Joint Taxation of Income and Wealth*

Mehmet Ayaz — LMU Munich

Dominik Sachs — University of St. Gallen

November 2021

Abstract

Empirically, income and wealth are positively correlated. A 'tagging' logic implies that taxing wealth more strongly for individuals with high income would be desirable. However, wealth and income are both endogenous variables and the simple tagging logic needs to be extended to account for this. If e.g., wealth taxes are increased only for individuals that earn more than \$100,000, this creates *two* distortions: not only the wealth accumulation margin gets distorted, but also the labor earnings margin for those that earn around \$100,000. We derive formulas for the excess burden and the welfare effects of such joints reforms and calibrate them to the U.S. Our first preliminary findings indicate that taxing wealth more progressively for individuals with high income is desirable despite the 'double distortion' it implies.

JEL classification: H21, H23

Keywords: Wealth Taxation, Capital Taxation, Income Taxation, Tax Reforms

1 Introduction

In this paper, we analyze the joint taxation of income and wealth. We show that conditioning the marginal tax rate of one variable on the other variable comes with tagging benefits if income and wealth are correlated with each other. However, an extra distortion arises due to the conditioning. We formalize this by studying joint tax reforms as in Golosov, Tsyvinski and Werquin (2014).

As an important benchmark, we consider the case where income and wealth are uncorrelated. In this case, there are no gains from tagging. Taxing wealth more strongly for high

*Contact: ayaz@econ.lmu.de and dominik.sachs@unisg.ch.

incomes does not bring any gains from tagging. However, a second distortion is created. In fact, for this benchmark case where income and wealth are independently distributed, we show that the excess burden from joint tax reforms is the sum of the excess burdens of separable tax reforms.

We then turn to the more realistic case that income and wealth are positively correlated. We derive formulas for the marginal excess burden and the welfare effects of joint reforms. We decompose these formulas into (i) a tagging component that captures the welfare gains from better targeting and (ii) a component that captures the 'double distortion' of such reforms.

Then, we bring our formulas to the data and measure revenue and welfare effects of introducing jointness in tax schedules in the U.S. We show that joint tax reforms have a marginal excess burden that is lower than income tax reforms, and higher than wealth tax reforms – at least if the elasticity of taxable wealth is sufficiently small.

We also try to answer the question of how joint tax reforms affect welfare without adopting a normative viewpoint. We do so by assuming the income and wealth tax schedules we observe are separately optimal. This assumption allows us to infer marginal welfare weights for any point in the joint income and wealth distribution. We show that joint tax reforms have positive welfare effects, even though separable tax reforms are not able to improve welfare by construction.

Interpretation of the positive welfare effects depends on the construction of joint tax reforms. If they are constructed by setting a wealth threshold for an income tax increase, the result stems from the shift of distortion to wealth taxation, which is less distortionary. If they are constructed by setting an income threshold for a wealth tax results increase, relieving those with high wealth and low income leads to positive welfare effects, because they have higher welfare weights relative to average. Last but not least, it needs not to be forgotten that the effect of a joint tax reform does not depend on the way it is constructed. Only the interpretation changes, as we start from different counterfactuals.

Related Literature. This paper brings together the insights of the nonlinear income taxation literature Mirrlees (1971); Saez (2001) and the literature on capital taxation. Regarding the latter, our paper builds in particular on Saez and Stantcheva (2018) and rely on sufficient statistics to measure individual responses to taxation. However, a crucial difference in our paper is that we consider possible non-separability in income and capital taxes. Concretely, we consider joint tax reforms as first formalized by Golosov, Tsyvinski and Werquin (2014).

2 Theory

In this section, we explore the effects of different types of tax reforms on government revenue. We assume that the model is static and takes a reduced-form approach to taxes. To do so, we rely on elasticities of income and wealth with respect to net-of-tax rates, which are sufficient statistics to measure the effects of a reform. These elasticities are defined as follows.

$$\varepsilon_{y,1-\tau_Y}(y,a) \equiv \frac{\partial y}{\partial(1-\tau_Y(y,a))}\frac{1-\tau_Y(y,a)}{y}, \qquad \varepsilon_{a,1-\tau_A}(y,a) \equiv \frac{\partial a}{\partial(1-\tau_A(y,a))}\frac{1-\tau_A(y,a)}{a},$$

where y represents income, a wealth, and $\tau_Y(y,a)$ and $\tau_A(y,a)$ represent marginal income tax rates with respect to income and wealth, respectively.

2.1 Uni- and Bivariate Tax Payment Functions

In a univariate tax payment function, total tax liability depends on a single variable. For example, taxable income by itself determines the total amount of income tax liability. The tax payment function may be an involved relationship between the taxable income and the income tax liability including progressivity, a linear relation that incorporates a constant marginal tax rate or even be a constant relation which means that the tax liability is the same for any level of taxable income.

If we allow another variable, such as wealth, to affect the overall tax liability, then the tax payment function is not a univariate function anymore. It is a bivariate function, that is the function has two arguments. Let us denote the bivariate tax payment function that depends on income and wealth as $T(y,a)$. Then, we can write the marginal tax rates as the partial derivatives with respect to individual arguments.

$$\tau_Y(y,a) = \frac{\partial T(y,a)}{\partial y} \quad \text{and} \quad \tau_A(y,a) = \frac{\partial T(y,a)}{\partial a},$$

The simplest way to obtain a bivariate tax payment function is to sum up two individual tax payment functions. In such a case, the tax payment function is called a *separable* function. For a separable tax payment function,

$$T(y, a) = T_Y(y) + T_A(a),$$

where $T_Y(\cdot)$ and $T_A(\cdot)$ represent the tax payment functions for income and wealth respectively.

If a bivariate tax payment function is separable, then it also exhibits independence for its marginal tax rates. That is, the marginal tax rate for one of the arguments does not depend on the other argument. The marginal income tax rate is the same for all wealth levels given income, and vice versa.

$$\tau_Y(y) = \tau_Y(y, a)\ \forall a \quad \text{and} \quad \tau_A(a) = \tau_A(y, a)\ \forall y$$

2.2 Tax Reforms

Elementary tax reforms, that increase the tax liability for certain income levels, are widely used to assess the effects of taxation. They are regarded as the building blocks of the tax payment functions because any tax payment function can be constructed by a linear combination of several elementary tax reforms. An elementary income tax reform that increases the income tax liability of the individuals who earn above a certain threshold is depicted below.

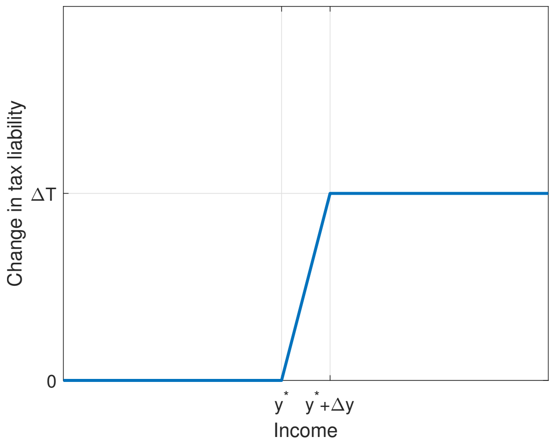

Figure 1: Elementary income tax reform. The figure shows the change in tax liability as a function of income.

There are two effects of an elementary income tax reform. Firstly, individuals who earn more than the income threshold pay more income taxes, while they are still facing the same marginal income tax rate. Without income effects, these individuals will not change their behavior as a result of the reform. Secondly, the individuals who earn around the income threshold face higher marginal tax rates, which will cause them to decrease their labor supply. These individuals will not experience a change in their utility due to the change in their behavior, given that the tax reform is sufficiently small.

The elementary tax reforms of one variable (i.e. separable reforms) can also be extrapolated to tax payment functions with more than one variable, see Golosov, Tsyvinski and Werquin (2014). We can interpret an elementary income tax reform as an increase in the total tax liability for those whose income is above the threshold of the reform, independent of the level of wealth. However, these separable reforms are no longer enough to construct any bivariate tax function. To do so, two additional types of elementary tax reforms have to be introduced. Starting from the separable income tax reform, one can set a wealth threshold, so that only the individuals who have more wealth is affected by the tax liability increase. This is called a *type-1 elementary tax reform*. On the other hand, the wealth threshold can be set, so that only the individuals who have less wealth is affected. This is called a *type-2 elementary tax reform*. The following figure illustrates how a separable income tax reform can be transformed into a type-1 tax reform.

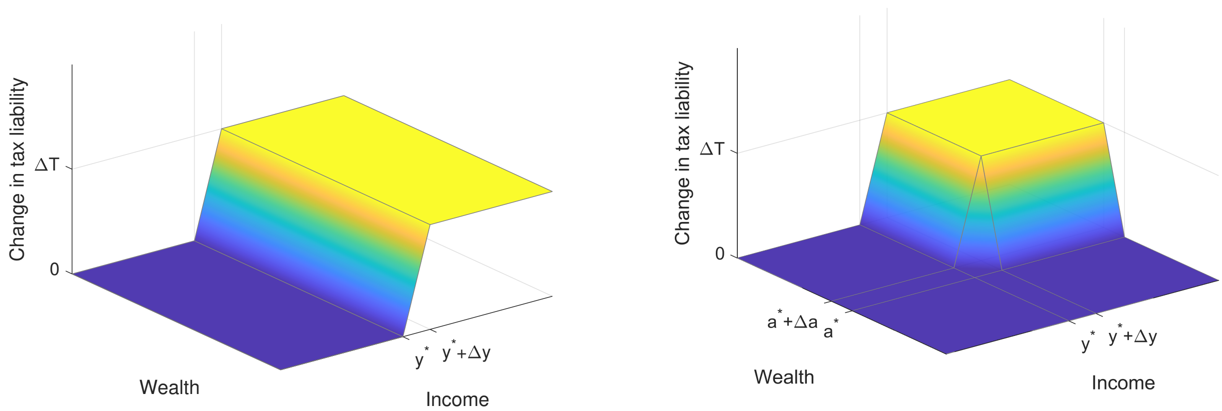

Figure 2: Relation between a separable income tax reform and a type-1 tax reform. The left-hand side panel shows a separable income tax reform that is extrapolated for a bivariate tax payment function. The right-hand side panel shows the same reform with a wealth threshold, which is a type-1 tax reform.

For separable tax reforms, three parameters are enough to fully describe an elementary tax reform. They are conventionally chosen as the location of the reform, the width of the reform, and the change in the marginal tax rate. Multiplication of the second and the third parameters yields the change in tax payment. That is, $\Delta T = \Delta\tau_Y \cdot \Delta y$. Even though the analysis carried out while they are approaching zero, it is crucial to define them to quantify the separate effects.

For joint tax reforms, one needs five parameters to fully describe an elementary reform as described above. Two parameters are needed to determine the location of the reform and three parameters are needed to determine the size of the reform. For joint tax reforms, it is more favorable to define the change in tax payment explicitly, instead of the change in the marginal tax rates. Then, the changes in the marginal income tax rate and the marginal wealth tax rate are implicitly defined by the width of the reform. That is, $\Delta\tau_Y = \Delta T/\Delta y$ and $\Delta\tau_A = \Delta T/\Delta a$.

There are three effects of a type-1 tax reform. Firstly, individuals who are above both thresholds pay more taxes. Absent any income effects, these individuals will not change their behavior and some tax revenue is raised mechanically. The increase in the tax revenue is given by

$$\Delta R_1^M(y^*, a^*) = \Delta T \int_{a^*}^{\infty} \int_{y^*}^{\infty} f(y,a)\, dy\, da, \tag{1}$$

where $f(y,a)$ is the joint probability density distribution of income and wealth, and subscript 1 indicates a type-1 tax reform.

The second effect is the decrease in the income of those who are at the income threshold and have more wealth than the wealth threshold. This is because that they face higher marginal income taxes. Thirdly, as an analog to the second effect, individuals who are at the wealth threshold and have more income than the income threshold will reduce their wealth.By combining these two effects, one can get the total substitution effect, and it is given by

$$\begin{aligned}\Delta R_1^S(y^*, a^*) = &- \int_{a^*}^{\infty} \tau_Y(y^*, a) \cdot \varepsilon_{y,1-\tau_Y}(y^*, a) \cdot \frac{y^*}{1-\tau_Y(y^*, a)} \cdot \frac{\Delta T}{\Delta y} \cdot f(y^*, a)\, \Delta y\, da \\ &- \int_{y^*}^{\infty} \tau_A(y, a^*) \cdot \varepsilon_{a,1-\tau_A}(y, a^*) \cdot \frac{a^*}{1-\tau_A(y, a^*)} \cdot \frac{\Delta T}{\Delta a} \cdot f(y, a^*)\, \Delta a\, dy. \end{aligned} \tag{2}$$

2.3 Marginal Excess Burden

The marginal excess burden is a powerful concept to assess the revenue effects of a policy change, in particular a tax reform. It is the amount of revenue that the government loses due to behavior changes as a ratio of how much the government gains mechanically.

For a type-1 tax reform that is described above, we can calculate the marginal excess burden by $MEB_1(y,a) = \Delta R_1^S(y,a)/\Delta R_1^M(y,a)$. The higher the ratio is, the more the government loses due to distortions of taxes. If the ratio is equal to one, then the tax reform is budget-neutral.

Proposition 1. *Assume that the initial tax payment function is separable, and the elasticity of one variable with respect to net-of-tax rate does not depend on the other variable. Then, the marginal excess burden of a type-1 joint reform is a weighted sum of the separable reforms that are carried out at the same* *level of income and wealth, respectively. That is*

$$MEB_1(y^*, a^*) = w_Y(y^*, a^*) \cdot MEB_Y(y^*) + w_A(y^*, a^*) \cdot MEB_A(a^*)$$

where $w_Y(y, a)$ and $w_A(y, a)$ are the weights that depend on the joint distribution of income and wealth. They are given by

$$w_Y(y^*, a^*) = \frac{Pr(a > a^*|y = y^*)}{Pr(a > a^*|y > y^*)}$$
$$w_A(y^*, a^*) = \frac{Pr(y > y^*|a = a^*)}{Pr(y > y^*|a > a^*)}$$

Proof. See Appendix A. □

The weights in Proposition 1 have a very intuitive interpretation. To understand $w_Y(y, a)$, assume that we construct the joint tax reform by setting a wealth threshold for a separable income tax reform. The wealth threshold will scale down both the mechanical effect and substitution effect of the separable reform. The nominator, $Pr(a > a^*|y = y^*)$, measures how many of the individuals, that are at the income threshold, are above the wealth threshold. In other words, it measures which fraction of the individuals are still affected by the substitution effect. Similarly, the denominator, $Pr(a > a^*|y > y^*)$, measures which fraction of the individuals are still affected by the mechanical effect. If the ratio is less than one, setting a threshold scales down the marginal excess burden. Figure 3 illustrates how $w_Y(y^*, a^*)$ calculated.

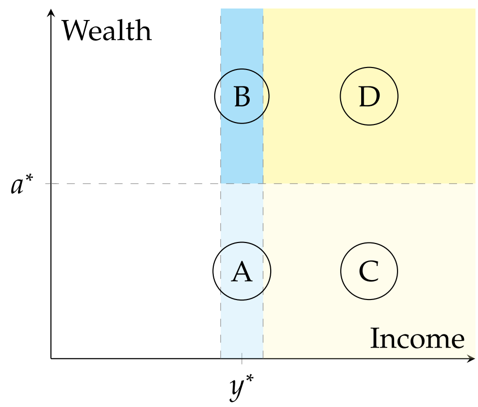

Figure 3: Illustration of how weights are calculated in Proposition 1. Let the circled letters represent the number of individuals in each region. Then $w_Y(y^*, a^*) = \frac{B/(A+B)}{D/(C+D)}$.

The second term, $w_A(y, a)MEB_A(y, a)$, is the adverse effect of setting a wealth threshold. Because wealth is not exogenous in the model, setting a wealth threshold distorts wealth accumulation. That is due to the increase in the wealth tax rate, which is unavoidable when setting a wealth threshold. The threshold creates an incentive to reduce total tax liability by reducing wealth.

This interpretation is very similar to the concept of *tagging*. Assuming that both income and wealth are correlated with the individual property that the government wants to redistribute across (e.g. ability), then tagging income taxation with wealth, or equivalently setting a wealth threshold, increases the effectiveness of taxation. However, there is one critical distinction in this case. The tagged variable is also endogenous, therefore responds to incentives. Then, one can interpret the weight, $w_Y(y,a)$[1], as the benefit of tagging by decreasing the marginal excess burden, and the whole second term, $w_A(y,a)MEB_A(y,a)$, as the costs of tagging due to endogeneity of the tagged variable.

Corollary 1. *If the distributions of income and wealth are independent of each other, then the marginal excess burden of a type-1 joint reform is the sum of the separable reforms that are carried out at the same level of income and wealth, respectively. That is*

$$MEB_1(y,a) = MEB_Y(y) + MEB_A(a),$$

Proof. For independent distributions $P(y|a) = P(y)$ and $P(a|y) = P(a)$. Then, $w_Y(y,a) = w_A(y,a) = 1$ $\square$

Corollary 1 shows that a type-1 tax reform is not better than either of the separable tax reforms in terms of revenue if the distributions are independent. The marginal excess burden of a type-1 tax reform is exactly the sum of two separable tax reforms. This is due to fact that these reforms introduce one more distortion, while independence ensures that there are no gains by setting a threshold.

Consider the case where a type-1 tax reform is constructed by setting a wealth threshold to a separable income tax reform. Initially, the wealth threshold scales down the initial mechanical and substitution effects of the separable reform at the same rate, thanks to the independence assumption between income and wealth. Then, the new distortion on wealth causes the marginal excess burden of a type-1 reform to be strictly greater than a separable income tax reform.

[1]It is less than one if income and wealth are positively correlated.

3 Empirics

The previous section showed us that type-1 tax reforms are strictly worse than their separable counterparts. However, this result heavily depends on the independence assumption between income and wealth. If this assumption is relaxed, the marginal excess burden of a type-1 tax reform is a weighted sum of those of separable reforms. These weights depend on the correlation between income and wealth in the economy.

It is an established fact in the literature that income and wealth are highly correlated with each other (Ríos-Rull and Kuhn, 2016). It is an empirical question whether this correlation is high enough to overturn the previous result. To answer that question, we turn to the 2016 wave of the Survey of Consumer Finances. The Survey of Consumer Finances is a survey of over 6000 United States households. It is conducted every three years as a cross-section and it should represent the whole U.S. population.

For our purposes, we define income as the sum of wages, salaries, and self-employment income, as well as capital income and transfers. Further, we define wealth as the sum of all financial and non-financial assets, excluding retirement savings, minus the total household debt. The wealth of a household represents its net worth.

	Income	Wealth
Mean	$106,000	$772,000
Median	$58,700	$141,880
Gini index	0.575	0.824
90-50 ratio	3.18	9.31
50-10 ratio	3.41	24.5
Correlation	0.484	

Table 1: Summary statistics of the distributions of income and wealth.

Table 1 shows some parameters about the distributions of income and wealth. We see that wealth is much more unequally distributed than income. More crucially, the Pearson correlation coefficient is equal to 0.48, indicating a high correlation between income and wealth.

The joint distribution of income and wealth is crucial to assess the effects of joint reforms. The revenue effects of joint reform compared to the separable reforms depend directly on the weights in Proposition 1. Moreover, these weights only depend on the shape of the joint distribution. The higher the correlation between income and wealth is, the lower the weights are.

Instead of relying on parametric function assumptions, we estimate it non-parametrically using the kernel density method. This method provides us with a density value for each income-wealth pair, (y, a). It also yields the marginal distributions of income and wealth.

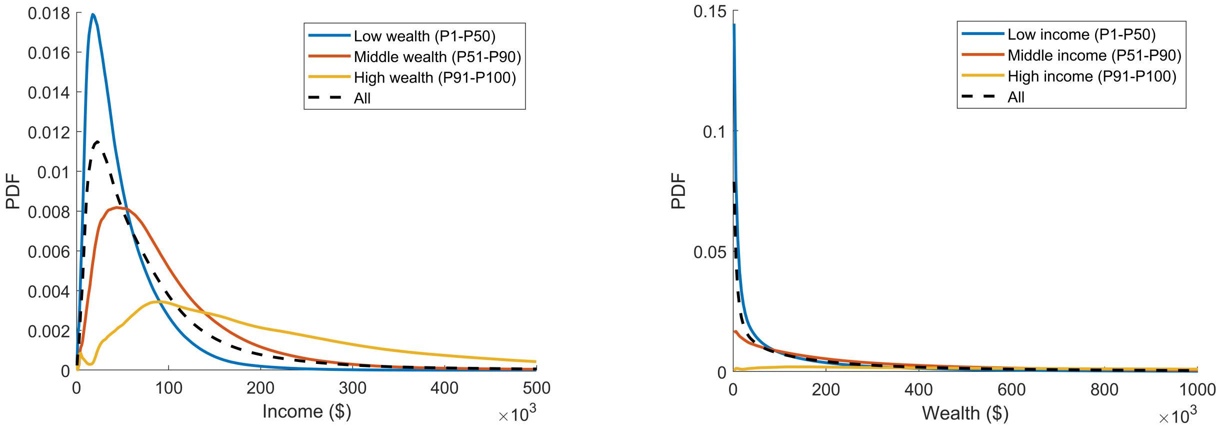

Figure 4: Conditional distribution of income and wealth. The left-hand side panel shows the income distribution conditional on being in a certain wealth group. Colored lines represent different wealth groups, whereas dashed black line represents the unconditional distribution. The right-hand side panel does the same thing for the wealth distribution conditional on income.

Figure 4 indicates that the distribution of income and wealth are not independent of each other. On the left-hand side panel, the blue line has a thinner tail compared to the dashed black line. This means that an individual who has below-median wealth is less likely to have high income compared to the unconditional distribution. On the other hand, having above-median wealth is correlated with having a high income. Moreover, the likelihood of having a high income is the highest for the highest wealth group. Similar correlation arguments can also be made for the wealth distribution as seen on the right-hand side panel.

3.1 Marginal Excess Burden

To compare marginal excess burdens of different reforms one has to make some assumptions about the initial tax schedule. Because the amount tax revenue of tax revenue lost depends on initial marginal tax rates. An increase in the marginal income or wealth tax rate causes people to decrease their before-tax income or wealth. This decrease in gross variables, then, leads to a decrease in tax revenue.

We assume that we are starting from two separable tax functions for income and wealth. This assumption allows us to assess the effects of deliberately introducing jointness to the tax schedule.

We use income tax simulator of the NBER[2] to estimate marginal income tax rates. The program calculates one's tax liability under US federal and state income tax laws. For our purposes, we omitted state income taxes and focused solely on federal income taxes. By observing the change in tax liability when income is increased by a small amount, we can infer the marginal tax rate that the individual faces.

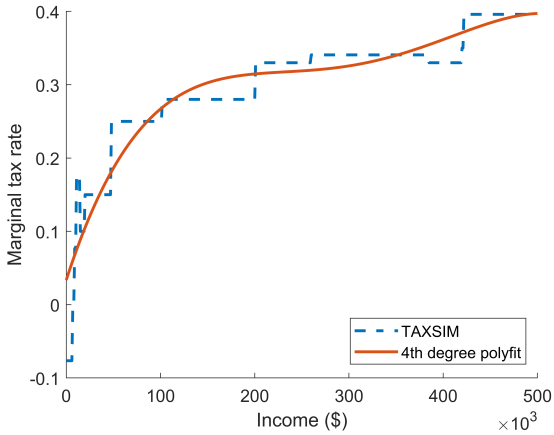

Figure 5: Estimating marginal income tax rate. Dashed blue line represent the marginal income tax rate according to the tax simulator. Red line is the polynomial fit of fourth degree.

Figure 5 shows the estimated marginal income tax rate for each income level. The fourth-order polynomial, the red line in the figure, is able to closely imitate the tax calculator output.

Currently, there is no wealth taxation in the US, even though there is an ongoing debate whether or not there should be. Therefore, we have to propose a hypothetical wealth taxation system. The corporate income tax rate in the US was 35% in 2016[3]. We can use the corporate income tax rate to calculate the equivalent wealth tax rate. If returns of capital is homogeneous, then a wealth tax rate, $\tau_A = r\tau_k/(1+r)$, is equivalent to the capital tax rate, τ_k. Assuming 3% annual interest rate, we use a constant wealth tax rate of 1%, as the initial wealth tax schedule.

The assumption of initial marginal income tax rates and marginal wealth tax rates allows us to carry out our analysis. First, we calculate the marginal excess burden for separable income and wealth tax reforms positioned at different levels. Then, we calculate the marginal excess burden of type-1 tax reforms positioned at different income-wealth pairs. A type-1 tax reform positioned at (y^*, a^*) increases marginal income tax rate at y^*, conditional on having higher wealth than a^*.Equivalently, it can be interpreted as the increase in marginal wealth tax rate at a^*, conditional on having a higher income than y^*.

Figure 6 shows the marginal excess burden of separable reforms as well as the marginal

[2]TAXSIM v32: `https://users.nber.org/~taxsim/taxsim32/`
[3]The data we use is from 2016.

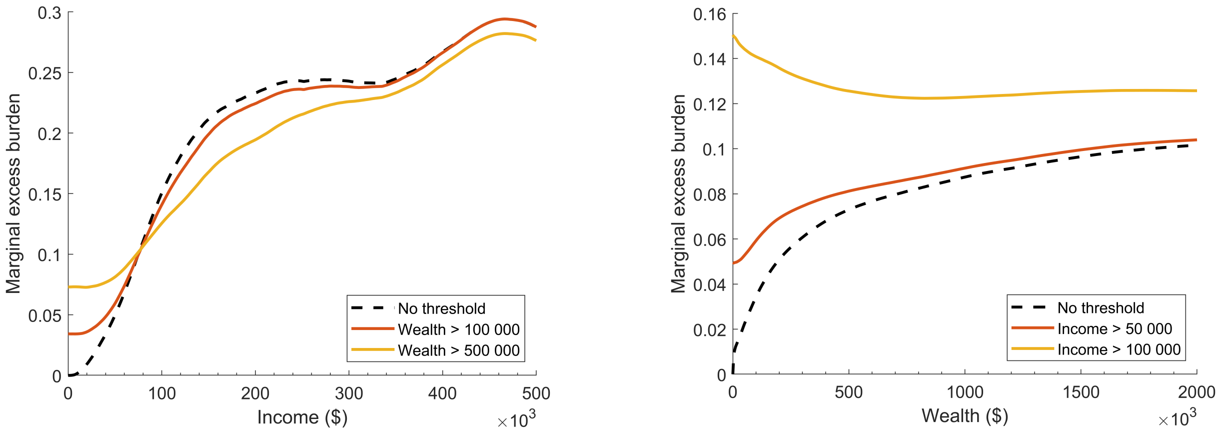

Figure 6: Marginal excess burden of various reforms. Left-hand side panel shows the separable income tax reform and type-1 reforms with two different wealth thresholds. Right-hand side panel shows the separable wealth tax reform and type-1 reforms with two different income thresholds.

excess burden of type-1 reforms with two different interpretations. In order to have a better understanding of the figure, consider a type-1 tax reform positioned at the point ($y^* = \$100,000$, $a^* = \$500,000$), whose marginal excess burden is equal to 12.6%. This reform is present in both panels.On the left-hand side panel it is interpreted as a marginal income tax rate increase at $y^* = \$100,000$, with a wealth threshold of $w^* = \$500,000$. Its marginal excess burden is given by the value of the yellow line at $y^* = \$100,000$. On the right-hand side panel, the same value is given by the yellow line at $w^* = \$500,000$.

When joint reforms are interpreted from an income tax perspective, setting a wealth threshold decreases the marginal excess burden of reforms. On the other hand, from a wealth tax perspective, setting an income threshold increases the marginal excess burden of the reforms. It is not complicated to reconcile these two facts with each other. Proposition 1 states that the marginal excess burden of type-1 reforms is a weighted sum of those of separable reforms. The weights are less than one for correlated variables, such as income and wealth. Therefore, joint reforms tend to average between two separable reforms. Moreover, the marginal excess burden of income tax reforms is generally higher than that of wealth tax reforms. These together result in joint reforms having a lower marginal excess burden than income tax reforms, and a higher marginal excess burden than wealth tax reforms.

3.2 Welfare Effects

After answering the question of how does joint tax reforms affect welfare, a natural next question to as is what about the well-being of those who pay more taxes. In this subsection, we try to answer that exact question.

Individuals who are affected by higher marginal tax rates, either on income or on wealth, do not experience any change in their utility thanks to the envelope theorem. The only relevant effect in terms of welfare is the effect through public funds. That is, the substitution effect in terms of revenue and welfare are equal to each other. For separable income tax reforms, the substitution effect is given by

$$\Delta W_Y^S(y^*) = \Delta R_Y^S(y^*) = \tau_Y(y^*) \cdot \varepsilon_{y,1-\tau_Y}(y^*) \cdot \frac{y^*}{1-\tau_Y(y^*)} \cdot \frac{\Delta T}{\Delta y} \cdot f_Y(y^*)\,\Delta y,$$

where $\Delta R_Y^S(y^*)$ is the revenue effect of a separable income tax reform at y^* due to substitution.

Individuals who are affected by higher tax liabilities, on the other hand, will experience a decrease in their utility due to the increase in average tax rates. However, one has to adopt a normative view on the social welfare objective to evaluate the effect of this change on welfare. It is not clear which view is optimal, or desired, for a given society. Temporarily, assume that $g(y,a)$ is the marginal welfare weight of an individual at (y,a). Then, the mechanical welfare effect of a separable income tax reform is given by

$$\Delta W_Y^M(y^*) = \Delta T \int_{y^*}^{\infty} (1 - g(y)) f_Y(y)\, dy,$$

where

$$g_Y(y) = \frac{\int_0^\infty g(y,a) f(y,a)\, da}{f_Y(y)}$$

is the average marginal welfare weight at a certain income level.

We refrain from adopting a social welfare objective. Instead, we turn to our estimates of two different separable tax schedules and assume that they are optimal as we observe. This assumption results in one key feature. Any sufficiently small separable tax reform has no effect on welfare by assumption. That is, mechanical effect and substitution effect are equal to each other for any separable tax reform, $W_Y^M(y) = W_Y^S(y)\ \forall y$. Then, we can estimate the marginal social welfare weight that is associated with a certain income level.

$$\overline{g}(y^*) = 1 - \frac{\tau_Y(y^*)}{1-\tau_Y(y^*)} \cdot \varepsilon_{y,1-\tau_Y}(y^*) \cdot \frac{y^* f_Y(y^*)}{1-F_Y(y^*)},$$

where

$$\overline{g}(y^*) = \frac{\int_{y^*}^{\infty} g_Y(y)\,dy}{1 - F_Y(y^*)}$$

is the average marginal welfare weight of those who have higher earnings than y^*.

After carrying out the same analysis for wealth taxation as well, we have estimates for welfare weights across income and across wealth, $g_Y(y)$ and $g_A(a)$. Then, we can extrapolate those into two dimensions and obtain $g(y, a)$, by carefully using the joint distribution of income and wealth, $f(y, a)$.

The welfare effect of a type-1 reform due substitution is equal to its revenue effect, as in separable reforms. That is, $\Delta W_1^S(y^*, a^*) = \Delta R_1^S(y^*, a^*)$, given in Equation (2). We can evaluate mechanical welfare effects using marginal welfare weights that are estimated. The total welfare effect of a type-1 reform is given by

$$\Delta W_1(y^*, a^*) = \Delta W_1^M(y^*, a^*) + \Delta W_1^S(y^*, a^*)$$

where

$$\Delta W_1^M(y^*, a^*) = \Delta T \int_{a^*}^{\infty} \int_{y^*}^{\infty} (1 - g(y, a)) f(y, a)\, dy\, da.$$

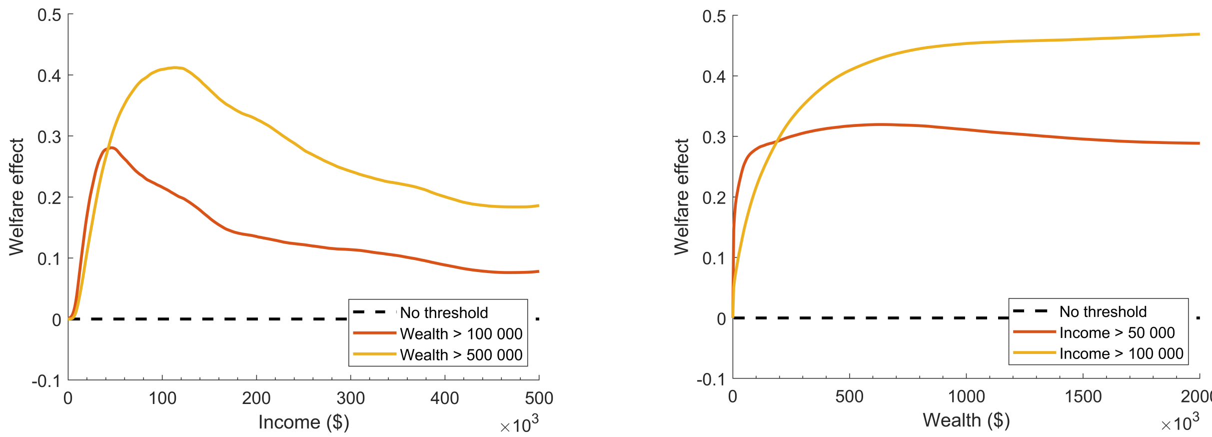

Figure 7: Welfare effect of various reforms. Left-hand side panel shows the separable income tax reform and type-1 reforms with two different wealth thresholds. Right-hand side panel shows the separable wealth tax reform and type-1 reforms with two different income thresholds.

Figure 7 shows the welfare effects of joint tax reforms as well as separable tax reforms with the same configuration as Figure 6. That is, the welfare effect of a joint tax reform positioned at the point ($y^* = \$100,000, a^* = \$500,000$) is equal to 40.9% for every dollar of mechanically raised tax revenue. On the left-hand side panel, it is given by the value of the yellow line at $y = \$100,000$. On the right-hand side panel, it is given by the value of the yellow line at $a = \$500,000$. We can see three important features from the figure.

First, it can be seen that both separable tax reforms have zero welfare effects. This is by construction due to separable welfare weights. By assuming that initial separable tax schedules are optimal, we ensure that the welfare effect of separable tax reforms is zero at any point.

More importantly, we see some positive welfare effects of joint reforms, even though there is no room for improvement by altering separable tax schedules. Setting a wealth threshold at $\$500,000$ increases welfare effect on income tax increases up to more than 40%. On the other hand, setting an income threshold at $\$100,000$ increases the welfare effect of wealth tax increases up to almost 50%.

To understand why setting a wealth threshold increases welfare for income tax increases, consider the change in distorted population. The distortion shifts from income to wealth when setting a wealth threshold. Since, wealth taxes less distortionary than income taxes, this shift results in lower distortion. Therefore, joint tax reforms have a welfare benefit when income tax reforms do not.

The channel is different when setting an income threshold for wealth tax increases. In this case, the shift of distortion is in the opposite direction. Distorting income instead of wealth increases overall distortion. This can also be seen looking at the revenue effects. Joint tax reforms have a higher marginal excess burden than separable wealth tax reforms. However, the question of who pays higher taxes is also important. Setting an income threshold relieves those with lower income from paying higher taxes. Since welfare weights are more heterogeneous across income than across wealth, those with low income and high wealth still have relatively high welfare weights. Relieving them from their higher tax liabilities by setting an income threshold increases welfare compared to a wealth tax reform without a threshold.

Figure 7 shows us that paying attention to only separable tax schedules, $T_Y(y)$ and $T_A(a)$, is not enough to secure an allocation that yields the highest possible welfare. One can achieve higher welfare levels by introducing jointness to the tax schedule so that the marginal tax rate of income depends on wealth and vice versa.

4 Conclusion

In this project, we study the joint taxation of income and wealth. We argue that a classical 'tagging' logic applies: the tail of the wealth distribution is much fatter if we condition on individuals with high incomes. Therefore the top wealth tax should be higher for high incomes ceteris paribus. Such 'double progressive' taxation, however, implies two distortions: the wealth accumulation and the earnings margin are distorted. Bringing our formulas to the data, indicate that likely the tagging argument dominates and it is desirable to have double progressive taxation despite the double distortion.

As a next step, we aim to understand in particular the role of cross-elasticities (how does income respond to changes in wealth taxes and how does wealth accumulation change in response to income taxes?) that we have so far implicitly assumed to be zero.

References

Golosov, Mikhail, Aleh Tsyvinski, and Nicolas Werquin. 2014. "A Variational Approach to the Analysis of Tax Systems." National Bureau of Economic Research Working Paper 20780.

Mirrlees, J A. 1971. "An Exploration in the Theory of Optimum Income Taxation 1,2." *Review of Economic Studies*, 38(2): 175–208.

Ríos-Rull, José-Víctor, and Moritz Kuhn. 2016. "2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics." *Federal Reserve Bank of Minneapolis Quarterly Review*, 37(1): 1–73.

Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies*, 68(1): 205–229.

Saez, Emmanuel, and Stefanie Stantcheva. 2018. "A Simpler Theory of Optimal Capital Taxation." *Journal of Public Economics*, 162: 120–142.

A Proof of Proposition 1

If the initial tax schedule is separable, then the marginal tax rate of one variable does not depend on the other variable. That is

$$\tau_Y(y) = \tau_Y(y,a)\ \forall a \quad \text{and} \quad \tau_A(a) = \tau_A(y,a)\ \forall y.$$

Using the above identity in Equation (2) yields

$$MEB_1(y^*,a^*) = y^*\varepsilon_y(y^*)\frac{\tau_y(y^*)}{1-\tau_y(y^*)}\frac{\int_{a^*}^{\infty} f(y^*,a)\,da}{\overline{F}(y^*,a^*)} + a^*\varepsilon_a(a^*)\frac{\tau_a(a^*)}{1-\tau_a(a^*)}\frac{\int_{y^*}^{\infty} f(y,a^*)\,dy}{\overline{F}(y^*,a^*)}, \tag{3}$$

where $\overline{F}(y^*,a^*)$ is the exceedance function for the joint distribution and given by

$$\overline{F}(y^*,a^*) \equiv \int_{a^*}^{\infty}\int_{y^*}^{\infty} f(y,a)\,dy\,da = Pr(y > y^* \ \wedge\ a > a^*).$$

We can simplify the expression for the MEB_1 further using the marginal excess burden expressions for separable reforms. They are given by

$$MEB_Y(y^*) = y^*\varepsilon_Y(y^*)\frac{\tau_y}{1-\tau_y}\frac{f_Y(y^*)}{\overline{F_Y}(y)},$$
$$MEB_A(a^*) = a^*\varepsilon_A(a^*)\frac{\tau_a}{1-\tau_a}\frac{f_A(a^*)}{\overline{F_A}(a)}.$$

where $f_Y(y)$ and $f_A(a)$ represent marginal probability density functions, and $\overline{F_Y}(y)$ and $\overline{F_A}(a)$ represent exceedance functions for income and wealth.

Plugging these expression into Equation (3) yields

$$MEB_1(y^*,a^*) = \frac{\dfrac{\int_{a^*}^{\infty} f(y^*,a)\,da}{f_Y(y^*)}}{\dfrac{\overline{F}(y^*,a^*)}{\overline{F_Y}(y^*)}}\cdot MEB_Y(y^*) + \frac{\dfrac{\int_{y^*}^{\infty} f(y,a^*)\,dy}{f_A(a^*)}}{\dfrac{\overline{F}(y^*,a^*)}{\overline{F_A}(a^*)}}\cdot MEB_A(a^*).$$

Replace the distributional quantities in the coefficients by probabilities.

$$MEB_1(y^*, a^*) = \frac{\dfrac{Pr(y = y^* \ \wedge \ a > a^*)}{Pr(y = y^*)}}{\dfrac{Pr(y > y^* \ \wedge \ a > a^*)}{Pr(y > y^*)}} \cdot MEB_Y(y^*) + \frac{\dfrac{Pr(y > y^* \ \wedge \ a = a^*)}{Pr(a = a^*)}}{\dfrac{Pr(y > y^* \ \wedge \ a > a^*)}{Pr(a > a^*)}} \cdot MEB_A(a^*).$$

Using the definition of conditional probabilities yields the result.

$$MEB_1(y^*, a^*) = \frac{Pr(a > a^*|y = y^*)}{Pr(a > a^*|y > y^*)} \cdot MEB_Y(y^*) + \frac{Pr(y > y^*|a = a^*)}{Pr(y > y^*|a > a^*)} \cdot MEB_A(a^*). \tag{4}$$