributive capacity of governments through several mechanisms. I then use empirical earnings' distribution and estimated migration elasticities to implement numerical calibrations and simulations. I use my formulas to quantify the welfare gains and losses of being in a tax competition union instead of a federal union, and show how these welfare effects vary along the earnings distribution. I show that the bottom fifty percent always loses from tax competition, and that being in a competition union rather than in a federal union could decrease poorer individuals welfare up to -20 percent.

*mathilde.munoz@psemail.eu. I thank my advisor Thomas Piketty for his continued guidance and supervision, and fascinating discussions about this project. I thank Antoine Levy for more than helpful comments and suggestions.

1 Introduction

Freedom of movement within the European Union is the cornerstone of union citizenship, and has been at the core of European integration since the Treaty of Rome in 1957. While free mobility of individuals between member states has now been effective for sixty years, tax coordination between European countries is still non-existent: member states set their tax rates at the national level, and the tax and transfers schedules remain outside the scope of the union policy.

In this paper, I explore the welfare consequences for European citizen of living in a free mobility union characterized by tax competition, rather than in a free mobility union with uniform taxation - a *federal union*. I start with a simple theoretical framework, where the free mobility union is composed by perfectly symmetric countries and where redistribution is fully consumed, that is to say making the (conservative) assumption that public spending does not produce anything else than immediate consumption for residents. In my model, the main difference between the competition and the federal union comes from tax-driven mobility. When countries set an uniform tax rate, individuals' location choices cannot be affected by differences in country-level taxes. By contrast, when countries engage in tax competition, individuals react to unilateral changes in taxation rates through migration. I start by deriving the optimal tax and transfer schedules when countries are competing and when countries are setting a uniform federal tax rate. I theoretically emphasize how mobility responses to taxation - *migration elasticities*- affect the redistributive ability of competing governments. The optimal tax formulas shed light on two main mechanisms that affect redistribution in the presence of tax-competition. Redistribution in the competition union is lowered because tax-driven migration reduces the amount of taxes that can be collected by the government, as individuals at the higher end of the income distribution respond to higher tax rates with emigration (*revenue-channel*). Redistribution in the competition union is also lower because higher rates of taxation increases the absolute number of transfer beneficiaries through mobility at the bottom of the income distribution, leading to a lower level of transfer per individual (*transfer channel*).

I then take the optimal tax formulas to the data, and quantify how individuals' welfare is changed between a tax competition union and a federal union. I focus my analysis on the *distribution* of the welfare effects of tax competition across earnings levels. I show that individuals are differently affected by tax competition depending on their income level, and that the magnitude of the welfare effects of tax competition varies with the intensity of tax-driven mobility and the redistributive tastes of the government. My results show that the bottom fifty percent always loses from tax competition, and that being in a competition union rather than in a federal union decreases poorer individuals welfare up to -20 percent.

This paper is related to a vast literature on the optimal taxes and transfers schedule, starting with the seminal work of Mirrlees [1971], studying how behavioural responses to taxation affects the optimal tax policy of governments. My analysis relies on the sufficient statistics approach developed by Piketty [1997] and Saez [2001], and extensively summarized in Piketty and Saez [2013]. The introduction of migration in the canonical model of optimal taxation dates back Mirrlees [1982], and has been generalized by the contribution of Lehmann et al. [2014] that shows how the shape of the optimal income tax schedule of a Rawlsian government is affected in the presence of migration.

2 Linear Tax Schedule

I start the analysis with a simple linear tax framework, that considerably simplifies the derivation of the optimal tax formulas but allows to capture the equity-efficiency trade-off at the heart of the optimal taxation problem in the presence of migration. As we shall see later, the linear tax problem is closely related to the non-linear tax problem, and it is therefore useful to derive the optimal tax and transfer in the case of a linear tax instrument.

2.1 Individuals' Problem

I consider individuals who are heterogeneous with respect to their preferences and skills. I follow the approach of Piketty and Saez [2013], where individual i has an utility $u_i(c_i, y_i)$ that is increasing in consumption and decreasing in earnings, as earnings require labour supply. Individuals are heterogeneous with respect to their skills w_i that are continuously distributed in the economy. There is a mass N_i of type-i individuals who are characterized by the same preferences and skills

$u_i(c_i, y_i)$. The total number of taxpayers in one country is given by $\sum_i N_i = N$, and the aggregated income denoted as $Y = \sum_i N_i y_i$. A type-i individual is endowed with skills w_i, and receives a pre-tax income y_i that is a combination of his exogeneous ability w_i and his amount of effort l_i, such that $y_i = l_i w_i$. The government observes pre-tax earnings, but the abilities of individuals are private information. The government sets a linear tax rate τ on observed earnings that is universally redistributed through a lump-sum transfer T_0, that can therefore be written as $\tau \times Y/N$. Individuals' budget constraint is therefore given by $c_i = (1-\tau)y_i + T_0$.

2.1.1 Labour Supply Decisions

Given their preferences and characteristics, individuals choose their labour supply at the intensive margin, which corresponds to their optimal amount of work l_i . Formally, they choose pre-tax earnings y_i that maximize $u_i(c_i = (1-\tau)y_i + T_0, y_i)$. Assuming no income effects, type-i individual utility $u_i(c_i, y_i)$ can be written as

$$u_i(c_i, y_i) = (1-\tau)y_i + T_0 - v_i(l_i) \tag{1}$$

The disutility from effort $v_i(l_i)$ is increasing and convex in effort l_i, and thereby in pre-tax earnings y_i. The individual-level optimality condition determines the earnings function $y_i(1-\tau)$, and the compensated labour supply elasticity captures the change in individual's earnings caused by a change in the net-of-tax rate $1-\tau$:

$$e_i = \frac{\partial y_i}{\partial(1-\tau)} \times \frac{1-\tau}{y_i} \tag{2}$$

The elasticity of earnings with respect to the net-of-tax rate e_i is structurally determined by individuals' preferences. When the tax system is linear, the individual chooses y_i that maximizes $y_i(1-\tau) + T_0 - v_i(y_i)$. The first order condition is simply given by $1-\tau = v_i'(y_i)$. The differentiation of the first order condition allows to link the definition of the elasticity of earnings e_i to the structure of individuals preferences such that $\frac{1-\tau}{y_i} \times \frac{\partial y_i}{\partial(1-\tau)} = \frac{v_i'(y_i)}{y_i v_i''(y_i)}$. By definition of the disutility of labour, the elasticity of gross earnings with respect to the net-of-tax rate is always positive.

2.1.2 Location Choices

In a free mobility union, individuals can move from one country to another. I assume that individuals make the decision to migrate conditionally on their labour supply decision. I start by considering two perfectly symmetric countries A and B that constitute the entire world economy. Agents have an idiosyncratic taste for residing in one country that is captured by the parameter θ_i^A for country A and $\theta_i^B$for country B. Migration is costly, and agents have to pay a migration cost m if they decide to migrate, meaning that m is equal to zero in the absence of migration. The utility of individuals residing in country A can therefore be written as $u_i^A = (1-\tau^A)y_i + T_0^A - v_i(l_i) + \theta_i^A - m$, and symmetrically in country B as $u_i^B = (1-\tau^B)y_i + T_0^B - v_i(l_i) + \theta_i^B - m$. Agents m

igrate from country A to country B if and only if they receive a higher utility in country B. Therefore, any agent residing in country A has to satisfy the following conditions:

$$u_i^A = (1-\tau^A)y_i + T_0^A - v_i(l_i) + \theta_i^A - m \tag{3}$$

$$u_i(c_i^A, y_i, \theta_i^A, m) \geq u_i(c_i^B, y_i, \theta_i^B, m) \tag{4}$$

Equation (3) and Equation (4) define together the mass of individuals in country A N_i^A in equilibrium. Equation (4) emphasizes how the taxation rate in country A affects location choices in this country, taking everything else as given. Migration decisions to country A are determined by the *overall tax liability* of individuals in this country, combining the amount of taxes paid $\tau^A y_i$ and transfers received T_0^A, by contrast to labour supply responses that are driven by marginal tax rates only in the absence of income effects. We can directly derive from Equation (4) that the density of individuals with type-i preferences who decide to locate in one country can be written as a function of the net-of-tax rate in this country such that $N_i^A(1-\tau^A)$ and $N_i^B(1-\tau^B)$[1]. The density of type-i individual in one country can be increasing or decreasing in the net-of-tax rate in this country depending on how type-i individual consumption is affected by the linear tax rate. Formally, the consumption of type-i individual $c_i = (1-\tau)y_i + T_0$ can be rewritten using the definition of T_0 as $c_i = y_i + \tau(Y/N - y_i)$. I follow Saez [2002] and define the *break-even point* as the income level

[1] In the linear model, writting the density function with respect to the net-of-tax rate rather than the consumption level considerably ease the problem exposure without loss of generality.

y_i such that $y_i = Y/N$ and at which transfers net of taxes are equal to zero. For any i such that $y_i < Y/N$, consumption is a *decreasing* function of the net-of-tax rate $1-\tau$, and N_i is therefore decreasing in the net-of-tax rate. Symmetrically, for any individual with $y_i > Y/N$, consumption is increasing in the net-of-tax rate and N_i is also increasing in $1-\tau$. Migration responses to taxation can be fully summarized in terms of elasticity concepts, and I define the migration elasticity as the change in the number of residents in one country when the retention rate is increased in this country:

$$\varepsilon_i = \frac{\partial N_i}{\partial(1-\tau)} \times \frac{1-\tau}{N_i} \tag{5}$$

The sufficient statistic ε_i summarizes the migration response of type-i individual to a change in the overall tax and transfer schedule at the income level y_i through a change in $1-\tau$. The intuition is that any increase in τ is redistributed to everyone through the universal demogrant T_0. In the absence of income effects, there is no labour supply changes implied by this additional redistribution. In the presence of tax-driven migration, this additional redistribution creates a behavioural response to taxation, even in the absence of income effects. As described by Equation (4) any unilateral change in the level of transfers in one country will affect location decisions through the implied change in utility differential. Therefore, the migration elasticity of type-i individuals ε_i captures the net effect of increasing the net-of-tax rate on location decisions, combining the effects of taxes and transfers on individuals' utility level. For individuals with $y_i < Y/N$, an increase in the net-of-tax rate $1-\tau$ leads to a net increase in consumption through transfers, by contrast to individuals with $y_i > Y/N$. These differential effects of $1-\tau$ on individuals' consumption and thus migration decisions enter in the model by being directly loaded in the sign of ε_i. An increase in the net-of-tax rate that translates to an increase in the level of transfer induces immigration of low income levels ($\varepsilon_i < 0$), and emigration of higher income individuals ($\varepsilon_i > 0$). Therefore, by contrast to the labour supply elasticity, the migration elasticity can either be positive or negative, depending on how individuals' earnings relate to the break-even point. If tax-driven response is exactly the same for all individuals, the migration elasticity will have the same value in absolute, but will be of opposite sign at each side of the break-even point. I show in the next section how the effects of tax-driven mobility on the number and the composition of tax payers separately affect the optimal tax rate set by the competing government.

In addition to taxation, location choices are of course also determined by the distribution of migration costs and idyosyncratic preferences. These parameters are taken as exogeneous to the tax policy, and are therefore not affected by changes in the net-of-tax rate.

2.2 Government Problem

The government sets the linear tax rate τ, and redistribute the collected revenue through a universal demogrant T_0. Summing individual earnings functions $y_i(1-\tau)$ over the total number N of taxpayers in the economy allows to obtain the aggregate earnings $Y = \sum_i N_i y_i$. Total income in the economy is thus determined by individual earnings and the number of taxpayers at each income level in the economy, that are both a function of $1-\tau$. It follows that the government budget constraint can be written as $R = Y(1-\tau)\tau$. This tax function sheds light on the effect of taxation on tax revenue. When the tax rate is equal to one, there is no incentives to work and the tax revenue is equal to zero. When the tax rate is equal to zero, aggregated earnings are maximized but cannot be redistributed. The guaranteed income level T_0 is determined in equilibrium by the total amount of tax revenue R and the linear tax rate set by the government.

2.2.1 Social Preferences

The government chooses the level of taxes τ in order to maximize a social welfare function. I follow the approach developped by Saez and Stantcheva [2016] and use the concept of *generalized social marginal welfare weights* where g_i measures how much the government values the marginal consumption of individual i. This formulation is conveniently very general, and the welfare weights are only defined up to a multiplicative constant as they measure only the relative value of consumption of individual i. Therefore, the government preferences for redistribution will be loaded in the weights g_i. The overall spectrum of possible preferences for redistribution, from low to infinite, will be loaded in the distribution of the weights g_i across earnings levels.

2.2.2 Tax Systems

I consider a free mobility union where symmetric countries can either compete or cooperate regarding the collection of their tax revenue. Symmetric countries are characterized by the same exogeneous distribution of skills and population size. Importantly, I start by assuming that there is no spillovers from integration, such that there is theoretically no differences between the autarky and the federal systems other than migration driven by exogeneous parameters such as migration costs and idiosyncratic preferences.

When the government is federal, it sets a uniform tax rate τ^f that is paid by everyone regardless of its residence in A or in B. As countries are perfectly symmetric, it is exactly equivalent to collect and redistribute the revenue at the country or the union level, as countries have the same average income level conditional on having the same federal tax rate τ^f. Reconsidering Equation (4) in the case where country A and country B impose the same federal rate, the difference in utility levels can only be driven by migration costs or individuals' preferences. It follows that in the federal union, the mass of taxpayers in each country is exogeneous to the taxation rate, as any change in the federal rate τ^f translates to a symmetric change in utility levels in both country, keeping the migration condition summarized by Equation (4) unchanged. Without any additional assumptions, as there is no tax-driven migration in the federal union, the federal tax rate is equal to the optimal tax rate in autarky.[2] The only behavioural response to taxes in the federal union is captured by the labour supply responses to taxation.

Rather than being part of a federal union, countries can choose to compete within the free mobility union. Tax competition means that countries set their respective tax rates and redistribute transfers separately, while individuals can freely locate in each country within the free mobility union. With competing countries, location decisions are affected by the competing linear tax rate set in country A τ_A^c and in country B τ_B^c as emphasized by Equation (4). Because of the tax competition, the population of taxpayers in each country is no longer independent from the taxation rate. The optimal tax rate of the competing economy τ^c is therefore affected by two behavioural responses to taxation: the intensive margin through labour supply responses to taxation, and the

[2] This is because wages are exogeneously determined, and even in the federal economy where migration may occur between two countries due to non-tax factors, a change in the tax rate applied to everyone does not distort migration decisions and is therefore not internalized in the government's maximziation problem. As we will see later, the independance of the federal tax rate to migration is likely to be changed in the case of endogeneous wages.

extensive margin through migration responses to taxation.

2.3 Optimal Linear Tax rates

In this section, I derive the optimal linear tax rate of the government in the two available systems: tax competition and federal union. I present the derivation of the optimal linear tax rates following the small tax deviation approach, but the formulas can also be derived by fully specifying the welfare maximization problem. The optimal linear tax rate is such that around the optimum no small reform can yield a welfare gain. The welfare gains from any tax deviation are quantified by weighting the money metric welfare gains or losses to each individual using these weights.

Proposition 1. *Optimal Linear Tax rate of the Federal Government:*

$$\tau^f = \frac{1-\bar{g}}{1+e-\bar{g}} \tag{6}$$

Proof. Where e denotes the income weighted average labour supply elasticity $e = \sum_i \frac{e_i N_i y_i}{Y}$ and $\bar{g}$ captures a weighted average of welfare weights $\bar{g} = \left(\sum_i N_i g_i y_i\right).\left(\sum_i N_i\right)/\left(\sum_i N_i g_i . \sum_i N_i y_i\right)$. The proof is formally derived in the Appendix, and intuitively below. □

Proposition 2. *Optimal Linear Tax Rate of the Competing Government:*

$$\tau^c = \frac{1-\bar{g}}{1-\bar{g}+e+\bar{\varepsilon}} \tag{7}$$

Proof. Where e denotes the income weighted average labour supply elasticity $e = \sum_i \frac{e_i N_i y_i}{Y}$, $\bar{\varepsilon}$ is a combination of the *income weighted* and *population weighted* average mobility elasticity such that $\bar{\varepsilon} = \sum_i \frac{\varepsilon_i N_i y_i}{Y} - \sum_i \frac{\varepsilon_i N_i}{N}$ and $\bar{g}$ captures a weighted average of welfare weights $\bar{g} = \left(\sum_i N_i g_i y_i\right).\left(\sum_i N_i\right)/\left(\sum_i N_i g_i . \sum_i N_i y_i\right)$, The proof is formally derived in the Appendix, and intuitively below. □

To derive the optimal tax rate in the presence of welfare weights, I consider an infra-marginal deviation in the tax rate $d\tau$ with no other effect on individuals' welfare than the effect on post-tax earnings. This is because of the classical envelop theorem argument, that implies that the change in individuals' labour supply after a small change in the tax rate does not change individuals' utility through disutility of work, as the optimal labour supply has been chosen at the optimum. The argument is the same for migration decisions. The welfare effect of the small tax deviation is therefore limited to its effects on post-tax earnings. The first effect of $d\tau$ on individuals' welfare is given by the increase in taxes paid by everyone $-\sum_i N_i g_i y_i d\tau$. The second effect on welfare is created by the change in transfers $\sum_i N_i g_i dT_0$. When N is exogeneous to the tax reform, the change in the universal demogrant is $dT_0 = dR/N$. In the case of a federal government, we can normalize the total population $N = \sum_i N_i$ to one without loss in generality, and use $dR = dT_0$. What is the effect of the small tax deviation on dR? The small tax reform creates a mechanical increase in tax revenue $d\tau Y$. As pre-tax earnings are endogeneously determined by the labour-leisure trade-off, $d\tau$ causes an additional change in pre-tax earnings because of behavioural responses to taxation. Using the definition of the labour supply elasticity, the change in tax revenue due to labour supply responses is $-\sum_i \frac{\tau}{1-\tau} N_i y_i e_i d\tau$. I rewrite this effect $-e\frac{\tau}{1-\tau} Y d\tau$ where $e = \sum_i \frac{N_i y_i e_i}{Y}$ is the income weighted labour supply elasticity. The total effect of the small tax change on tax revenue is therefore $dR = Y d\tau (1 - e\frac{\tau}{1-\tau})$. Using the expression for dR derived before, and the fact that at the optimum the net welfare effect of $d\tau$ is zero gives $\sum_i N_i g_i y_i d\tau = \sum_i N_i g_i (1 - e\frac{\tau}{1-\tau}) Y d\tau$, which is equivalent to $1 - e\frac{\tau}{1-\tau} = \bar{g}$ with $\bar{g} = (\sum_i N_i g_i y_i) . (\sum_i N_i) / (\sum_i N_i g_i . \sum_i N_i y_i)$, that is a simple discretization of the standard formula $\bar{g} = (\int_i g_i y_i) / (\int_i N_i g_i . \int_i N_i y_i)$ with population normalized to one developed in Piketty and Saez [2013] and Saez and Stantcheva [2016].

When countries are competing, the total number of taxpayers becomes endogeneous to the tax system. The small tax deviation creates two behavioural responses to taxation: individuals respond to the tax reform through labour supply changes and migration decisions. The envelop theorem holds for location choices. Because the tax deviation considered is small enough, there is no effect on individuals' welfare through the change in migration decisions implied by $d\tau$. How is welfare changed in the presence of tax-driven migration by the small tax reform? Similarly than in the federal case, the welfare effect can be decomposed between the additional taxes paid for everyone in the economy and the change in the universal demogrant dT_0.

By contrast to the analysis in the federal union, in the presence of tax competition, the total mass of individuals in the economy cannot be normalized to one without making the restrictive assumption that migration decisions change the composition of the population keeping the total number of taxpayers constant. In the competing union, tax-driven migration modifies the amount of transfers received by residents *(i)* by changing the amount of taxes that can be collected and *(ii)* by changing the number of transfer beneficiaries among which the tax revenue is split. I show formally in the Appendix A.1 and Appendix A.2 how the *revenue-maximizing rate* differs from the *transfer-maximizing rate* because of this *transfer channel* that changes the absolute number of individuals who share the government revenue. Below, I develop the same intuition by using the small tax deviation approach.

How is the amount that can be redistributed among residents changed by the small tax reform $d\tau$ in the competing union? Using the definition of the migration elasticity, the change in the mass of type-i taxpayers after a small tax deviation is given by $-\sum_i \frac{N_i}{1-\tau}\varepsilon_i d\tau$, where ε_i is allowed to be positive or negative. This migration response of type-i individuals generates a change in taxes collected equal to $-\sum_i \frac{N_i}{1-\tau}\varepsilon_i d\tau \times y_i \times \tau$, as individuals come or leave with their overall tax liability $y_i\tau$. This term captures the *revenue effect* of tax-driven migration. In the presence of tax competition, any change in the linear tax rate changes the amount collected by the government because of the gais (or losses) of tax liabilities through mobility. The amount of revenue that can be redistributed to individuals in the economy is also changed by the absolute number of beneficiaries that is endogeneously affected by the reform. When the number of taxpayers is changed by $d\tau$, the reform generates a fiscal gain through the change in the absolute number of transfers' beneficiaries $\sum_i \frac{N_i}{1-\tau}\varepsilon_i d\tau \times T_0$. Note that for individuals below the break-even point, this term is negative ($\varepsilon_i < 0$) and captures the additional redistribution cost of bottom earners who move to the country where the transfer is increased. The overall effect of tax-driven migration on the amount that can be redistributed to everyone remaining in the country is therefore given by $-\left(\sum_i \frac{\tau}{1-\tau}\varepsilon_i N_i y_i d\tau - \sum_i \frac{\tau}{1-\tau}\frac{N_i}{N}\varepsilon_i Y d\tau\right)$. Summing this to the mechanical change in tax revenue of residents $Yd\tau$ and the labour supply effect $\sum_i \frac{N_i}{1-\tau}e_i d\tau \times y_i \times \tau$ gives the total change in the amount that can be redistributed to the total mass residents $d\tau \times Y \times (1 - \frac{\tau}{1-\tau}\sum_i \frac{e_i N_i y_i}{Y} - \frac{\tau}{1-\tau}\sum_i \frac{\varepsilon_i N_i y_i}{Y} + \frac{\tau}{1-\tau}\sum_i \frac{\varepsilon_i N_i}{N})$ and each individual remaining in the economy has a change in transfer received equal to $\frac{1}{N} \times d\tau \times Y \times (1 - \frac{\tau}{1-\tau}\sum_i \frac{e_i N_i y_i}{Y} - \frac{\tau}{1-\tau}\sum_i \frac{\varepsilon_i N_i y_i}{Y} + \frac{\tau}{1-\tau}\sum_i \frac{\varepsilon_i N_i}{N})$. Denoting $\varepsilon = \sum_i \frac{\varepsilon_i N_i y_i}{Y}$ the income-weighted average migration elasticity, $e = \sum_i \frac{e_i N_i y_i}{Y}$ the income-weighted labour supply elasticity and $\varepsilon_p = \sum_i \frac{\varepsilon_i N_i}{N}$ the population-weighted average migration elasticity, the transfer maximizing rate is such that $\frac{1}{N} \times d\tau \times Y \times (1 - \frac{\tau}{1-\tau} e - \frac{\tau}{1-\tau}\varepsilon + \frac{\tau}{1-\tau}\varepsilon_p) = 0$, which is equivalent to $\frac{\tau}{1-\tau} = \frac{1}{e+\bar{\varepsilon}}$where $\bar{\varepsilon} = \varepsilon - \varepsilon_p$ is a combination of the income-weighted and population-weighted average mobility elasticity. I discuss in the next paragraph the underlying mechanisms captured by this aggregated mobility parameter.

Let's finally consider the welfare maximizing linear tax rate such that the welfare gain of $d\tau$ is zero. How should we compute the welfare effect of $d\tau$ in the competing union? I discuss in details in the AppendixA.3 the normative challenges related to welfare aggregation, and definition, in an union with migration, because of the endogeneous size of the population. I derive as a baseline specification the welfare maximizing linear rate as the linear tax rate that maximizes the welfare of residents. The formal derivation of the optimal linear tax rate is presented in the Appendix A.3, and can also be derived using the small perturbation approach. The small tax deviation generates a loss in welfare for individuals remaining in the country after the reform due to the increase in taxes paid equal to $\sum_i N_i y_i g_i d\tau$, and a change in welfare due to the change in transfers received equal to $\sum_i N_i g_i \frac{1}{N} \times d\tau \times Y \times (1 - \frac{\tau}{1-\tau} \sum_i \frac{e_i N_i y_i}{Y} - \frac{\tau}{1-\tau} \sum_i \frac{\varepsilon_i N_i y_i}{Y} + \frac{\tau}{1-\tau} \sum_i \frac{\varepsilon_i N_i}{N})$. The total welfare effect of the small tax deviation is therefore $\sum_i N_i g_i d\tau \times Y \times \frac{1}{N} \times (1 - \frac{\tau}{1-\tau} \sum_i \frac{e_i N_i y_i}{Y} - \frac{\tau}{1-\tau} \sum_i \frac{\varepsilon_i N_i y_i}{Y} + \frac{\tau}{1-\tau} \sum_i \frac{\varepsilon_i N_i}{N}) - \sum_i N_i g_i y_i d\tau$. Summing the two welfare effects to zero yields the optimal linear tax tax formula with welfare weights described in 2 and derived formally in the the Equation (24) of the Appendix section A.3. Importantly, the averaged welfare weight $\bar{g} = (\sum_i N_i g_i y_i) . (\sum_i N_i) / (\sum_i N_i g_i . \sum_i N_i y_i)$ depends of the densities of residents N_i that are taken as given for the aggregation of welfare.[3]

The optimal linear rate of the competing union is a function of the mobility parameter $\bar{\varepsilon}$ that is a combination of the income-weighted and population-weighted mobility parameter. Note that the case where the *absolute* number of taxpayers is unchanged, $\sum \varepsilon_i \frac{Ni}{N} = 0$ and we are back to

[3]I discuss in the Appendix A.3 the normative challenges related to the aggregation of welfare in the open economy, explained by the fact that in the open economy, the total welfare can theoretically be increased by (i) increasing the consumption of individuals in this country but also by (ii) increasing the number of individuals who enter in the sum of individuals' welfare. To avoid any considerations due to population size other than its effects on the amount of transfers that can be redistributed, I consider a government that maximizes the welfare of a given population, taking into account the effect of tax-driven mobility responses on the consumption of this population. Typically, this welfare function would correspond to a government that would maximize the welfare of non-movers, taking into account the effects of movers on non-movers consumption through the amount of redistribution that can be acchieved. I discuss this assumption in the Appendix, and derive in Equation (26) how the linear tax rate would be changed if one would to relax this assumption.

case where the optimal tax rate is only a function of the income-weighted mobility elasticity. In this specific case, the revenue-maximizing rate is equivalent to the transfer-maximizing rate, as in the classical federal case. In the general case, the effect of tax-driven mobility can now be decomposed between two terms. The first term captures the effect of migration on the tax revenue, through the *income weighted mobility parameter* $\varepsilon = \sum_i \varepsilon_i \frac{N i y_i}{Y}$. The second term captures the effect of tax-driven mobility on the absolute number of taxpayers through the *population-weighted mobility parameter* $\varepsilon_p = \sum_i \varepsilon_i \frac{Ni}{N}$. The net effect of migration on the optimal linear tax rate is therefore summarized by $\bar{\varepsilon} = \varepsilon - \varepsilon_p$. The first term captures the *revenue channel* that is to say the change in tax revenue collected caused by mobility responses to taxation. The second term captures the *transfer channel,* that is to say the change in the number of transfer beneficiaries caused by tax-driven mobility. The net effect of type-i individuals' mobility on the universal demogrant thus depends of the importance of their relative income compared to their relative weight in the population. Said differently, the government weights the mobility response of type-i individual by taking the difference between type-i individuals' fiscal gain and cost.

To illustrate the implications of the weighting of the migration elasticity ε_i, I discuss the implications of two extreme assumptions on the distribution of the mobility parameter ε_i. Let's first make the assumption that only bottom earners react to taxation through migration. This could be the case if, for instance, top earners have a very strong attachement to their national labour market, and do not react to taxation through migration, while bottom earners can easily move across borders. For bottom earners, consumption is an increasing function of the linear tax rate τ, and a decreasing function of the net-of-tax rate. Therefore, the stock of bottom earners N_b is a *decreasing* function of the net-of-tax rate, and it follows that ε_b is negative. What would be the optimal linear tax rate of the government in the case where tax-driven mobility is exclusively coming from bottom earners that would change their location decisions if transfers are increased? With mobility responses concentrated at the bottom of the distribution, as earnings of bottom earners are close to zero, the uniform mobility parameter is $\bar{\varepsilon} = -\frac{\varepsilon_b N_b}{N}$. As ε_b is negative, the uniform mobility parameter $\bar{\varepsilon}$ is positive and the resulting optimal linear tax rate in competition is lowered by tax-driven migration coming from the bottom of the distribution. What happens in the opposite situation, when tax-driven mobility only comes from the very top of the income distribution? This assumption could be verified if bottom earners have very strong migration costs while rich people can easily change their residence country. At the top of the income distribution, consumption is a decreasing function of the linear tax rate, and thus a decreasing function of the net-of-tax rate. Top earners' mobility elasticity ε_t is thus always positive. When the country considered is large enough, the population weight of very high earners becomes negligible, and the uniform mobility parameter entering in the optimal tax formula can thus be approximated by $\bar{\varepsilon} = \varepsilon_t N_t y_t / Y$ that is always positive. The main take away from these two examples is that no matter towards what side of the earnings distribution ε_i is skewed, the resulting optimal linear tax rate in competition is always lowered by tax-driven migration, leading to less redistribution in the competition union compared to the federal union. The mechanisms leading to less redistribution in these two extreme cases are different, and both emphasize the trade-offs faced by governments competing in a free mobility union with no cooperation. In the case where only bottom earners move in response to tax changes, the optimal amount of redistribution is exclusively lowered by the *transfer channel* of tax-driven mobility, that is to say the additional immigration of individuals who benefit in net of the tax and transfer system after an increase in the taxation rate. In the case where mobility responses to taxes are only coming from the top of the income distribution, the optimal amount of redistribution is exclusively limited by the *revenue channel* of tax-driven mobility, that is to say the amount of tax collected that is lost because of the emigration response to the increase in the tax rate.

2.4 The Welfare Effects of Tax Competition

I now turn to the quantification of the welfare effects of tax competition. For this purpose, I use the theoretical formulas derived in the previous section to quantify the welfare of individuals in the two available tax systems: tax competition and federal union. As described before, individuals derive an utility $u_i(c_i, y_i)$ that is decreasing with earnings due to disutility for work, and increasing in consumption. The welfare effect of tax competition compared to the federal union will be given by the change in individuals' utility from one system to another. This change in tax system will affect individuals' utility through three channels.

The choice of tax system will first affect individuals' pre and post tax earnings. The optimal tax rates set in each of the two systems differ because of the migration parameter, leading to different amount of taxes paid for a given level of income. In addition, due to labour supply responses to taxation, the differences in tax rates between the two systems will also lead to differences in individuals' pre-tax income.

Second, the change in labour supply induced by the change in tax systems will affect individuals' welfare through disutility for work.

Third, the choice of tax system will affect the amount of transfers received by individuals. The choice of tax system will affect pre-tax aggregated income that can be taxed by the governement, because of the changes in individuals' labour supply decisions. The choice of tax system will also affect the amount of transfers received by individuals through the change in the rate at wich the aggregated pre-tax earnings can be taxed in order to be redistributed to everyone.

Importantly, as I start by considering two *perfectly symmetric* competing countries, the competition tax rates set in equilibrium are perfectly similar. In this case, the density of tax-payers in each tax bracket is supposely unchanged, because the neighbouring country exactly mimic the other country tax policy. The competing tax rates are thus similar in the symmetric equilibrium. This implies that in the symmetric equilibrium, there are no welfare costs of tax competition through the change in taxpayers densities, transfers beneficiaries or migrations costs. The only difference with the federal union is the change in the optimal linear tax rates, as government take into account the fact that individuals can react to taxation through migration, without anticipating the tax rate set by the competing country, as in a very crude illustration of a Nash equilibirum. The computed welfare costs therefore correspond to the welfare effects of tax competition through the *migration threat*. Even if there is ultimately no tax-driven migration in the symmetric equilibrium, the welfare is changed through the change in the tax and transfer schedule implied by tax competition, and the fact that government internalizes individuals' migration threat. The symmetric equilibrium analysis is therefore very useful to estimate a *lower bound* for the welfare effects of tax competition, and to emphasize how competition affects individuals' welfare only through the incentives given to the government to lower its tax rate because of the competition. I will investigate the welfare effects of tax competition in an asymmetric equilibrium with endogeneously changed densities in the future.

2.4.1 Methodology

There are three key factors that determine the optimal linear tax and transfer schedules, and that are necessary to implement welfare calibrations: the behavioural labour supply and migration elasticities, the redistributive tastes of the government, and individuals' underlying preferences that determine the behavioural elasticities.

To quantify the welfare effect of the choice of tax system, it is necessary to make some functional form assumptions regarding the primitives of the model, that is to say individuals' utility functions. I start with a standard quasi-linear utility function with no income effects

$$u_i(c_i, l_i) = c_i - \frac{l_i^{1+k}}{1+k} \tag{8}$$

In that case, the compensated labour supply elasticity is equal to $\frac{1}{k}$, and the value of parameter k is chosen in order to be consistent with empirical values of e. Individuals have heterogeneous abilities. Formally, they are endowed with skills w_i such that for every individual $y_i = w_i l_i$. Using the first order condition of the individual problem, it is possible to express the earnings as a function of the labour supply elasticity, the tax rate and individuals' ability:

$$y_i = w_i^{e+1}(1-\tau)^e$$

In the absence of income effects, the pre-tax earnings of individuals are not affected by the level of the universal transfer. I follow the approach developed in Saez [2001] that consists in using this expression to retrieve the exogeneous distribution of skills using the observed distribution of earnings, the current tax rate and a chosen distribution of e. I use the current distribution of earnings in France taken from the World Inequality Database and an approximation of the actual linear tax rate of 50 percent, that roughly corresponds to the share of national income that is taxed. With the calibrated exogeneous distribution of skills at hand, it is possible to compute the welfare of individuals under different tax systems (federal or competition), and scenarios (varying elasticities values and distribution and government redistributive tastes), taking the distribution of skills as fixed conditionally on the distribution of labour supply elasticities. This methodology allows to take into account all changes in the earnings distribution (and thus collected tax revenue) that are caused by changes in the tax rates due to different tax systems considered, but also by different assumptions on elasticities' distribution. Regarding the value of the labour supply elasticity e, I use for the calibrations a constant value of 0.25 that is in line with the value widely used and estimated in the literature.

Regarding preferences for redistribution, a first case to consider is the most redistributive government, that is to say a Rawlsian government that only values the welfare of the bottom fifty percent such that $g_i = 1$ for any i in the bottom fifty percent while $g_i = 0$ for anyone else. It is then possible to consider different shape of the government preferences for redistribution, through variations in the value of the parameters g_i, and therefore $\bar{g}$. I consider two types of government: a *highly redistributive government* that values the welfare of each individual in the bottom fifty percent five times more than the welfare of individuals in the other deciles, and a *moderately redistributive government* that values the welfare of each individual in the bottom fifty percent two times more than individuals in the other deciles.

The last parameter needed, and the most central in the analysis, is the migration elasticity parameter. As showed in the previous section, the optimal linear tax rate depends on the overall mobility parameter $\bar{\varepsilon}$. This global mobility parameter is determined by the income-weighted average mobility elasticity and the population-weighted average mobility elasticity, that both depend on one relevant sufficient statistics: the migration elasticity with respect to the net-of-tax rate at each income level. The policy-relevant parameter is therefore ε_i, the elasticity of the *stock* of type-i individuals with respect to the net-of-tax rate. As emphasized by Kleven et al. [2019], there is a lack of empirical evidence on the empirical value of ε_i, especially for broad labour market segments and low levels of income. Importantly, as underlined by Kleven et al. [2019], ε_i is not a structural parameter but is affected by many environmental factors, such as the size of jurisdictions, current differences in tax rates, and levels of cooperation. The values of the elasticity ε_i may be varying over time, and across countries. For now, there is little empirical evidence on *cross-country* mobility responses to taxation. This is mainly because data on international migration flows are very hard to obtain, and because tax changes and mobility responses are likely to be endogeneous, and it is therefore difficult to find an empirical setting allowing to estimate a causal effect of taxation on mobility.

Two seminal contributions have managed to get around these empirical challenges by using original tax reforms and individual-level data for specific occupations allowing to track individuals' residence. Kleven et al. [2013] use data on the international carreer of football players in order to track their mobility choices, whileAkcigit et al. [2016] make use of international patents data to infer inventors' residence mobility. Kleven et al. [2013] and Akcigit et al. [2016] estimate sizeable elasticities of migration for top earners with specific occupations (football players and inventors), and find that mobility responses to taxation are especially large for foreigners, with elasticities around one, or above. This finding is confirmed by Kleven et al. [2014] who study the effect of a preferential tax scheme targeted on top earners immigrants in Denmark, and find a migration elasticity of 1.5. The reason why these studies have distinguished the mobility responses to taxation between foreigners and domestic is because they originally exploited quasi-natural variations stemming from tax reforms targeted on foreigners. However, the parameter of interest for the revenue-maximizing government is the elasticity of the overall stock of top earners, rather than the elasticity of the flows, or the foreigners elasticity. The stock elasticity will of course be lower, as by definition it relates to a larger base. For instance, Kleven et al. [2013] estimate that the elasticity of the number of football players with respect to the net-of-tax rate (the *uniform* elasticity) is between 0.1 and 0.4 on average, while the migration elasticities of foreigners alone is 0.7. Another strand of literature has focused on within-country mobility responses to taxation, exploiting the effects of regional-level variations in tax rates on individuals' mobility. Studying within-US mobility of inventors, Moretti and Wilson [2017] estimate an elasticity of the flow of inventors with respect to personal income tax rate of 1.5, that translates to a lower stock elasticity of between 0.4 and 0.5. In a recent contribution, Agrawal and Foremny [2018] exploit regional-variations in the the level of personal income tax rates within Spain and find that the elasticity of the stock of top earners is around 0.8. In a recent work, Muñoz [2019] estimates migration elasticities for the top ten percent employees of 26 European countries. The results show that the location choices of European top ten percent employees are significantly affected by variations in top income tax rates. This translates to a large elasticity of migration of foreigners (around 1.5), and a much lower uniform migration elasticity, that is between 0.1 and 0.4 on average.[4]

[4]The estimated elasticities are presented in Table B.I for a given set of European countries. The migration elaticities for the top ten percent range on average from 0.15 to up to 0.8 for some countries, and are heterogenous across member states, due to countries sizes and tax bases characteristics. For France, the migration elasticity for the top decile is estimated in the range [0.30;0.45].

Regarding migration responses at the bottom of the income distribution, there are very few empirical studies that have tried to quantify migration responses to taxation for all earners, and even fewer studies that have looked at migration responses of bottom earners to transfers. Some papers have however found that elderly migration within the US may have been partially driven by taxes and state-level policies in terms of amenities (Conway and Rork [2006], Conway and Houtenville [2001]). Overall, migration responses of low and middle earners to taxes and transfers remains a blackbox.

Regarding the results of the empirical literature described above, I consider an interval for the value of ε_i of [0.1;0.4]. These values are small, as they are far from the unity, and are close to the values that have been estimated by the literature for standard labour supply elasticities. Of course, the magnitude of ε_i could be much higher if the migration area is restricted to very small jurisdictions, or a highly integrated set of countries. In this case, the migration elasticity could be closer to the higher stock elasticity that has been estimated for within-country mobility responses to taxes, as in Martinez [2017] or Agrawal and Foremny [2018]. As I lack evidence on the value of ε_i for the entire range of income levels, I will start with the assumption that mobility responses are constant across earnings levels. I will then relax this assumption, and investigate various scenarios regarding the distribution of the mobility elasticity with income (the semi-elasticity).

2.4.2 Results

The baseline results of the numerical calibrations are presented in Table 1 and Figure 1 using the French earnings distribution. I use as the baseline scenario a constant labour supply elasticity of 0.25 and a migration elasticity that is constant across earnings types, meaning that all the ε_i have the same *absolute* magnitude. This implies that at all income levels, individuals have the same migration response to a change in their consumption caused by a change in taxes, except that tax changes have an opposite effect on individuals' consumption depending on which side of the break-even point they are. Using the theoretical formulas of the optimal linear tax rates, I compute the optimal linear tax rates in the federal and the competing unions for different redistributive tastes and various values for the migration elasticity. Given these optimal tax rate, I use the first order condition of the individual problem together with the exogeneous skills distribution to compute their optimal amout of labour supply and pre-tax earnings under each tax systems and scenarios. The transfers for each tax systems and scenarios are determined by the sum of these pre-tax earnings. I finally compute the welfare of individuals under each tax system using the utility specification presented in Equation (8). The welfare effect of tax competition is given by the change in welfare from going from a federal union to a competition union. These changes are summarized for the bottom ten and fifty percent in Table 1. I show the full distribution of welfare gains and losses created by tax competition across all earnings deciles in Figure 1. In Table B.III, I relax the assumption of constant elasticities and investigate the special case where tax-driven migration is only present at the top of the income distribution.

3 Non Linear Tax Schedule with Discrete Income Brackets

To take into how the tax progressivity of tax systems may be affected by the presence of tax competition, I develop a discrete version of the Mirrlees model, following Piketty [1997] and Saez [2002], in the presence of migration. The discrete non linear case has the advantage to be more tractable than the non linear model with continuous types, and to shed light on marginal tax rates at in the top and bottom brackets. Compared to the linear analysis developed before, it also allows to better take into account how the progressivity of the tax system can be affected by the magnitude and the distribution of migration elasticities.

3.1 Baseline Framework

Agents are indexed by k and are endowed with continuously distributed skills w_k, but there is a finite numbers of tax brackets, or occupations, $i = 0, ..., I$. Each tax bracket provides a wage y_i for $i = 0, .., I$, with $y_0 = 0$. Earnings y_i are increasing with i, and I start by assuming that there is a perfect substitution of labour types in the production function, implying that pre-tax wages are fixed. The government cannot directly observe individuals' skills and has to condition taxation on the observable income levels. The tax function is non-linear, and depends on the level of earnings, such that individuals in bracket-i have an overall tax liability $T(y_i) = T_i$.

Similarly than before, individuals have a utility function $u^k(c_i, k(i))$ that is a function of their tax-bracket choice $k(i)$ and the after-tax income level in their bracket c_i. Given their abilities, preferences and the tax and transfer schedule $(c_0, ..., c_I)$, agents choose their bracket i in order to maximize their utility. There is a population h_i of agents in each bracket $i \in I+1$, and $\sum_i h_i = N$ is the total population in the country. As in the linear model, agents respond to distortions created by taxation through labour supply changes, that are captured by their choices of income brackets. Labour supply decisions of individuals are thus loaded in the function $h_i(c_0, c_1 ..., c_I)$. I assume that the tastes for work embodied in the individual utilities are smoothly distributed so that the aggregate functions h_i are differentiable. As before, I consider as an important simplification the case with no income effects. In this case, increasing all after-tax consumption levels by a constant amount does not affect the distribution of individuals across brackets.

In this model, $T(y_i)$ embeds all taxes and transfers received by each individual. Compared to the simplistic linear case studied before, the more complex non-linear tax schedule allows to explore at a finer level variations in the profile of transfers depending on the level of revenues. The non-linear tax system is characterized by two key concepts. First, the universal demogrant $-T(0) = T_0$ that is distributed to everyone. Second, the marginal tax rate $T'(y_i)$ that captures the taxation on transitions from one bracket to another. The marginal tax rates, also called the phasing-out rates, allow to take into account the distributive effects of income taxation in the presence of behavioural responses to taxation. It also allows to capture at which rate the lumpsum grant is taxed away, and how the tax liability increases with earnings. A negative value for $T_i(y_i)$ means that individuals receive a net transfer from the government, and has to be distinguished from a negative value for the taxation rate on occupations transitions defined by $(T_i - T_{i-1})/(c_i - c_{i-1})$. There is an i for which $T_i = 0$, and as in the linear case I call this income level the break-even point. An important feature of the optimal tax problem is that it does not produce an explicit formula for the optimal transfer -T(0). The guaranteed income is determined in general equilibrium, and results from the optimal tax and transfer schedule T_i, and the empirical densities h_i determined by the tax schedule. The amount of taxes and transfers in each bracket T_i are set by the government in order to maximize a total welfare function. The government budget constraint is a function of the tax schedule and the endogeneously determined density of individuals in each income bracket:

$$R = \sum_{i=0}^{I} h_i T_i \tag{9}$$

Where R is exogeneously determined. There are several types of welfare functions that can be considered regarding the optimal tax problem. I start by studying a revenue-maximizing government. Then I consider a government that maximizes the total welfare in the economy, using the concept of generalized social welfare weights, where g_i captures the weight given to additional consumption for individuals in the tax bracket i, as in the linear analysis.

I extend the model to the case where individuals can react to taxation with migration. Similarly than in the linear model, in the presence of tax competition, taxation affects individuals' choices at the intensive margin regarding their choice of income bracket, and at the extensive margin through their migration choices. Conditional of being in the bracket i, individuals choose to migrate from A to country B if their utility is higher in country B. The migration condition considered in the linear case is unchanged, except that the tax system is now non-linear, and T_i directly loads the total tax liability of type-i individual:

$$u_i^A = y_i - T_i^A - v_i(y_i, l_i) + \theta_i^A - m \tag{10}$$

$$u_i(c_i^A, y_i, \theta_i^A, m) \geq u_i(c_i^B, y_i, \theta_i^B, m) \tag{11}$$

As before, the migration condition establishes that location choices are driven by differences in tax liabilities between the two countries, and the density of individuals in one country is therefore a function of its tax liability in this country. As in the linear framework, in the presence of tax competition the number of individuals in the national bracket i becomes a function of the tax and transfer schedule in this country $h_i(c_i)$. Migration decisions are thus driven by average tax liabilities, by contrast to occupation decisions that are driven by taxation on transitions from one bracket to another. The migration responses to taxation can be summarized in terms of elasticity concepts. By contrast to the linear case, individuals' consumption does not depend on a net-of-tax average rate, but of the amount of taxes paid and transfers received T_i. Therefore, I define the migration elasticity as the change in the density of type-i individuals locating in country A when their *disposable income* in country A is increased by one percent:

$$\xi_i = \frac{\partial h_i}{\partial c_i} \times \frac{c_i}{h_i} \tag{12}$$

Note that ξ_i is similar to ε_i for individuals with above the break-even point, and of opposite sign for individuals with income levels below the break-even point. As I define the migration elasticity with respect to consumption in the linear case, it is positive at all income levels.

3.2 Intensive Model

In this section, I present the canonical intensive model first developed by Piketty [1997] and Saez [2002] where individuals respond to taxation through labour supply choices only. In this model, a change in consumption level in any bracket i relative to another bracket $i-1$ induces individuals to switch from bracket i to bracket $i-1$. For simplicity, I assume that agents can only choose between adjacent occupations, and therefore, h_i is only a function of c_i, c_{i+1} and c_{i-1}. I define the elasticity of the number of individuals in bracket i with respect to the differences in consumption $c_i - c_{i-1}$

$$\eta_i = \frac{\partial h_i}{\partial(c_i - c_{i-1})} \times \frac{(c_i - c_{i-1})}{h_i} \tag{13}$$

As outlined by Piketty [1997], η_i captures the transition of individuals to bracket $i-1$ to bracket i when the difference in consumption between the two brackets is increased. The parameter η_i captures the participation of each individual in bracket i, and can be easily linked to the earnings elasticity e_i. Following Saez [2002], I use the relationship $\eta_i y_i = e_i(y_{i-1} - y_i)$. Hence, with intensive responses at the labour supply margin, a change in the tax liability in the bracket i will affect the transition rate between the bracket i and the adjacent occupations. The maximization of the government tax revenue leads to the first order condition:

$$h_i = T_{i-1}\frac{\partial h_{i-1}}{\partial(c_i - c_{i-1})} - T_{i+1}\frac{\partial h_{i+1}}{\partial(c_{i+1} - c_i)} + T_i\frac{\partial h_i}{\partial(c_i - c_{i-1})} - T_i\frac{\partial h_i}{\partial(c_{i+1} - c_i)}$$

The optimal tax liability of the revenue-maximizing government is given by:

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{h_i + h_{i+1} + ... + h_I}{h_i \eta_i} \tag{14}$$

The proof is formally derived in the Appendix A.4.1. Using τ_i the implicit marginal tax rate on bracket i such that $\tau_i = (T_i - T_{i-1})/(Y_i - Y_{i-1})$, where $1 - \tau_i = c_i - c_{i-1}/Y_i - Y_{i-1}$, and $a_i = Y_i/(Y_i - Y_{i-1})$, we obtain the formula for the optimal marginal tax rate on bracket i in the case where individuals can only respond to taxation through labour supply choices. This corresponds to the case of a federal government composed by symetric countries. When countries set the same tax and transfer schedule and are symetric such that before-tax salaries are equal, there is no differences in consumption between home and abroad that is affected by taxation. Therefore, migration decisions are independent from T_i, and do not affect the optimal tax formula.

Proposition 3. *Optimal Marginal Tax Rate of the Revenue-Maximizing Federal Government:*

$$\tau_i^f = \frac{h_i + h_{i+1} + ... + h_I}{h_i + h_{i+1} + ... + h_I + h_i a_i e_i} \tag{15}$$

Proof. The proof is formally derived in the Appendix. □

As outlined by Saez [2002], in the absence of extensive margin responses to taxation, the optimal tax liabilities are always increasing with i, and negative marginal tax rates are therefore never optimal.[5] As a result, the marginal tax rate in the first bracket is very high, and is maximal in the Rawlsian case with high redistributive taste. In complement to the formal maximization of the government problem given in the Appendix, it is possible to provide a simple and intuitive proof of Equation 15 by studying a small deviation in the tax schedule. Consider a small change dT for all brackets $i, i+1, ..., I$. This change in taxation changes $c_i - c_{i-1}$, leaving all other differences in consumption levels unchanged. This change in tax liabilities induces a mechanical increase in collected revenue equal to $(h_i + h_{i+1} + ... + h_I)dT$. The change in taxation also induces a behavioral response through the change in transition from bracket i to $i-1$. Using the definition of the participation elasticity, the mass of taxpayers in bracket i changes by $dh_i = -h_i \eta_i dT/(c_i - c_{i-1})$, inducing a loss in tax revenue of $dhi(T_i - T_{i-1})$. Summing the behavioural and mechanical effects to zero, we retrieve the formula for the optimal tax formula in the pure intensive model.

3.3 Extensive Model

I now turn to the extension of the canonical model, allowing migration responses to taxation. To emphasize how the tax driven mobility affects the non linear tax schedule, I start by considering

[5]The fact that negative marginal tax rates are never optimal would plausibly hold even considering participation margin at the bottom of the income distribution in that model. This is because the government is Ralwsian, which implies that the underlying welfare weights are very high for unemployed, and lower for poor workers.

the pure extensive model, where individuals can only respond to taxation through migration. This means that given their location decision, individuals' earnings are fixed. The only effect of T_i on h_i is therefore through migration responses to taxation summarized by Equation (10). The revenue-maximizing government chooses the optimal T_i taking into account the endogeneous changes in h_i due to migration responses to taxation, and the optimal tax and transfer schedule for type-i individuals satisfies:

$$\frac{T_i}{y_i - T_i} = \frac{1}{\xi_i} \tag{16}$$

The proof of Equation Equation (16) is derived in the Appendix. I give a simple intuition of the formula studying a small deviation in the tax schedule when individuals respond to taxation through migration only. In the pure extensive model, the change in tax liability T_i has an effect on migration decisions of individuals in bracket i, but does not affect transition to adjacent brackets. The change in T_i produces a mechanical increase in collected tax revenue $h_i dT_i$. The reform also creates a behavioral effect through migration, with a change in taxpayers mass of $dh_i = -h_i dT_i / c_i \xi_i$. Each individual emigrating from the country induces a loss of its overall tax liability T_i, and the overall behavioral effect is therefore $-h_i dT_i / c_i \xi_i T_i$. Summing behavioural and mechanical effects to zero gives the formula for the optimal tax and transfer schedule for each income bracket i. The marginal tax rate from bracket i to bracket $i-1$ is therefore given by

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{1}{ci - c_{i-1}}\left(\frac{y_i}{1+\xi_i} - \frac{y_{i-1}}{1+\xi_{i-1}}\right)$$

3.4 Optimal Linear Tax Rate in Tax Competition

I finally put together the pure intensive and extensive model to consider the case where individuals respond to taxation through migration and labour supply behavioural reponses. With a sligh abuse of notation, I rewrite the population function of taxpayers in the country as $h_i(c_i - c_{i-1}, c_{i+1} - c_i, c_i)$. The first two terms capture the effect of taxation on transition to adjacent tax brackets, while the last term captures the effect of taxation on utility differentials between home and abroad through consumption at home c_i. The derivation of the government tax revenue with respect to T_i is given by:

$$h_i = (T_i - T_{i-1})\frac{\partial h_i}{\partial(c_i - c_{i-1})} + (T_{i+1} - T_i)\frac{\partial h_i}{\partial(c_{i+1} - c_i)} + T_i\frac{\partial h_i}{\partial c_i}$$

Making use of the set of first order condition and the fact that $\frac{\partial h_i}{\partial(c_i - c_{i-1})} = \frac{-\partial h_{i-1}}{\partial(c_i - c_{i-1})}$, I obtain an expression for the optimal non linear tax rate chosen by a revenue-maximizing government in the competition union.

Proposition 4. *Optimal Revenue-Maximizing Marginal Tax Rate in Tax Competition:*

$$\tau_i^c = \frac{h_i(1 - b_i\xi_i) + h_{i+1}(1 - b_{i+1}\xi_{i+1}) + ... + h_I(1 - b_I\xi I)}{h_i(1 - b_i\xi_i) + h_{i+1}(1 - b_{i+1}\xi_{i+1}) + ... + h_I(1 - b_I\xi_I) + h_i a_i e_i} \tag{17}$$

Proof. The optimal tax rate formulas are formally derived in the Appendix. □

The formula for the optimal non linear tax rate in tax competition can also be retrieved by using a small deviation in the tax schedule. I consider a small deviation dT for all tax bracket $i, i+1, ..I$. As in the pure intensive model, this change in taxation modifies $c_i - c_{i-1}$, leaving all other differences in consumption levels unchanged. The change in tax liabilities induces a mechanical increase in collected revenue equal to $(h_i + h_{i+1} + ... + h_I)dT$. The change in taxation also induces a behavioral response through the change in transition from bracket i to $i-1$. Using the definition of the participation elasticity, the mass of taxpayers in bracket i changes by $dh_i = -h_i\eta_i dT/(c_i - c_{i-1})$, inducing a loss in tax revenue of $dhi(T_i - T_{i-1})$. In the presence of tax competition, there is an additional effect on tax revenue due to migration responses to taxation. As migration decisions are driven by overall tax liabilities, the change dT creates a migration response in all brackets affected by the change. Using the definition of the migration elasticity, the change in the number of taxpayers due to the tax reform can be written $-dT(\frac{h_i}{c_i}\xi_i + \frac{h_{i+1}}{c_{i+1}}\xi_{i+1} ... + \frac{h_I}{c_I}\xi_I)$. Any individual migrating from the country imposes a loss in tax revenue equal to its overall tax liability, such that the effect of tax-driven migration on tax revenue is equal to $-dT(\frac{h_i}{c_i}T_i\xi_i + \frac{h_{i+1}}{c_{i+1}}T_{i+1}\xi_{i+1} ... + \frac{h_I}{c_I}T_I\xi_I)$. Note that by contrast to the linear case, the *revenue and transfer channels* are simultaneously captured by the overall tax liability T_i, that can either be positive or negative. Summing the behavioural and mechanical effects to zero yields the optimal tax formula.

Letting one of the two elasticities e_i and ξ_i tend to zero in Equation (17), we retrieve the optimal formula of the pure extensive and intensive models. The optimal tax formulas make use of two fundamental parameters that relate to the two distortions implied by the intensive and extensive behavioural responses to taxation. The labour supply responses is weighted by the discrete equivalent of the usual Pareto parameter $a_i = Y_i/(Y_i - Y_{i-1})$ that captures the relative gain from the income bracket transition taxed at the marginal rate τ_i, while the migration response is weighted by the parameter $b_i = T_i/(y_i - T_i)$, emphasizing how location choices are driven by average tax rates. The migration wedge b_i is negative for individuals with income level below the break-even point, leading the overall migration response $b_i\xi_i$ of bottom earners to be negatively weighted in the optimal tax formula, making the link between the optimal tax formula derived in the linear framework.

The optimal marginal rates in the discrete linear case are, as it is well known, U-curved, with high and decreasing marginal rates at the bottom of the distribution, and increasing marginal rate at the top of the distribution. Of particular interests are the optimal top and bottom marginal tax rates. The optimal formulas in Equation (17) relate to the extreme case of a revenue-maximizing government, where the social planner exclusively values redistribution towards zero earners. As emphasized by Piketty and Saez [2013], this specific case of social preferences is likely to generate high values for the optimal bottom marginal tax rate τ_1. This is because increasing the transfers by increasing the phase-out rate produces a moderate behavioral cost, as individuals who decide to leave the labor force would have had low earnings should they work. As a result, in the presence of extensive and intensive responses to taxation, the optimal phase-out rate chosen by the Ralwsian government is positive, and very high.

As emphasized by Proposition 4, tax competition modifies the optimal tax and transfers schedule compared to the federal system summarized by Proposition 3, to an amount that is proportional to the migration tax wedges b_i that are negative for low income levels, and the mobility elasticities ξ_i. In the non-linear case, the effect of migration on the universal demogrant through the amount of taxes paid and the amount of transfer received is directly loaded in the term T_i that captures the net amount of taxes and transfers received or paid by the individual in bracket i. To simply illustrate this fact, and link the migration wedge b_i to the trade-off between the income and population weighted mobility parameter faced by the government in the linear case, I present the derivation of the optimal non-linear tax rate in the presence of migration by making the distinction between the amount of taxes paid $\tilde{T}_i$ and the amount of transfer received by everyone T_0. The resulting optimal linear tax rate is the same than the one presented in Proposition 3 with alternative notations. In this case, the government seeks to maximize $\frac{1}{\sum_i h_i}\sum_i h_i\tilde{T}_i$, where the total number of taxpayers $N=\sum_i h_i$ is endogeneously affected by the tax schedule through migration responses to taxation. With $N=\sum_i h_i$ and $R=\sum_i h_i\tilde{T}_i$, the first order condition of the government with respect to the amount of taxes $\tilde{T}_i$ can be rewritten as $\frac{1}{N^2}(\frac{\partial R}{\partial \tilde{T}_i}\times N-\frac{\partial N}{\partial \tilde{T}_i}\times R)=0$ which is equivalent to $\frac{\partial R}{\partial \tilde{T}_i}=\frac{\partial N}{\partial \tilde{T}_i}\times\frac{R}{N}$. Intuitively, the government first order condition with respect to $\tilde{T}_i$ indicates that at the optimum, the change in tax revenue due to a distortion in type-i individuals tax liability has to be offset by the change in transfer caused by the change in the number of type-i taxpayers implied by the tax reform, such that the net effect of the reform is equal to zero in the optimum. As before, I consider a small change $d\tilde{T}$ on tax brackets $i,i+1...,I$. The change $d\tilde{T}$ causes a change in the density of taxpayers in all brackets affected by the tax reform that is equal to $d\tilde{T}(-h_i\frac{\xi_i}{c_i}-h_{i+1}\frac{\xi_{i+1}}{c_{i+1}}-...-h_I\frac{\xi_I}{c_I})$. Each migration response to taxation induces a fiscal lost equal to individual's overall tax liability. When the absolute number of taxpayers is changed by the small tax deviation, there is an additional effect of migration on the government tax revenue, through the change in the number of transfers' benificiaries, and each individual emigration yields a fiscal gain equal to the universal demogrant $T_0=-T(0)$. It follows that the net effect of migration responses to taxation of $d\tilde{T}$ is equal to $-d\tilde{T}(h_i\frac{\xi_i}{c_i}(\tilde{T}_i-T_0)+h_{i+1}\frac{\xi_{i+1}}{c_{i+1}}(\tilde{T_{i+1}}-T_0)+...+h_I\frac{\xi_I}{c_I}(\tilde{T}_I-T_0))$. The overal effect of $d\tilde{T}$ on the universal demogrant is therefore $d\tilde{T}(h_i+h_{i+1}+...+h_I-\frac{T_i-T_{i-1}}{c_i-c_{i-1}}h_i\eta_i-\xi_i\frac{\tilde{T}_i-T_0}{c_i}h_i-\xi_{i+1}\frac{\tilde{T_{i+1}}-T_0}{c_{i+1}}h_{i+1}-...-\xi_I\frac{\tilde{T}_I-T_0}{c_I}h_I)$. The mobility response of individuals in each bracket-i is weighted by $b_i=(\tilde{T}_i-T_0)/c_i=T_i/c_i$, meaning that the trade off between the revenue and the transfer channel of tax driven migration responses emphasized in the linear case is captured by the migration wedge b_i in the non linear case.

3.5 Welfare Weights

I finally turn to the derivation of the optimal non linear tax tax and transfers schedule relying on the more general concept of generealized social marginal welfare weights. These formulas will be used in the numerical simulations, as they allow to capture the effects of government's tastes for redistribution on the optimal tax rates, and ultimately on the welfare effects of tax competition. As I will show later, the formulas of the optimal non linear tax rate with discrete earnings and welfare weights allow to emphasize simply the effects of tax-driven mobility on redistribution. I use the concept of generalized marginal social welfare weights, to be fully consistent with the approach presented for the linear framework. As before, the government attributes a weight g_i to type-i individuals' consumption, and the optimal tax schedule is such that any small tax deviation is welfare neutral.

Proposition 5. *Optimal Non-Linear Marginal Tax Rates in Federal Union:*

$$\tau_i^f = \frac{h_i(1-\bar{g}_i)+h_{i+1}(1-\bar{g}_{i+1})+...+h_I(1-\bar{g}_I)}{h_i(1-\bar{g}_i)+h_{i+1}(1-\bar{g}_{i+1})+...+h_I(1-\bar{g}_I)+h_i a_i e_i} \tag{18}$$

Proof. With $\bar{g}_i = g_i/(\sum_{m=0}^{I} h_m g_m \times N)$. The proof is derived below. □

Proposition 6. *Optimal Non-Linear Marginal Tax Rates in Tax Competition:*

$$\tau_i^c = \frac{h_i(1-b_i\xi_i-\bar{g}_i)+h_{i+1}(1-b_{i+1}\xi_{i+1}-\bar{g}_{i+1})+...+h_I(1-b_I\xi_I-\bar{g}_I)}{h_i(1-b_i\xi_i-\bar{g}_i)+h_{i+1}(1-b_{i+1}\xi_{i+1}-\bar{g}_{i+1})+...+h_I(1-b_I\xi_I-\bar{g}_I)+h_i a_i e_i} \tag{19}$$

Proof. With $\bar{g}_i = g_i/(\sum_{m=0}^{I} h_m g_m \times N)$. The proof is derived below. □

As in the linear case, I derive the optimal non linear tax rate using the small perturbation method with generalized social marginal weights, where the optimal tax schedule is such that no welfare gain can be acchieved through a small reform.[6] Consider again a small tax deviation dT on tax brackets $i, i+1..., I$. The deviation causes a mechanical increase in tax revenue $dT(h_i + h_{i+1} + ... + h_I)$. In addition to the mechanical change in revenue collected due to the tax reform, the small tax deviation creates behavioural responses to taxation. In the federal union, there is no migration responses to taxation, and the only behavioural response to taxation is through labour supply responses to the tax reform. The tax reform dT modifies transition to bracket $i-1$ to bracket i, and the density of taxpayers in the bracket i is changed by the amount $dh_i = -h_i\eta_i dT/(c_i - c_{i-1})$ at a net fiscal cost $(T_i - T_{i-1})$. The total effect of the reform on

[6] See Saez and Stantcheva [2016] for a discussion on the local derivation of the optimuum with generalized social marginal weights.

the government tax revenue is therefore $dT(h_i + h_{i+1} + ... + h_I - \frac{T_i - T_{i-1}}{c_i - c_{i-1}} h_i \eta_i)$. What is the effect on individuals welfare of the reform in the federal union? By definition of the general welfare weights, any increase in type-i individuals' consumption has a value g_i for the government. From bracket i to I, individuals have to pay additional taxes and the effect on their welfare is $-dT(h_i g_i + h_{i+1} g_{i+1} + ... + h_I g_I)$.The tax reform also increases the amount of transfer received by everyone such that the effect of the tax reform on total welfare through the change in tax revenue collected is given by $-\frac{dT}{N}(\sum_{m=0}^{I} h_m g_m (h_i + h_{i+1} + ... + h_I - \frac{T_i - T_{i-1}}{c_i - c_{i-1}} h_i \eta_i))$. At the optimum, the welfare effects sum to zero and the optimal marginal tax rate in the federal union is given by $\frac{\tau_i^f}{1-\tau_i^f} = \frac{h_i(1-\bar{g_i}) + h_{i+1}(1 - \bar{g_{i+1}} h_{i+1}) ... + h_I(1-\bar{g_I})}{h_i a_i e_i}$ with $\bar{g_i} = g_i / (\sum_{m=0}^{I} h_m g_m \times N$ the normalized welfare weight, and where the population N can be normalized to one without loss in generality. Note that this formula is a discretization of the optimal non linear tax formula provided in Saez and Stantcheva [2016].[7]

I now turn to the evaluation of the optimal non linear tax rate in the case where individuals respond to taxation through migration. As before, the tax reform dT modifies transition to bracket $i-1$ to bracket i, and the density of taxpayers in the bracket i is changed by the amount $dh_i = -h_i \eta_i dT/(c_i - c_{i-1})$ at a net fiscal cost $(T_i - T_{i-1})$. Second, the change dT causes a change in the number of taxpayers in all brackets affected by the tax reform, as migration decisions are driven by overall tax liabilities. This change in the mass of taxpayers is equal to $dT(-h_i \frac{\xi_i}{c_i} - h_{i+1} \frac{\xi_{i+1}}{c_{i+1}} - ... - h_I \frac{\xi_I}{c_I})$. Each migration response induces a net fiscal cost of T_i for the government and it follows that the net effect of migration responses to the reform on the government revenue is $-dT(h_i \frac{\xi_i}{c_i} T_i + h_{i+1} \frac{\xi_{i+1}}{c_{i+1}} T_{i+1} + ... + h_I \frac{\xi_I}{c_I} T_I)$. The overal effect of dT on the government tax revenue is $dT(h_i + h_{i+1} + ... + h_I - \frac{T_i - T_{i-1}}{c_i - c_{i-1}} h_i \eta_i - \xi_i \frac{T_i}{c_i} h_i - \xi_{i+1} \frac{T_{i+1}}{c_{i+1}} h_{i+1} - ... - \xi_I \frac{T_I}{c_I} h_I)$. This increase in tax revenue is rebated lump-sum such that the small reform is budget neutral. What are the effects on welfare of the tax reform dT? The increase in taxes for individuals in brackets $i, i+1, ..., I$ generates a change in the universal demogrant for all individuals in the economy,

[7] The formula can be rewritten as $\frac{\tau_i^f}{1-\tau_i^f} = \frac{\sum_{m>i} h_m - \bar{G}_i}{h_i a_i e_i}$ where $\sum_{m>i} h_m$ denotes the mass of taxpayers with income above the income level where the small tax reform applies and $\bar{G}_i = \sum_{m>i} h_m g_m / (\sum_{m=0}^{I} h_m g_m \times N)$ is the discrete equivalent of the average welfare weight parameter in t$\bar{G}(y) = \int_{\{i:y_i>y\}} g_i d_i / (\int_i g_i d_i)$ with population normalized to one.

and this welfare effect can be written $\frac{-dT}{N}(\sum_{m=0}^{I} h_m g_m(h_i + h_{i+1} + ... + h_I - \frac{T_i - T_{i-1}}{c_i - c_{i-1}} h_i \eta_i - \xi_i \frac{T_i}{c_i} h_i - \xi_{i+1} \frac{T_{i+1}}{c_{i+1}} h_{i+1} - ... - \xi_I \frac{T_I}{c_I} h_I)$. The welfare effect for individuals who have to pay the increase in taxes that is equal to $-dT(h_i g_i + h_{i+1} g_{i+1} + ... + h_I g_I)$. Summing the welfare effects to zero, we obtain the formula for the optimal non linear tax rate in tax competition presented in Equation (19).

Comparing Equation (19) with Equation (18) allows to see how mobility responses to taxation affects the implicit weights given by the government to mobile individuals. Because of b_i, the implicit welfare weight given to mobile individuals is *lowered* for individuals below the break-point, and *increased* for individuals above. The optimal tax formulas in the non-linear discrete framework therefore allow to emphasize how tax competition modifies implicitly the redistributive preferences on the government, and how these *mobility-adjusted* welfare weights may have different values, but also different distribution, because of the transfer channel of tax-driven mobility discussed in the previous section.

3.6 Effect of Tax Competition on Individuals' Welfare

I then turn to the numerical simulations of the optimal non linear marginal tax rates schedules. The numerical simulations for the non linear tax schedule are more complex, as they require to find the non linear tax schedule and the distribution of earnings that simultaneously satisfy the government first order condition. In order to compute the welfare effects of tax competition in the non linear tax schedule, I first conduct numerical simulations in order to find the optimal tax and transfer schedules in the federal and competing unions. For this purpose, I use a fixed-point algorithm such that the optimal tax formulas in Equation (19) and Equation (18) and the optimal conditions of individuals summarized by the behavioural elasticities in Equation (13) and Equation (12) are simultaneously satisfied. Then, I compute individuals' welfare taking into account the change in the taxes and transfer schedules and the changes in labour supply implied by the changes in taxes. I describe these two steps in details below.

I use a discrete grid of earnings with eight tax brackets based on the same empirical French earnings distribution used in the linear framework. I define h_i as the number of individuals whose earnings fall in the range $[y_i - (y_i - y_{i-1})/2; y_i + (y_{i+1} - y_i)/2]$. The resulting discretized distribution of earnings used for the numerical simulations is presented in Table B.II. In the discrete non linear model, the earnings grid and income levels are fixed, while intensive labour supply responses are loaded in the endogeneously determined population functions h_i. The functional forms chosen for the population functions h_i need to be consistent with the structure of behavioural elasticities defined in Equation (13) and Equation (12) and should coincide with empirical populations h_i^0 when the tax schedule is equal to the actual tax schedule. As in the linear framework, with symmetric countries, there is no tax-driven migration in equilibrum and the effect of migration on taxpayers' population and densities can be ignored. From Equation (13) it is possible to write:

$$h_i = h_i^0 \left(\frac{c_i - c_{i-1}}{c_i^0 - c_{i-1}^0} \right)^{a_i e_i} \tag{20}$$

Where $(c_{i-1}^0, c_i^0 ..., c_I^0)$ are the actual after tax schedules. The after-tax schedule used for the simulation is a very simple approximation of the real current after tax schedule, with a linear tax rate of 50 percent and a constant transfer of 5,000 euros. However, the results of the numerical simulations show very little sensitivity to the initial tax schedule used to solve the model. Using the functional form for h_i and the exogeneously chosen g_i, I find the tax and transfer schedules such that the optimal conditions of the government summarized in Proposition 5 and Proposition 6 and the behavioural responses summarized by Equation (20) are simultaneously satisfied.

With these optimal non-linear tax schedules at hand, I turn to the welfare analysis. As in the linear case, individuals' welfare is computed using the functional form described by Equation (8). The change in welfare from the federal union to the competition union is caused by the change in the optimal tax and transfer schedule and the change in labour supply that is loaded in the change of the endogeneous mass of tax payers h_i. The results of the numerical simulations are presented in Table 2 and Figure 2, while the shape of the optimal tax schedule is displayed in Figure 3.

4 Conclusion

This paper quantifies the welfare effects of tax competition. The results show that individuals in the bottom fifty percent of the income distribution always lose from tax competition, and would always be better off in a federal union. Their loss in welfare ranges on average from -10 to -20 percent, depending on the redistributive tastes of the government and the strenght of mobility responses to taxation. By contrast, higher income earners benefit from tax competition, as taxes are lowered by mobility responses to taxation when countries engage in tax competition.

These welfare estimates are based on three restrictive assumptions that imply that they are a lower bound for the real welfare effects of tax competition. The first restrictive assumption is that redistribution is viewed as not productive, as the government only reallocates consumption across individuals. In the case where public spending would generate externalities, say through investment in education or health, the welfare effects of tax competition may be increased. The second assumption is that the analysis is performed in a perfectly symmetric union, where there is no migration in equilibrium because the neighbouring countries mimic their tax policies. The welfare analysis therefore quantifies the effects of migration threat on individuals welfare, rather than the fully specified effects of migration. In the case of the migration equilibrium with asymmetric countries, earnings and transfers will be changed in equilibrium. The last assumption is that wages are assumed exogeneously fixed. When tax-driven mobility changes pre-tax earnings, individuals' welfare will be affected by tax-competition through the effects of mobility on pre-tax earnings. These three assumptions will be relaxed to complete this ongoing work.

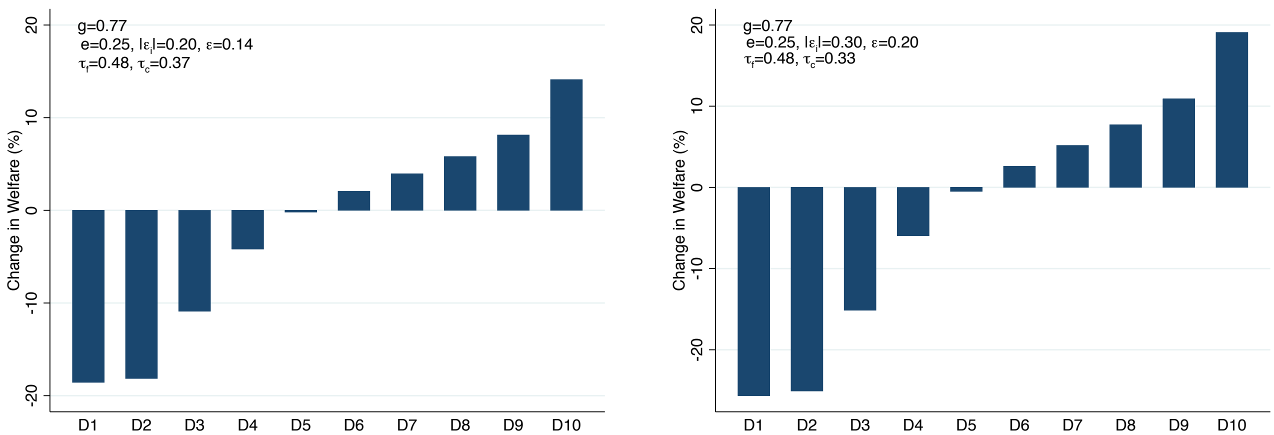
Table 1: **Effects of Tax Competition on Optimal Taxes and Welfare With a Linear Tax Schedule**

	Scenario 1 Elasticities e=0.25 $\bar{\varepsilon}$=0.07 $\lvert\varepsilon_i\rvert$=0.1		Scenario 2 Elasticities e=0.25 $\bar{\varepsilon}$=0.14 $\lvert\varepsilon_i\rvert$=0.2		Scenario 3 Elasticities e=0.25 $\bar{\varepsilon}$=0.20 $\lvert\varepsilon_i\rvert$=0.3		Scenario 4 Elasticities e=0.25 $\bar{\varepsilon}$=0.27 $\lvert\varepsilon_i\rvert$=0.4	
I- Optimal Linear Tax Rates	Federal	Competition	Federal	Competition	Federal	Competition	Federal	Competition
Rawlsian	0.73	0.68	0.73	0.64	0.73	0.60	0.73	0.57
Highly Redistributive	0.65	0.59	0.65	0.54	0.65	0.50	0.65	0.47
Mod. Redistributive	0.48	0.42	0.48	0.37	0.48	0.33	0.48	0.30
II- Welfare effect of Tax Competition (%)	Bottom 10	Bottom 50	Bottom 10	Bottom 50	Bottom 10	Bottom 50	Bottom 10	Bottom 50
Rawlsian	-2.7	-0.7	-6.0	-1.9	-10.3	-4.0	-12.8	-5.2
Highly Redistributive	-5.4	-2.5	-10.7	-5.3	-15.7	-7.9	-20.1	-10.5
Mod. Redistributive	-10.2	-5.7	-18.6	-10.5	-25.7	-14.5	-31.7	-18.0

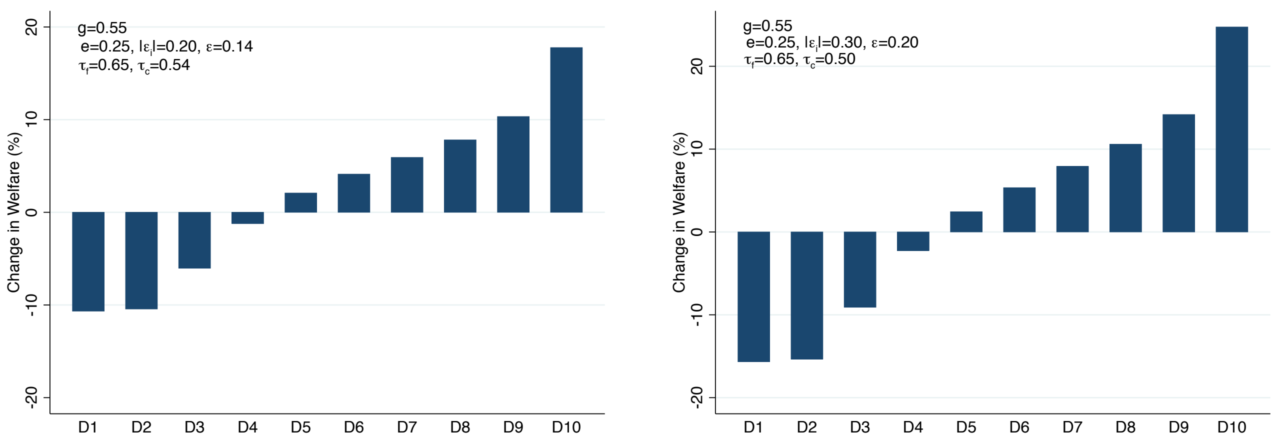
Notes: This Table summarizes the effects of tax competition on optimal tax rates and welfare. The optimal linear tax rates are computed following the formulas presented with more details in the text and presented in Proposition x and Proposition x. The elasticity e_i is the elasticity of type-i individuals gross earnings y_i with respect to the net-of-tax rate $1-\tau$. The elasticity ε_i is the elasticity of the number of type-i residents N_i with respect to the net-of-tax rate $1-\tau$. As described with more details in the text, ε_i is negative for all individuals who have an income level that is lower than the average income in the economy (break-even point). For the calibrations presented in the Table above, the migration responses to taxation are assumed to be constant across all earnings levels, that is to say of similar absolute value, meaning that all individuals in the population have the same migration response to an increase of their consumption through a change in taxes. The parameter e is the income weighted average labour supply elasticity $\sum_i((N_i y_i)/Y) \times e_i$ and the parameter $\bar{\varepsilon}$ is the combination of the income weighted and population weighted average mobility elasticity $\bar{\varepsilon} = \sum_i((N_i y_i)/Y) \times \varepsilon_i - \sum_i(N_i/N) \times \varepsilon_i$. The average welfare weight $\bar{g}$ captures the redistributive preferences of the government. The moderately redistributive government values the welfare of individuals in the bottom fifty percent two times more than the welfare of individuals in the other deciles with a corresponding $\bar{g} = 0.77$. The highly redistributive government values the welfare of individuals in the bottom fifty percent five times more than the welfare of other deciles with a corresponding $\bar{g} = 0.55$. The Ralwsian government only values the welfare of individuals in the bottom fifty percent. The welfare of each individual is computed using the utility specification $u_i = (1-\tau)y_i + T_0 - 1/(1+1/e_i) \times l_i^{1+1/e}$. Pre-tax earnings are endogeneously determined and follow the first order condition of the individual $y_i = w_i^{1+e}(1-\tau)^e$ using an exogeneous distribution of skills for w_i calibrated using the current distribution of French labour earnings combined with a current linear tax rate of 50 percent, displayed in Table B.II. The welfare effect of tax competition is the variation in percentage of individuals' welfare from a federal union to a competition union. A negative welfare variation means that individuals would be better off in a federal union.

Figure 1: **Distribution of Welfare Gains and Losses from Tax Competition with a Linear Tax Schedule**

Panel A. Moderately Redistributive Government



Panel B. Highly Redistributive Government



Notes: This graph shows the distribution of the welfare effects of tax competition across labour earnings' deciles. The welfare effect of tax competition is the variation in percentage of individuals' welfare from a federal union to a competition union. A negative welfare variation means that individuals would be better off in a federal union. The moderatively redistributive government values the welfare of individuals in the bottom fifty percent two times more than individuals in higher income deciles. The highly redistributive government values the welfare of individuals in the bottom fifty percent five times more than individuals in the higher deciles. The tax system consists in a linear τ paid on income and a universal demogrant redistributed to everyone. The parameter ε_i is the elasticity of migration with respect to the net-of-tax rate, while $\bar{\varepsilon}$ is a combination of income-weighted and population-weighted average migration elasticity. See the note below Table 1 for more details on the computation of the optimal tax rates and individuals' welfare.

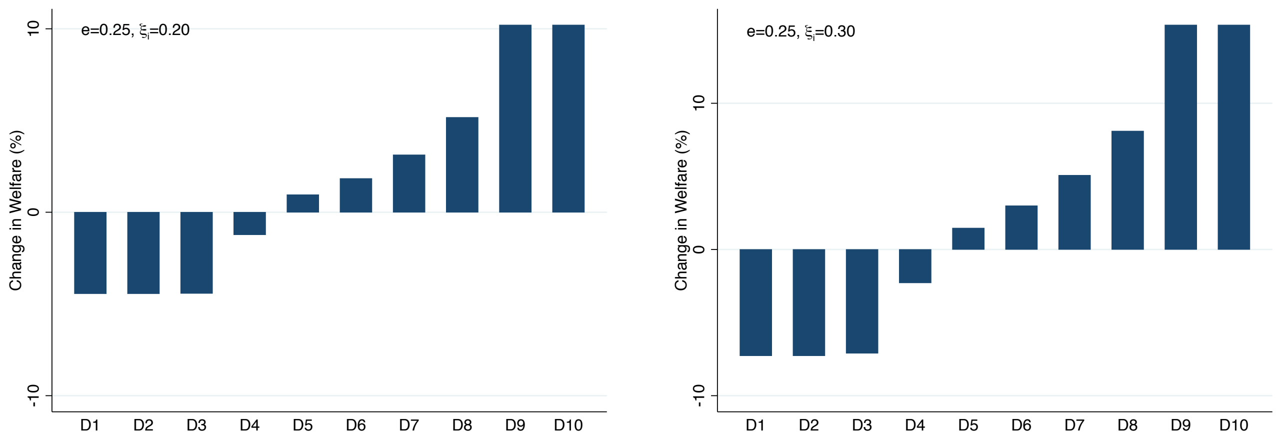
Table 2: **Effects of Tax Competition on Optimal Taxes and Welfare With a Non Linear Tax Schedule**

	Scenario 1 Elasticities e=0.25 ξ=0.1		Scenario 2 Elasticities e=0.25 ξ=0.2		Scenario 3 Elasticities e=0.25 ξ=0.3		Scenario 4 Elasticities e=0.25 ξ=0.4	
I- Average Marginal Tax Rates *(Marginal Tax Rate in the Top Bracket in Parentheses)*	Federal	Competition	Federal	Competition	Federal	Competition	Federal	Competition
Rawlsian	.69	.68	.69	.67	.69	.66	.69	.64
	(.64)	*(.63)*	*(.64)*	*(.61)*	*(.64)*	*(.59)*	*(.64)*	*(.55)*
Highly Redistributive	.62	.60	.62	.58	.62	.56	.62	.53
	(.57)	*(.54)*	*(.57)*	*(.51)*	*(.57)*	*(.48)*	*(.57)*	*(.45)*
Mod. Redistributive	.43	.41	.43	.39	.43	.37	.43	.35
	(.37)	*(.35)*	*(.37)*	*(.31)*	*(.37)*	*(.29)*	*(.37)*	*(.26)*
II- Welfare effect of Tax Competition (%)	Bottom 10	Bottom 50	Bottom 10	Bottom 50	Bottom 10	Bottom 50	Bottom 10	Bottom 50
Rawlsian	-.7	-.3	-.8	-.4	-2.9	-2.4	-4.8	-4.3
Highly Redistributive	-2.7	-2.1	-4.5	-3.4	-7.3	-5.9	-11.7	-9.5
Mod. Redistributive	-4.5	-3.4	-8.3	-5.0	-12.8	-9.3	-14.5	-11.5

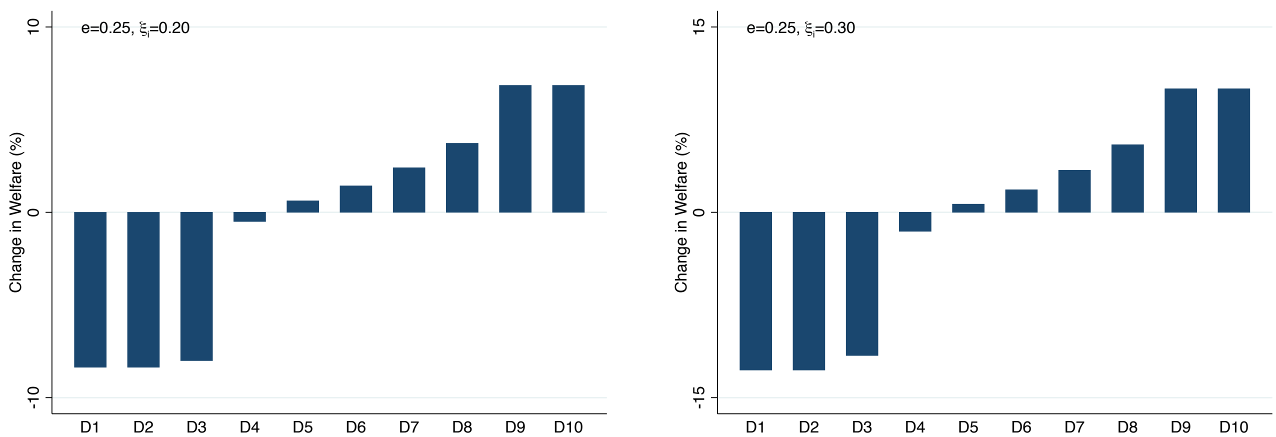
Notes: This Table summarizes the effects of tax competition on optimal tax rates and welfare. The optimal non linear tax rates are computed following the formulas presented with more details in the text and presented in Proposition 6 and Proposition 7. The numerical simulations use a discrete grid of earnings with eight income tax brackets taken from the empirical distribution of labour earnings in France and displayed in Table B.II. The elasticity of migration with respect to taxation ξ_i and the labour supply elasticity e_i are taken as constant across individuals and denoted ξ and e. The average marginal tax rates reports the average marginal tax rates across income tax brackets weighted by income in each bracket. The optimal marginal tax rates in the top income tax bracket are also reported in parentheses. The welfare is computed following Equation (8) and the endogeneous densities are determined following the functional form detailed in Equation (20). The moderately redistributive government corresponds to a government that values the welfare of individuals in the bottom fifty percent two times more than the welfare of individuals in higher earnings' deciles. The highly redistributive government values the welfare of individuals in the bottom fifty percent five times more than the welfare of individuals in higher income deciles. The welfare effect of tax competition is the variation in percentage of individuals' welfare from a federal union to a competition union. A negative welfare variation means that individuals would be better off in a federal union.

Figure 2: **Distribution of Welfare Gains and Losses from Tax Competition with a Non-Linear Tax Schedule**

Panel A. Highly Redistributive Goverment



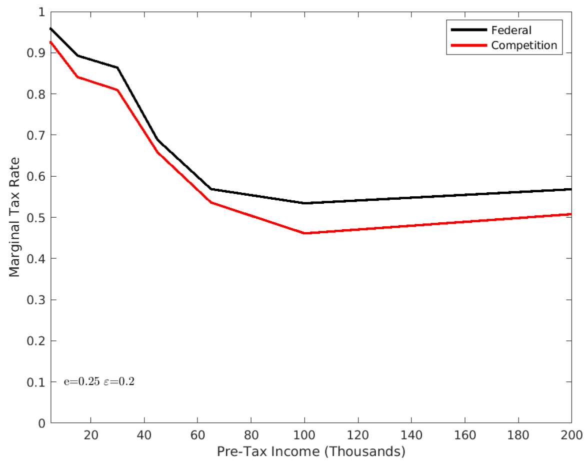
Panel B. Moderately Redistributive Government

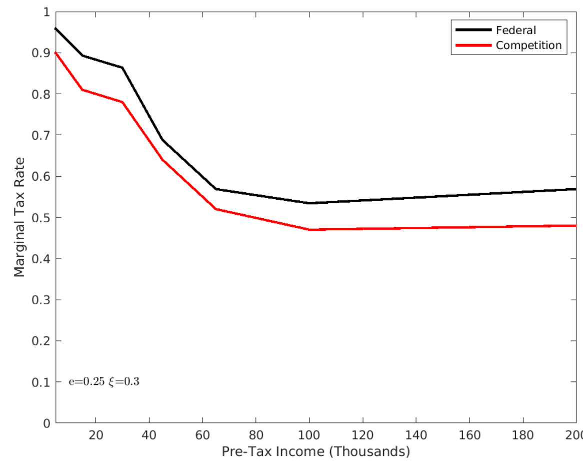


Notes: This graph shows the distribution of the welfare effects of tax competition across labour earnings' deciles. The welfare effect of tax competition is the variation in percentage of individuals' welfare from a federal union to a competition union. A negative welfare variation means that individuals would be better off in a federal union. The moderatively redistributive government values the welfare of individuals in the bottom fifty percent two times more than individuals in higher income deciles. The highly redistributive government values the welfare of individuals in the bottom fifty percent five times more than individuals in the higher deciles. The tax system is non linear. The parameter ξ_i is the elasticity of migration with respect to the disposable income, and is taken as constant across earnings' deciles. See the note below Table 2 for more details on the computation of the optimal tax rates and individuals' welfare.

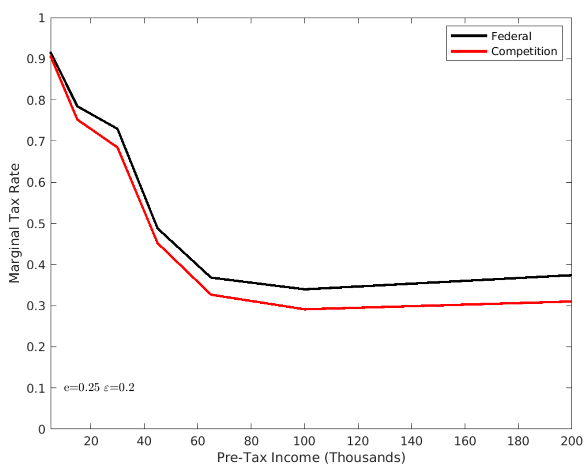
Figure 3: **Optimal Non-Linear Tax Schedules**

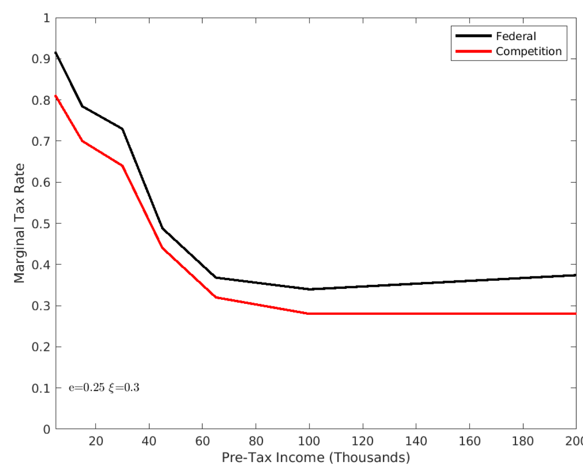
Panel A. Highly Redistributive Goverment





Panel B. Moderately Redistributive Government





Notes: This Figure shows the optimal marginal tax rates schedule after the numerical simulations of Proposition 5 and Proposition 6.

References

David R. Agrawal and Dirk Foremny. Relocation of the rich: Migration in response to top tax rate changes from spanish reforms. *Review of Economics and Statistics*, 2018.

Ufuk Akcigit, Salomé Baslandze, and Stefanie Stantcheva. Taxation and the international mobility of inventors. *American Economic Review*, 106(10):2930–81, 2016.

Karen Smith Conway and Andrew Houtenville. Elderly migration and state fiscal policy: Evidence from the 1990 census migration flows. *National Tax Journal*, pages 103–123, 2001.

Karen Smith Conway and Jonathan C. Rork. State "death" taxes and elderly migration—the chicken or the egg? *National Tax Journal*, pages 97–128, 2006.

Henrik Kleven, Camille Landais, Mathilde Muñoz, and Stefanie Stantcheva. Taxation and migration: Evidence and policy implications. *prepared for the Journal of Economic Perspectives*, 2019.

Henrik J. 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, 2013.

Henrik J. Kleven, Camille Landais, Emmanuel Saez, and Esben A. Schultz. Migration and wage effects of taxing top earners: Evidence from the foreigners' tax scheme in denmark. *Quarterly Journal of Economics*, 129(1):333–378, 2014.

Etienne Lehmann, Laurent Simula, and Alain Trannoy. Tax me if you can! optimal nonlinear income tax between competing governments. *Quarterly Journal of Economics*, 129(4):1995–2030, 2014.

Isabel Martinez. Beggar-thy-neighbour tax cuts: Mobility after a local income and wealth tax reform in switzerland. 2017. LISER Working Paper Series 2017-08.

James A. Mirrlees. An exploration in the theory of optimum income taxation. *Review of Economic Studies*, 38(2):175–208, 1971.

James A. Mirrlees. Migration and optimal income taxes. *Journal of Public Economics*, 18(3): 319–41, 1982.

Enrico Moretti and Daniel J. Wilson. The effect of state taxes on the geographical location of top earners: Evidence from star scientists. *American Economic Review*, 107(7):1858–1903, 2017.

Mathilde Muñoz. Do european top earners react to labour taxation through migration? 2019. Working Paper.

Thomas Piketty. La redistribution fiscale face au chômage. *Revue Française d'Economie*, 1997.

Thomas Piketty and Emmanuel Saez. Optimal labor income taxation. In A. Auerbach, R. Chetty, M. Feldstein, and E. Saez, editors, *Handbook of Public Economics*, volume 5, pages 391–474. Elsevier, 2013.

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68, 2001.

Emmanuel Saez. Optimal income transfer programs: intensive versus extensive labor supply responses. *The Quarterly Journal of Economics*, 117(3):1039–1073, 2002.

Emmanuel Saez and Stefanie Stantcheva. Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, 106(1):24–45, 2016.

A Proofs

A.1 Revenue-Maximizing Linear Tax Rate

Federal Government The federal government maximises the tax revenue function, that does not depend on the number of taxpayers because of the absence of tax-driven migration in the federal union. For notation purposes, I present the problem such that there are N_i individuals characterised by preferences $u^i(c,y)$, and Y denotes the sum of earnings' function over individuals with various preferences $Y = \sum_i N_i y_i$.

The first order condition with respect to goverment rax revenue $R = \sum_i N_i y_i (1-\tau)\tau = Y(1-\tau)\tau$ is given by $(d\tau/d\tau)Y - (dY/d(1-\tau))\tau = 0$. Using the definition of the labour supply elasticity,we can show that:

$$\tau^f = \frac{1}{1+e}$$

Where e is the income weighted elasiticities such that $e = \sum \frac{N_i y_i}{Y} e_i$.

Competing Government In the presence of tax competition, the number of taxpayers becomes a function of the net-of-tax rate determined in tax competition. As a result, the government tax revenue can be written $R = \sum_i N_i(1-\tau) y_i(1-\tau)\tau$. The first order condition with respect to the tax rate is $\sum_i [(d\tau/d\tau) y_i N_i - (dy_i/d(1-\tau))\tau N_i - (dN_i/d(1-\tau))\tau y_i] = 0$. Using the definition of ε_i and e_i, we obtain:

$$\tau^f = \frac{1}{1+e+\varepsilon}$$

Where is e and ε are the income weighted elasiticities such that $e = \sum_i \frac{y_i N_i}{Y} e_i$ and $\varepsilon = \sum_i \frac{y_i N_i}{Y} \varepsilon_i$. The optimal tax can be easily retrieved by studying a small deviation in the tax schedule τ. Consider an infra-marginal change in the linear tax schedule $d\tau$. The small tax deviations induces a change in the government tax revenue equal to $d\tau Y$, due to a mechanical increase in tax revenue. As pre-tax earnings are endogeneously determined by a labour-leisure trade-off, the reform causes an aggregated change in earnings $-e\frac{\tau}{1-\tau} Y d\tau$. In the presence of tax competition, individuals have an extensive margin of response to the tax change through migration. Individuals react to

$d\tau$ through an additional migration effect $-\varepsilon\frac{\tau}{1-\tau}Y d\tau$, that captures mobility response to the net effect of the reform on their post-tax earnings. The total effect on tax revenue is therefore given by $dR = (1 - e\frac{\tau}{1-\tau} - \varepsilon\frac{\tau}{1-\tau})Y d\tau$ in the competing union, and $dR = (1 - e\frac{\tau}{1-\tau})Y d\tau$ in the federal union. Summing behavioural and mechanical effects to zero yields the inverse tax rate formula for the Laffer rate that maximizes tax revenue.

A.2 Transfer Maximizing Rate

Let's now consider the case where the government wants to maximize the amount of transfer to the poorest individuals (that is to say the government is Rawlsian). In the absence of tax competition, the population can always be normalized to one without loss of generality, and the revenue-maximizing rate corresponds to the optimal linear rate chosen by the Rawlsian government. In the presence of tax competition, individuals are able to respond to taxation trough migration, and the absolute number of taxpayers may be changed by these migration responses. Therefore, the amount that can be redistributed to individuals can be indirectly affected by the number of individuals in the country through the tax-driven migration channel. As a result, when the absolute number of taxpayers in changed by a change in the linear tax rate, the Ralwsian linear rate is no longer equivalent to the revenue-maximizing rate. The optimal linear rate of the Rawlsian government maximizes $T_0 = \frac{1}{N}\sum_i \tau y_i N_i$. Denoting $R = \sum_i \tau y_i N_i$, the first order condition with respect to τ is given by $\frac{dR \times N - dN \times R}{N^2}$. Formally, the FOC is:

$$\frac{1}{N}(\sum_i y_i N_i - \sum_i \frac{\tau}{1-\tau} e_i y_i N_i - \sum_i \frac{\tau}{1-\tau}\varepsilon_i y_i N_i) + \frac{1}{N^2}(\sum_i \frac{\tau}{1-\tau}\varepsilon_i N_i \sum_i y_i N_i) = 0 \qquad (21)$$

The equation can be rewritten as

$$1 - \sum_i \frac{\tau}{1-\tau}\frac{e_i N_i y_i}{Y} - \sum_i \frac{\tau}{1-\tau}\frac{\varepsilon_i y_i N_i}{Y}) + \frac{1}{N}\sum \frac{\tau}{1-\tau}\varepsilon_i N_i = 0$$

And it follows that the linear tax rate that maximizes the amount of transfer is given by

$$\frac{\tau}{1-\tau} = \frac{1}{e + \bar{\varepsilon}}$$

Where $e = \sum \frac{N_i y_i}{Y} e_i$ and $\bar{\varepsilon} = \sum \frac{N_i y_i}{Y} \varepsilon_i - \sum_i \frac{N_i}{N} \varepsilon_i$.

A.3 Optimal Linear Tax Rate with Welfare Weights

In this section, I consider the case where the government maximizes the total welfare in a country. The welfare function is modeled as a sum of weighted utilities, where the welfare weights capture the social preferences of the government, and are exogeneously determined. In each country, there is a mass N_i of type-i individuals characterized by the same preferences and the same income y_i. In the federal symmetric union, the size of the population does not matter for the welfare maximization and can always be normalized to one without loss of generalities. Without tax-driven migration, the welfare maximized is always the welfare of individuals located in the country, as location choices are exogeneous to the coordinated tax policy. Any change in the tax rate will not affect the government welfare function through the number of individuals entering in this sum. In a free mobility union with competing countries, the maximization of total welfare for the optimal tax policy becomes less evident, as the competing government could aim to maximize the welfare of its nationals, initial residents, or may want to take into account the welfare of residents arriving after a change in the tax rate. The issues related to which individuals should be included in the social welfare functions of competing countries is normative, and beyond the scope of this paper. I discuss below two alternative welfare functions and their implications for the optimal linear tax rate set at the optimuum.

Formally, the government attributes a general welfare weight g_i to the utility of type i individuals in the economy such that the optimal linear tax rate maximizes $\sum_i N_i g_i u_i(c_i, y_i)$. Because of the envelop theorem, the government can ignore the effect of τ through y_i at the optimum. When countries compete, the density N_i is affected by tax policy. Therefore, the total welfare of the government may be affected through two channels by tax competition. The sum of weighted utilities may be changed by *(i)* the intensive change in welfare through the change in taxes paid and transfers received but also through *(ii)* the extensive change in the number of individuals that enter in the welfare maximization of the government due to migration responses to taxation. This can be simply observed by taking the first order condition of the government social welfare function with respect to $1 - \tau$:

$$-\sum_i N_i g_i \frac{\partial u_i(c_i)}{\partial(1-\tau)} - \sum_i g_i u_i(c_i) \frac{\partial N_i}{\partial(1-\tau)} \tag{22}$$

The first term captures the effect of tax-driven migration on the level of consumption in the country, because mobility responses to taxation affect (1) taxes collected by the government and (2) the number of individuals in the country who have to split the amount collected by the government. The second term captures the effect of mobility responses to taxation of the number of individuals entering in the welfare function of the government. In the case where the welfare maximization is endogeneous to the population changes, the government may have the incentives to increase the number of taxpayers to increase the total welfare in the economy. Said differently, the government could have the incentive to maximize the amount of individuals entering in the country in order to maximize the total sum of welfare, rather than maximizing the amount of welfare for a given population size.

In the main specification, I ignore the second term of Equation (22) in order to avoid considerations related to the size of the population that maximizes the weighted sum of utilities. Rather, I consider a government that maximizes the welfare of a given population N_i, that is endogeneously determined in equilibrium, but that is taken as given for the welfare aggregation. With this specification, the number of individuals leaving country A only affect the total welfare in country A through the effect of their migration on the level of tax revenue and transfer for individuals residing in country A. The fact that individuals leave country A and thus the total welfare of country A does not matter, the government only cares about individuals residing in country A. Similarly, the total welfare in country A cannot be increased by the entry of new residents would increase the total welfare in country A by increasing the number of individuals entering in the sum. In spirit, this approach would consist in a government that maximizes the welfare of non-movers, taking into account the effect of movers on non movers utility through the revenue and transfer effects. Using the quasi-linearity in consumption because of the absence of income effects, it is possible to write the first order condition with respect to the linear tax rate as:

$$\sum N_i g_i \left(-y_i + \frac{1}{N}\left(Y - \sum_i \frac{\tau}{1-\tau} e_i y_i N_i - \sum_i \frac{\tau}{1-\tau} \varepsilon_i y_i N_i\right) + \frac{Y}{N^2} \sum_i \frac{\tau}{1-\tau} N_i \varepsilon_i\right) = 0 \tag{23}$$

$$\sum_i N_i g_i y_i . \sum_i N_i = \sum_i N_i g_i (Y - \sum_i \frac{\tau}{1-\tau} e_i y_i N_i - \sum_i \frac{\tau}{1-\tau} \varepsilon_i y_i N_i + Y \sum_i \frac{\tau}{1-\tau} \frac{N_i}{N} \varepsilon_i$$

$$\frac{\sum_i N_i g_i y_i . \sum_i N_i}{(\sum_i N_i Y_i) . \sum_i N_i g_i} = 1 - \frac{\tau}{1-\tau} e - \frac{\tau}{1-\tau} \bar{\varepsilon} \tag{24}$$

Denoting $\bar{g} = \frac{\sum N_i g_i y_i . N}{Y . \sum N_i g_i}$, we obtain the optimal linear tax rate formula $\tau^c = \frac{1-\bar{g}}{1-\bar{g}+e+\bar{\varepsilon}}$ where $\bar{\varepsilon}$ is a combination of the income weighted migration elasticity $\varepsilon = \sum_i \frac{N_i y_i \varepsilon_i}{Y}$ and the population-weighted migration elasticity $\varepsilon_p = \sum_i \frac{N_i \varepsilon_i}{N}$. In the case where the absolute number of taxpayers is unchanged by tax-driven migration (only the composition of the population changes), the population-weighted elasiticity is zero, and the optimal linear tax rate only depends on the standard income-weighted average mobility parameter ε, similarly than for the revenue-maximizing government. Importantly, the terms $\sum N_i g_i y_i$ and $\sum N_i g_i$ of the average welfare weight $\bar{g}$ depends on the densities that are taken as exogeneous in the government for the welfare aggregation (for instance, densities of stayers).

Endogeneous Size of the Welfare Sum Let's now explore the case where the government maximizes the sum of weighted utilities taking into account the change in the composition of the population and thus the set of individuals entering in the sum of welfare. With the envelop theorem and quasi-linearity in consumption, this is equivalent to $\sum_i N_i g_i \left[(1-\tau) y_i + \frac{\tau \sum_i N_i y_i}{\sum_i N_i} \right]$. The first order condition of the government with respect to τ would be given by $\sum_i N_i g_i \frac{\partial u_i(c_i)}{\partial \tau} + \sum_i g_i u_i(c_i) \frac{N_i}{\partial \tau}$. With these social preferences, the endogeneous change of N_i caused by tax-driven mobility will affect the total welfare through its effect on taxes and transfer as captured by Equation (23), and an additional effect through the change in the number of individuals who compose the welfare sum in equilibrium. The first order condition is more precisely given by:

$$\sum_i N_i g_i \left(-y_i + \frac{1}{N} \left[Y - \sum_i \frac{\tau}{1-\tau} e_i y_i N_i - \sum_i \frac{\tau}{1-\tau} \varepsilon_i y_i N_i \right] + \frac{Y}{N^2} \sum_i \frac{\tau}{1-\tau} N_i \varepsilon_i \right) - \sum_i g_i u_i(c_i, y_i) \frac{\partial N_i}{d(1-\tau)} = 0 \tag{25}$$

As in the previous section, the first term of Equation (25) captures the effect of tax-driven migration on welfare through its effects on residents' consumption: *(i)* a change in taxes paid *(ii)* a change in the amount of transfers received because of change in tax liabilities (*revenue effect*) and change in absolute number of transfer beneficiaries (*transfer channel*). The formula is augmented by an additional term capturing the effect of tax-driven migration of the amount of individuals who enter in the total welfare of the country. The underlying intuition is that any change in the tax rate τ causes a change in the total welfare through the amount of individuals who leave the country and somehow *"take their welfare"* with them. This effect is of magnitude $\frac{\partial N_i}{d(1-\tau)}$ that captures the magnitude of migration responses to taxation, and has a welfare cost $g_i u_i$ as any type-i individual leaving the country decreases the sum of total welfare by its consumption weighted by its corresponding welfare weight. Note that for individuals below the break even point, N_i is an increasing function of the net-of-tax rate, and therefore an increase in τ increases the total welfare by including immigrants in the welfare sum. The FOC of the government can be rewritten:

$$\sum_i N_i g_i \left(-y_i + \frac{1}{N}\left[Y - \sum_i \frac{\tau}{1-\tau} e_i y_i N_i - \sum_i \frac{\tau}{1-\tau}\varepsilon_i y_i N_i\right] + \frac{Y}{N^2}\sum_i \frac{\tau}{1-\tau} N_i \varepsilon_i\right) - \sum_i g_i u_i(c_i, y_i)\frac{N_i}{1-\tau}\varepsilon_i = 0$$

$$\sum_i N_i g_i \left(-y_i + \frac{1}{N}\left[Y - \sum_i \frac{\tau}{1-\tau} e_i y_i N_i - \sum_i \frac{\tau}{1-\tau}\varepsilon_i y_i N_i\right] + \frac{Y}{N^2}\sum_i \frac{\tau}{1-\tau} N_i \varepsilon_i\right) - \sum_i g_i \frac{N_i}{1-\tau}\varepsilon_i \left[(1-\tau)y_i + \tau Y/N\right] = 0$$

$$\sum_i N_i g_i \left(-y_i + \frac{1}{N}\left[Y - \sum_i \frac{\tau}{1-\tau} e_i y_i N_i - \sum_i \frac{\tau}{1-\tau}\varepsilon_i y_i N_i\right] + \frac{Y}{N^2}\sum_i \frac{\tau}{1-\tau} N_i \varepsilon_i\right) - \sum_i g_i N_i \varepsilon_i y_i - \sum_i g_i N_i \frac{\tau}{1-\tau}\varepsilon_i \frac{Y}{N}) = 0$$

$$\sum_i N_i g_i y_i (1+\varepsilon_i) = \sum_i N_i g_i \left(\frac{1}{N}\left[Y - \sum_i \frac{\tau}{1-\tau} e_i y_i N_i - \sum_i \frac{\tau}{1-\tau}\varepsilon_i y_i N_i\right] + \frac{Y}{N^2}\sum_i \frac{\tau}{1-\tau} N_i \varepsilon_i\right) - \sum_i N_i g_i \frac{\tau}{1-\tau}\varepsilon_i \frac{Y}{N}$$

$$\frac{\sum_i N_i g_i y_i (1+\varepsilon_i) \times N}{\sum N_i y_i} = \sum_i N_i g_i \times \left(1 - \frac{\tau}{1-\tau} e - \frac{\tau}{1-\tau}\varepsilon + \sum_i \frac{\tau}{1-\tau}\frac{N_i}{N}\varepsilon_i\right) - \frac{\tau}{1-\tau}\sum_i N_i g_i \varepsilon_i$$

$$\frac{\sum_i N_i g_i y_i (1+\varepsilon_i) \times N}{\sum N_i y_i . \sum_i N_i g_i} = 1 - \frac{\tau}{1-\tau}\left(e - \varepsilon + \varepsilon_p - \sum_i \frac{N_i g_i \varepsilon_i}{\sum_i N_i g_i}\right) \tag{26}$$

How is the optimal linear tax rate changed if the size of the welfare sum, or said differently the population that is taken into account in the total welfare of one country, is changed by taxation? There is a first change through the average welfare weight parameter, that is now affected by the migration elasticity. The welfare weight of individuals is augmented (or lowered) by the strength of their mobility elasticity. The optimal tax rate is also directly affected by an additional term

$\varepsilon_w = \sum_i \frac{N_i g_i \varepsilon_i}{\sum_i N_i g_i}$ that is a *welfare-weighted average mobility elasticity*, capturing the effect of tax-driven mobility on total welfare through its effect on the number of individuals included in the welfare definition. For individuals with a negative mobility elasticity, this term is positive, meaning that the government has incentives to *increase* the linear tax rate in order to attract bottom earners and to capture their additional welfare. To summarize, when the government maximizes the total welfare in the country by also maximizing the amount of individuals included in the computation of the welfare, there are three main effects of tax-driven migration on the optimal linear tax rate. First, tax-driven migration changes the revenue collected in equilibrium through the *revenue channel* that is captured by the *income-weighted* parameter ε. Second, tax driven migration changes the amount that can be redistributed to everyone remaining in the country through the *transfer channel* that is captured by the *population weighted* parameter ε_p. These effects are the one affecting welfare through residents' consumption, as in Equation (24). Third, tax-driven migration changes the total welfare through the changed number of individuals included in the welfare aggregation in equilibrium. This *size channel* affects the optimal linear tax rate through the *welfare weighted* parameter ε_w that captures the amount of welfare that can be attracted or loss due to the absolute change in the number of individuals that enter in the government sum of weighted utilities.

A.4 Formal Derivation of the Non Linear Optimal Tax Rates

A.4.1 Intensive Model

The Ralwsian government maximizes the tax revenue $R = \sum_{i=0}^{J} T_i h_i$, given that h_i is a function of $(c_i - c_{i-1}, c_{i+1} - c_i)$. The first order condition is given by the system of equation:

$$h_i = \sum_{j=0}^{I} \frac{-dh_j}{dT_i} T_j = \sum_{j=0}^{I} \frac{dh_j}{dc_i} T_j$$

As individuals can only choose between adjacent occupation, it is easy to rewrite the first order condition such that:

$$h_i = T_{i-1} \frac{\partial h_{i-1}}{\partial (c_i - c_{i-1})} - T_{i+1} \frac{\partial h_{i+1}}{\partial (c_{i+1} - c_i)} + T_i \frac{\partial h_i}{\partial (c_i - c_{i-1})} - T_i \frac{\partial h_i}{\partial (c_{i+1} - c_i)} \tag{27}$$

Using $\partial h_{i+1} / \partial (c_{i+1} - c_i) = -\partial h_i / \partial (c_{i+1} - c_i)$, we obtain:

$$h_i = (T_i - T_{i-1})\frac{\partial h_i}{\partial(c_i - c_{i-1})} + (T_{i+1} - T_i)\frac{\partial h_i}{\partial(c_i - c_{i+1})}$$

Using Equation 27 for $i = i+1...I$ and the participation elasticity $\eta_i = \partial h_i / \partial(c_i - c_{i-1}) \times (c_i - c_{i-1}) \times h_i$ we obtain:

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{h_i + h_{i+1} + ... + h_I}{h_i \eta_i}$$

To express the optimal tax schedule as a function of the standard labor supply elasticity e_i, I follow Saez (2002) and use $(y_i - y_{i-1})\eta_i = e_i y_i$, that yields to $\eta_i = e_i y_i / (y_i - y_{i-1})$. Using τ_i the marginal tax rate on bracket i such that $\tau_i = (T_i - T_{i-1})/(Y_i - Y_{i-1})$, where $1 - \tau_i = c_i - c_{i-1}/Y_i - Y_{i-1}$, we obtain the formula for the optimal marginal tax rate on bracket i. As outlined by Saez (2002), in the absence of extensive margin responses to taxation, the optimal tax liabilities are always increasing with i, and negative marginal tax rates are therefore never optimal. As a result, the marginal tax rate in the first bracket is very high, and is maximal in the Rawlsian case with high redistributive taste.

A.4.2 Extensive Model

Let's consider now the case where individuals respond to taxation through migration. Conditional of being in the bracket i, individuals can choose to migrate from A to B if $u_i(c_i^B, y_i) \geq u_i(c_i^A, y_i)$. In that case, the fraction of individuals in a given tax bracket h_i is a function of the overall tax schedule in the bracket i T_i. Consider first the case where there are only extensive margin responses to taxation. In that case, the number of individuals in the tax bracket i is only a function of the overall tax liability T_i that determines migration decisions. The system of first order conditions follows:

$$h_i = \frac{\partial h_i}{\partial c_i} T_i \tag{28}$$

Making use of the migration elasticity formula:

$$\frac{T_i}{y_i - T_i} = \frac{1}{\xi_i} \tag{29}$$

A.4.3 Mixed Intensive and Extensive Model

I finally put together intensive and extensive responses to taxation. In this case, individuals in bracket i can choose adjacent occupations, but can also choose to locate abroad. The first order condition becomes:

$$h_i = (T_i - T_{i-1}) \frac{\partial h_i}{\partial (c_i - c_{i-1})} + (T_{i+1} - T_i) \frac{\partial h_i}{\partial (c_{i+1} - c_i)} + T_i \frac{\partial h_i(c_i)}{\partial (c_i)} \tag{30}$$

This relationship illuminates the effects of a tax reform at the extensive and intensive margins. The two first terms of Equation 30 capture the effect of a change in the marginal rate on transitions between brackets. The last term captures the extensive margin on individuals that decide to migrate after that their overall tax liabilitie T_i has been changed. It is easy to derive the optimal top marginal tax rate from Equation 30 :

$$\frac{T_I - T_{I-1}}{c_I - c_{I-1}} = \frac{1}{a_I e_I} \left(1 - \frac{T_I}{c_I} \xi_I \right)$$

That simplifies to the following formula, with $b_I = T(y_I)/(y_I - T(y_I))$ that captures the overall wedge of labour taxation at income level y_I and $a_I = y_I/(y_I - y_{I-1})$ the local discrete pareto parameter.

$$\frac{\tau_I}{1 - \tau_I} = \frac{1 - b_I \xi_I}{a_I e_I} \tag{31}$$

For any other income level, making use of the set of first order conditions, the optimal tax formula is given by:

$$\frac{\tau_i}{1 - \tau_i} = \frac{h_i(1 - b_i \xi_i) + h_{i+1}(1 - b_{i+1} \xi_{i+1}) + ... + h_I(1 - b_I \xi_I)}{a_i h_i e_i} \tag{32}$$

If we let one of the two elasticities e_i and ξ_i tend to zero in Equation 32, we retrieve the optimal formula of pure extensive and intensive models. The model mixing labour supply and migration responses is similar to the pure intensive model, with weights on income groups replaced by $b_i \xi_i$, implicitly attributing more welfare weights to individuals who are more likely to respond to taxation through migration.

A.5 Non Linear Discrete Tax Rates with Social Marginal Welfare Weights

I consider alternatively the formal derivation of the optimal tax formulas in the case where the the government maximizes a welfare function. Individual k in the bracket i derives an utility $u^k(c_i, i)$ that is a function if his after-tax income in the bracket i c_i and of its tax bracket choice $i = 0, ..., I$. For a given tax schedule, there is a population $h_i(c_0, ..., c_I)$ of individuals who choose to be in the bracket i. I denote $k(i)$ the tax bracket choice of individual k. The government chooses the tax schedule $(T_0, ..., T_I)$ such that it maximizes the total welfare:

$$SWF = \sum_k w_k u_k$$

$$\sum h_i T_i \geq R$$

Denoting p the multiplier of the government budget constraint, the first order condition yields

$$(1 - G_i)h_i = \sum_{m=0}^{I} \frac{\partial h_m}{\partial c_i} T_i \tag{33}$$

With $G_i = \frac{1}{ph_i} \sum_{k \in bracket} w_k G'(u^k) u^k_c(c_i, k(i))$. Similarly than before, I can derive the optimal tax schedules in the federal and competition union with these alternative social welfare weights G_i. With the assumption that intensive responses only occur between adjacent occupations, the first order condition can be rewritten in the federal union as

$$(1 - G_i)h_i = (T_i - T_{i-1}) \frac{\partial h_i}{\partial (c_i - c_{i-1})} + (T_{i+1} - T_i) \frac{\partial h_i}{\partial (c_{i+1} - c_i)} \tag{34}$$

In the competition union, there is an additional behavioural response to taxation, and the number of individuals in each bracket h_i is affected through migration responses to taxation such that:

$$(1 - G_i)h_i = (T_i - T_{i-1}) \frac{\partial h_i}{\partial (c_i - c_{i-1})} + (T_{i+1} - T_i) \frac{\partial h_i}{\partial (c_{i+1} - c_i)} + T_i \frac{\partial h_i(c_i)}{\partial (c_i)} \tag{35}$$

B Additional Tables and Figures

Table B.I: **Migration Elasticities of Top Ten Percent Employees**

Countries	Lower bound	Upper Bound
Belgium	.19	.27
Germany	.16	.24
France	.32	.45
Italy	.05	.07
Luxembourg	.26	.37
Poland	.12	.18
Portugal	.10	.15
Spain	.25	.34
Switzerland	.28	.41
United Kingdom	.52	.83

Notes: This Table summarizes the elasticity of the number of top ten percent employees with respect to the top net-of-tax rate estimated by Muñoz [2019] for the period 2009-2015 using individual-level data from the European Labour Force Survey. The empirical strategy exploits within-country variations in top marginal tax rates coming from differences in propensities to be treated by top marginal tax rates between individuals of different earnings levels. Lower bounds are computed using the 8th decile as a control group not affected by top marginal tax rates changes. Upper bounds are compited using the 5th decile as a control group not affected by top marginal tax rates changes.

Table B.II: **Empirical Earnings Distribution Calibration**

Income level (euros)	Density Weight	Cumulative Density
0	.2	.2
5,000	.05	.25
15,000	.11	.36
30,000	.14	.5
45,000	.23	.73
65,000	.17	.9
100,000	.08	.98
200,000	.02	1

Notes: This Table shows the discretized empirical earnings distribution used for the numerical simulations of the non-linear tax and transfer schedule. The data is based on the distribution of labour factor income for France provided by the World Inequality Database. I define the density at each average income level as the density of individuals whose earnings fall in the range $[y_i - (y_i - y_{i-1})/2; y_i + (y_{i+1} - y_i)/2]$.

Table B.III: **Tax-Driven Migration Only at the Top With Linear Tax Schedule**

	Scenario 1 Elasticities e=0.25 $\bar{\varepsilon}$=0.02 $\varepsilon_{D_1-D_9}$=0, $\varepsilon_{D_{10}}$=0.1		Scenario 2 Elasticities e=0.25 $\bar{\varepsilon}$=0.04 $\varepsilon_{D_1-D_9}$=0, $\varepsilon_{D_{10}}$=0.2		Scenario 3 Elasticities e=0.25 $\bar{\varepsilon}$=0.06 $\varepsilon_{D_1-D_9}$=0, $\varepsilon_{D_{10}}$=0.3		Scenario 4 Elasticities e=0.25 $\bar{\varepsilon}$=0.09 $\varepsilon_{D_1-D_9}$=0, $\varepsilon_{D_{10}}$=0.4	
I- Optimal Linear Tax Rates								
	Federal	Competition	Federal	Competition	Federal	Competition	Federal	Competition
Rawlsian	0.73	0.71	0.73	0.70	0.73	0.68	0.73	0.67
Highly Redistributive	0.65	0.63	0.65	0.61	0.65	0.59	0.65	0.57
Mod. Redistributive	0.48	0.45	0.44	0.37	0.42	0.33	0.48	0.40
II- Welfare effect of Tax Competition (%)								
	Bottom 10	Bottom 50	Bottom 10	Bottom 50	Bottom 10	Bottom 50	Bottom 10	Bottom 50
Rawlsian	-.7	-0.2	-1.7	-.3	-2.6	-.6	-3.6	-.9
Highly Redistributive	-1.7	-.7	-3.5	-1.6	-5.2	-2.4	-6.8	-3.24
Mod. Redistributive	-3.2	-1.9	-6.7	-3.7	-9.7	-5.4	-12.5	-7.0

Notes: This Table shows the welfare effects of tax competition when migration responses to taxation are concentrated at the top of the income distribution (top decile) but are zero everywhere else. This implies that $\bar{\varepsilon} = \varepsilon_{D10} - \varepsilon_{p,D10}$, as $\varepsilon_i = 0$ for every individuals who are not in the top decile. The methodology for welfare computation is described extensively in the note below Table 1.