The Overall Redistributive Effect and Progressivity of the Portuguese Tax-Transfer System

Mariana Isabel Sanches Felizardo

152221006

Dissertation written under the supervision of Professor Nuno Alves

Dissertation submitted in partial fulfilment of requirements for the MSc in Economics - Major in Macroeconomic Policy, at Católica-Lisbon School of Business & Economics

April 2023

The Overall Redistributive Effect and Progressivity of the Portuguese Tax-Transfer System

Mariana Felizardo

Abstract

The redistribution effect of the tax-transfer system is a first-order issue for policymakers. The degree of progressivity of taxes and transfers plays a central role in this respect. We apply the Kakwani index to evaluate the redistributive effect and progressivity of the overall tax and transfer system - including cash transfers, direct taxes and indirect tax liabilities - for Portugal in 2010 and 2015. To the best of our knowledge, this computation is novel for the Portuguese case. The first stage of our research project involves using parametric statistical techniques to link expenditure data from the HBS dataset with information on income and transfers from the EU-SILC. After building this new dataset, we assess both the overall redistributive effect and progressivity of the tax-transfer system, as well as the contribution of each tax and benefit instrument. Our empirical analysis shows that the overall redistributive effect translates into an inequality reduction of approximately 6 and 7 Gini percentage points, respectively for 2010 and 2015. The decrease in inequality is mainly driven by direct taxes, and this impact increased between 2010 and 2015. Corroborating the existing literature, cash benefits and direct taxes are progressive, while indirect taxes are regressive. The overall tax-transfer system is progressive. From 2010 to 2015, the Portuguese tax-transfer system became less progressive, yet the overall redistribution effect increased, mostly due to increased average direct tax rates.

Keywords: Kakwani Index, Redistributive Effect, Progressivity, Direct Taxes, Indirect Taxes, Transfers, Parametric Statistical Matching, Portugal

The Overall Redistributive Effect and Progressivity of the Portuguese Tax-Transfer System

Mariana Felizardo

Resumo

O efeito redistributivo de um sistema fiscal é uma questão primordial para os decisores políticos e o grau de progressividade dos impostos e das transferências desempenha um papel central a este respeito. Nesta dissertação, aplicamos o índice Kakwani para avaliar o efeito redistributivo e a progressividade do sistema global de impostos e transferências - incluindo as prestações sociais em dinheiro e impostos diretos e indiretos - para Portugal, em 2010 e 2015. No nosso melhor conhecimento, é a primeira vez que esta análise é feita para o caso português. A primeira etapa do nosso projeto envolve a utilização de métodos estatísticos paramétricos para combinar dados sobre a despesa do IDEF com dados sobre o rendimento e transferências das famílias do EU-SILC. Após a construção deste novo conjunto de dados, avaliamos tanto o efeito redistributivo total e a progressividade do sistema fiscal, como a contribuição de cada instrumento de impostos e benefícios. A nossa análise empírica mostra que o efeito global redistributivo se traduz numa redução da desigualdade de aproximadamente 6 e 7 pontos percentuais de Gini, respetivamente para 2010 e 2015. A diminuição da desigualdade é maioritariamente impulsionada pelos impostos diretos, e este impacto aumentou entre 2010 e 2015. Corroborando a literatura existente, as transferências e os impostos diretos são progressivos, enquanto os impostos indiretos são regressivos. O sistema fiscal é progressivo. De 2010 a 2015, este tornou-se menos progressivo, porém o efeito redistributivo total aumentou, devido sobretudo a um aumento da taxa média de impostos diretos.

Palavras-chave: Índice de Kakwani, Efeito Redistributivo, Progressividade, Impostos Diretos, Impostos Indiretos, Transferências, Modelo Paramétrico de Emparelhamento Estatístico, Portugal

Acknowledgements

First and foremost, I would like to start by thanking my supervisor, Professor Nuno Alves, for all the support, guidance, and time throughout this process.

I would also like to thank Paula Antão, as well as my colleagues at the Bank of Portugal who so dearly welcomed me and daily share their knowledge with me.

Quoting Thomas Aquinas, ”there is nothing on this earth more to be prized than true friendship” and for that, I am grateful to Luana Santos and Guilherme Viegas. A special appreciation to Pedro Baptista for his continuous unwavering encouragement and precious insights that helped me through all the difficulties.

Last, but definitely not least, my gratitude goes to my parents and my grandparents for being an unbelievable support and my greatest admirers throughout my entire academic journey. This endeavour would not have been possible without them.

Contents

List of Figures

List of Tables

1 Introduction

> In technology, we spend so much time experimenting, fine-tuning, getting the absolute cheapest way to do something – so why aren't we doing that with social policy?
>
> *Esther Duflo*

Economic policy in general and social policy in particular aim to improve welfare in a society. Even though there is not a single measure of welfare, most indicators rely on the measurement of the levels and distribution of variables such as income, consumption or wealth. Using these variables to perform economic analyses allows us to better understand the reality we live in.

Market forces typically generate high levels of inequality of income and wealth. These high disparities are unwarranted not only because of normative reasons but also because they arguably threaten economic growth, social cohesion and ultimately democracy (Ritzen and Woolcock, 2000). Governments thus aim to directly affect income distribution through the conduct of fiscal policy, to ensure a distribution of produced goods aligned with social preferences. Redistribution is valued by society not only for strictly utilitarian reasons - assuming that the marginal utility of consumption of individuals is decreasing with the level of income - but mainly with a view to correct distortions in income distribution arising from the lack of effective equality of opportunity among citizens. In this regard, both tax and transfer policies play a prominent role and need to be carefully designed to jointly maximize efficiency and equity. [1]

In Portugal, the concern for income redistribution is explicitly recognized in the Constitution. Article 104 states that the personal income tax shall be progressive and aim at decreasing inequalities. However, the redistributive impact of fiscal policy is not confined to direct taxation but it also includes indirect taxes and transfers to households.

The importance of evaluating the redistribution effects and progressivity of taxes alongside benefits has long been recognized in literature, *viz* Podger et al. (1980). The reasons are straightforward: first, the main policy instruments used by governments to accomplish a redistribution of resources are taxes and transfers; second, tax and transfer systems usually overlap, so they can

[1] One important distinction to be made is the one between equality and equity since these two concepts are often mistakenly used. The definitions commonly associated with the words show interchangeability rather than sharp contrast. Indeed, they are commonly used to define one another. The Oxford English Dictionary gives "fairness, impartiality, equity" as a definition of "equality" and defines "equity" as "the quality of being equal or fair." In other words, the equality of income distribution is essentially a matter of fact, and it is, therefore, objective. The equity of the same distribution is a matter of ethical judgment, and it is, therefore, subjective. In this research project, we shall refer to the concept of equality or inequality.

strengthen or counteract each other, affecting important micro- and macroeconomic variables, such as GDP or savings; and third, welfare spending *i.e.*, social benefits or transfers, are funded through taxation.

In this dissertation, we evaluate the redistributive effect and progressivity of the overall tax and transfer system in Portugal. To the best of our knowledge, this computation is novel for the Portuguese case. The reason for this gap is that there is no comprehensive data source that simultaneously covers information on direct and indirect taxes paid by households, alongside transfers received. The first stage of our research project thus involves using statistical techniques to link expenditure data from the Household Budget Survey (HBS) with information on income and transfers from the European Union Statistics on Income and Living Conditions (EU-SILC), as proposed by Serafino and Tonkin (2017). We are able to build the database for two years: 2010 and 2015 (with respective income reference years 2009 and 2014, respectively). After building our new dataset, we assess both the overall redistributive effect and progressivity of the tax and transfer system, as well as the contribution of each tax and benefit instrument, for both years. Our computations follow the methodology proposed by Jenkins and Hérault (2021). We will respectively use the terms transfers and benefits, and pre- or post-fisc and pre- or post-tax, interchangeably.

Our key findings are as follows. First, the overall redistributive effect of the Portuguese tax-transfer system translates into an inequality reduction of approximately 6 and 7 Gini percentage points, respectively for 2010 and 2015. In 2015, the inequality reduction in terms of the Gini index is mainly driven by direct taxes - personal income tax (PIT), social security contributions (SSC) and wealth taxes. Second, corroborating the existing literature, cash benefits and direct taxes are progressive, while indirect taxes are regressive, both for 2010 and 2015. The tax-transfer system as a whole is progressive. Third, from 2010 to 2015, the Portuguese tax-transfer system became less progressive, yet the overall redistribution effect is greater in 2015, mostly due to the increase in direct tax rates.

The remainder of the paper proceeds as follows. Section 2 reviews the relevant literature on progressivity measurement and redistributive effects, including empirical studies for Portugal. Section 3 describes the datasets used and presents the methodology employed to build our joint dataset on taxes and transfers. Section 4 presents the empirical results on the progressivity of the tax-transfer system in Portugal. Section 5 compares the results for 2010 and 2015. Lastly, section 6 concludes and suggests some avenues for future research.

2 Literature Review

Throughout the years, there have been numerous studies on the measurement of tax progressivity, with a variety of indices being proposed and used to assess the degree to which a tax system is progressive or regressive. Even though many of the classical works on the measurement of tax progressivity originate from the 1950s, they are still of great relevance today. Local tax progression measures – particularly those relating to tax elasticity and residual income elasticity – were the very first ones of this vast literature. Dalton (1922) and Musgrave and Thin (1948) are two examples. The former proposes the use of the Gini coefficient as a measure of income inequality, which has since become one of the most widely used measures in economics and other social sciences. The latter argues that the marginal tax rate is an important measure of progressivity because it determines either the incentive to work or to save. Both early works rely on measurements of pre- and post-tax income inequality, which may alternatively be stated as weighted deviations between actual taxation and hypothetical proportional taxation that is revenue-neutral. However, tax schedules that have lower average tax rates for certain income ranges may be seen as more progressive than those with consistently increasing tax rates. This disadvantage prompted the creation of the global tax progression measures, which are based on the dominance relationships of Lorenz Curves.

The measurement of progressivity and vertical equity moved forward significantly in the middle of the 1970s, following slightly earlier advances in the measurement of inequality, as is the case of Fellman (1976), Jakobsson (1976) and Kakwani (1977a). Later on, Kakwani (1977b), Suits (1977), and Reynolds and Smolensky (1977) developed influential indices of tax progressivity and vertical equity.

The findings of Kakwani (1977a,b, 1984) have been noteworthy in shaping research and policy discussions on poverty and inequality. They provide a comprehensive overview of the methods and approaches used to measure and decompose income inequality and poverty and highlight the policy implications of these measures. They also introduce new techniques for decomposing income inequality and poverty, which have since been widely used in empirical research and policy analysis. The vertical redistribution effect of income taxes and cash benefits is a principle of distributive justice that is rarely questioned as such, which Kakwani deftly and skillfully expressed in terms of two independent concepts: the average tax (benefits) rate, reflecting the overall tax (benefits) level, and a measure of tax (benefits) progression, which summarizes how individual tax payments stray from proportionality. This means that the distributional effect of taxation is induced by changes in the average tax rate, by changes in the degree of progressivity or by both.

Kiefer (1984) criticizes the above-mentioned measures of tax progressivity for failing to take into account the distribution of income. He argues that measures such as the average tax rate, marginal tax rate, and effective tax rate are not sufficient for assessing the progressivity of a tax

system, as they do not consider how the burden of taxation is distributed across different income groups. The author also criticizes the use of standard measures of progressivity, such as the Kakwani index and the Suits index, which are based on the concept of vertical equity. By contending that these measures do not account for differences in the shape of the income distribution, Kiefer believes they can produce misleading results when applied to tax systems with non-linear tax schedules. Furthermore, the author suggests that some progressivity measures may be overly influenced by extreme values in the income distribution, which can result in misleading conclusions about the overall progressivity of a tax system. Kiefer's distributive tax progressivity indices, by considering the entire income distribution, provide a more comprehensive and accurate measure of tax progressivity.

Indeed, Kiefer's critiques have been subject to some debate and discussion in the field of public economics. Authors such as Tony Atkinson, Peter Lambert, and Amartya Sen also contributed to the debate with different arguments. They defend that the Kakwani and Suits indices are not intended to be comprehensive measures of tax progressivity but rather specific tools for assessing the extent to which a tax system is vertically equitable. These measures are designed to capture the degree to which tax burdens are distributed in proportion to income, rather than to provide a complete picture of the distributional effects of a tax system. [2] Moreover, those indices can be useful in situations where the distribution of income is relatively stable, as they are less sensitive to changes in the shape of the income distribution. In contrast, Kiefer's distributive tax progressivity indices may be more sensitive to changes in the distribution of income, which could lead to some volatility in the results.

Despite the ongoing debate on the measurement of tax progressivity, Kakwani's decomposition of redistributive effects into vertical and reranking indices has become a cornerstone in the literature. This is evidenced by the numerous empirical studies that have employed it, as well as the many extensions and improvements that have been proposed by its advocates. The reason for the popularity of the Kakwani index lies in its comprehensive nature, simplicity, ease of computation, and its ability to provide straightforward policy interpretation. A recent application of these techniques can be found in Jenkins and Hérault (2021).

With regard to Portugal, some individual analyses of the redistributive effects and progressivity of tax and transfer instruments were conducted as well. Martins (2016) explores the evolution of the Portuguese personal income tax system progressivity over the period of 2005 through 2013 and finds that progressivity had very modest variations from 2005 to 2012, whilst the jump from 2012 to 2013 was the largest in the timespan considered, mostly due to changes in personal income tax schedule. Matos (2017) uses microdata from the Household Budget Survey (HBS) to measure the

[2] We can apply the famous quote by Mark Twain, "To a man with a hammer, everything looks like a nail". to this debate. One shall not fall into Maslow's Law, a cognitive bias that involves an over-reliance on a familiar tool. In our case, the tool is the use of the same index of progressivity for every purpose.

extent of progressivity of income taxes and of the value-added tax (VAT), together with social security contributions. Converting the net-at-source income into gross income and using the Reynolds-Smolensky and Kiefer indices, the author finds that personal income tax is progressive with withheld taxation being more progressive than final taxes and the VAT is found to be regressive. Araújo (2019) and Braz and Correia da Cunha (2009) focus on the analysis of indirect taxes. The former uses the income-based Kakwani progressivity index on HBS data and concludes that the global indirect tax system in Portugal is regressive; the latter uses the Household Expenditure Survey (HES) to analyse the redistributive impact of VAT in Portugal and finds three sets of results. First, Braz and Correia da Cunha (2009) concludes the relationship between the share of expenditure on classes of goods and services and net income distribution to be negative for goods such as food and it is positive for services such as recreation and culture, restaurants and hotels, and services which are exempt or taxed at the reduced or intermediate VAT rates, particularly education. Second, they find that VAT appears to be a slightly progressive tax relative to expenditure. Third and last, the authors conclude that horizontal and reranking effects are minor in the case of VAT in Portugal.

Reis (2021) empirically assesses the effect that tax decisions have had on the degree of progressivity, redistribution, equity and social welfare, using traditional indicators of local and global progressivity. The author concludes that both the degree of progressivity and redistribution of the personal income tax have increased, between 1997 and 2018, along with vertical equity and horizontal inequality. Moreover, the existing structure of increasing marginal rates by income bracket is found to be superior in terms of social well-being to a proportional structure.

3 Data, Method and Descriptive Statistics

Assessing tax progressivity indices to measure inequality and redistribution requires detailed microdata with a breakdown of household income. The standard source in the literature is household surveys. We combine two household surveys in order to perform our analysis: the first is the European Union Statistics on Income and Living Conditions (EU-SILC); the second is the Household Budget Survey (HBS).

The remainder of this section is organized as follows. We start by briefly presenting each database and the income definitions that will be used in the analysis (subsection 3.1). Then we describe the computation of indirect taxes using the HBS (subsection 3.2). Subsection 3.3 details the methodology to merge both databases and subsection 3.4 presents some descriptive statistics of the final dataset.

3.1 EU-SILC and HBS databases

The EU-SILC database is the prime source for analyzing household income, poverty, social exclusion and living conditions at the European level. The EU-SILC has detailed information on wages, pensions, transfers, direct taxes and social contributions at the individual and household levels. This dataset provides us with two types of timely and comparable data: cross-sectional and longitudinal - we will use the former for Portugal. This sample survey is produced annually. Given our final aim to merge the EU-SILC and HBS databases, we will only employ the surveys for the years 2010 and 2015 (the ones for which the HBS was conducted). In both databases, the information on the income of individuals and households refers to the previous year (2009 and 2014). We have information on 5,182 and 8,740 households for 2010 and 2015, respectively corresponding to 13,368 and 21,965 individuals.

The HBS - in Portuguese, *Inquérito às Despesas das Famílias (IDEF)* - is a statistical survey carried out by Statistics Portugal that collects data every five years on the income and expenditure of households. The survey typically includes information on the sources and amounts of income, as well as the types and amounts of goods and services that households purchase. The purpose of the survey is to periodically provide detailed information on the consumption and saving patterns of households, as well as on their socio-economic characteristics. Moreover, it provides information on food quantities. We will use the HBS to compute the indirect taxes paid by households. Our sample contains information on 9,447 and 11,334 households for 2010 and 2015, respectively corresponding to 24,323 and 28,982 individuals.

Income Definitions

Our analysis of the progressivity of benefits and direct taxes is based on four main income aggregates: original, gross, disposable, and post-tax. For the purpose of this study, the definitions of income follow from those set out by Canberra Group (2011), with some variations adopted by Alves (2012). We define each of these aggregates below.

Original Income

We start by constructing the original income aggregate, also referred to as market income – income from both the labour and the capital markets. We define original income as the sum of income from employment, household production, investments and savings, net transfers between households and pension income. The reason for including pension income – old age and survivor benefits – in the original income is that otherwise many retirees would have incomes equal to zero. From an economical point of view, the methodological choice of excluding pension incomes would be unreasonable given how mature the Portuguese Social Security System is. We also included transfers paid between households in order to compute net transfers, since these transfers have a different nature from other transfers *i.e.*, cash benefits. According to Eurostat, the weight of pensions in total social cash benefits is around two-thirds in the European Union as a whole and 70 per cent in Portugal. In the EU-SILC dataset (2015), these benefits amount to approximately 80 per cent of the original income for 15 per cent of the households and they are equal to the original income for 10 per cent of the households.

Gross Income

The gross income variable is the original income plus any social benefits they receive in the form of cash payments relating to unemployment, illness or disability, family support, social exclusion, education, and others in the scope of housing (the latter are the only ones which are not strictly cash transfers).

Disposable Income

By definition, disposable income corresponds to the total amount of money that a person or household has available to spend or save after paying direct taxes and other required deductions, such as social security contributions. Therefore, our disposable income variable is computed as gross income less personal income taxes (PIT), social security contributions (SSC) and taxes on wealth. Even though in the EU-SILC database it is not possible to separate PIT from SSC, that is not a problem since both should be included in an overall tax progressivity analysis (Piketty and Saez, 2006). Additionally, the income taxes reported in the survey only encompass the taxes actually paid each year and do not take into account any tax refunds or adjustments related to the reference year. This is a limitation of the information in the database.

POST-TAX INCOME

Post-tax income, also referred to as take-home pay, is the disposable income less the household's indirect tax liabilities, *i.e*, the annual amount of VAT and excise duties paid by the households.

EQUIVALIZED INCOME TRANSFORMATION

We take the household as the income-receiving unit. In our analysis, we focus on individuals and assume that each person receives the equivalized income of the household to which they belong, based on the OECD modified equivalence scale. This equivalization takes into account differences in household size and composition. This is a conventional assumption in the literature to avoid an under- or overestimation of the degree of income inequality in a society.

3.2 Household's Indirect Tax Liabilities

As previously stated, the EU-SILC database provides all the variables needed to compute the original income, gross income and disposable income. In order to compute the post-tax income, we also need the annual amount of indirect taxes paid by households. We resort to the HBS to fill this gap.

In the HBS database, goods and services are aggregated according to the Classification of Individual Consumption by Purpose (COICOP), which is a system for classifying and aggregating individual consumption expenditure data. It is used by the European Union (EU) and other organizations to compare and analyse consumer spending patterns across countries and over time. COICOP is based on the International Classification of Individual Consumption According to Purpose (ICCOP), which was developed by the United Nations Statistics Division and consists of a hierarchical structure of categories and subcategories that covers all types of individual consumption expenditure, including goods and services purchased by households and non-profit institutions serving households. Our calculations use the COICOP disaggregated at 2-digits (see table A1).

As in the case of Araújo (2019), we use the model proposed by Decoster (2005) to match each category of goods and services in our sample to their corresponding indirect tax rates. Since we do not have information on prices, it is important to state that all calculations that follow assume that the prices of the goods and services do not vary, meaning that the prices set out by the producers are constant.

The following expression (1) gives the relationship between the consumer price of good i, q_i, and the producer price, p_i and shows that VAT is paid on both the producer price and the excise components.

$$q_i = (1+t_i)(p_i + a_i + v_i p_i), \tag{1}$$

where t_i corresponds to the VAT rate, a_i denotes the excise per unit, and v_i the excise expressed as a share of producer price, p_i.

However, our interest is on the household's indirect tax liabilities and these not only depend on VAT rates but also depend on the expenditure pattern, *i.e* the quantity purchased of good i, which we denote by x_i. Therefore, the tax liability on good i is given by:

$$T_i = (q_i - p_i)x_i. \tag{2}$$

Given that HBS does not observe producer prices but gives the expenditures of households, we substitute the producer price, p_i in equation (2) by its expression in (1):

$$p_i = \frac{q_i}{(1+t_i)(1+v_i)} - \frac{a_i}{(1+v_i)}. \tag{3}$$

Equation (3) is the producer price in terms of the consumer price. Plugging it into equation (2), we get an expression for the tax liability solely in terms of the parameters of the tax system and observable expenditures:

$$T_i^h = \left(\frac{t_i}{(1+t_i)(1+v_i)}\right) e_i^h + \left(\frac{v_i}{1+v_i}\right) e_i^h + \left(\frac{a_i}{1+v_i}\right) x_i^h, \tag{4}$$

where e represents the expenditure of household h on good i and is given by

$$e_i^h = q_i^h x_i^h. \tag{5}$$

Equation (4) was used to compute the household's indirect tax liabilities in the HBS survey. The first term refers to the VAT component and the second and third ones to the excise component.

The tax liabilities for the twelve commodity aggregates displayed in the first column of table A1 are based on the following approach: tax liabilities for commodity aggregate J (eg. Health) for each household h are given by:

$$T_J^h = \sum_{j \in J} T_j^h, \tag{6}$$

where j corresponds to the subsets of each commodity aggregate J. Finally, we obtain T_J as the result of the sum up T_J^h across households, giving the total indirect tax revenue collected on commodity aggregate J.

Limitations and Assumptions of the HBS dataset

The bottom-up technique we use has some limitations with respect to computing the amounts for all excise taxes at the individual level because there is no data on the quantity purchased. To

give an example, in the case of the tax on tobacco products (IT), we have assumed an average price of 5€ per pack of cigarettes.

Also, we do not have information on the tax on oil and energetic commodities (ISP) and vehicles-related taxes, such as *Imposto Único de Circulação* (IUC) and *Imposto Sobre Veículos* (ISV) – they both depend on the characteristics of the vehicle. In any case, our final values of share of excise duties in total fiscal revenues do not differ substantially from the ones of the National Accounts of INE (we estimate a share of 12% and the one from National Accounts is 16% in 2015).

3.3 Parametric Statistical Matching

In order to compute post-tax incomes we need to merge the information from EU-SILC and HBS databases. To this aim, we resort to parametric statistical matching, which is a statistical method used to join two or more datasets that have different variable structures but are related to the same underlying population. It involves creating a statistical model based on one dataset and using that model to impute missing values of a variable in the other dataset. The main advantage of parametric statistical matching is that it can be used to create a complete dataset from incomplete data, which is useful in our analysis. That being said, some assumptions about the data are required, such as that each dataset follows a particular probability distribution – a normal distribution or a binomial distribution – and the parameters of the distribution can be estimated from the available data.

3.3.1 Harmonization of Variables

The harmonization of variables is a crucial step to combine two different datasets. It involves identifying the variables in each dataset that represent the same concept or information and then ensuring that they have consistent names, formats, and values so that they can be matched together.

As underlined by Serafino and Tonkin (2017), there are several comparable variables in HBS and EU-SILC datasets. Our variables of interest are the ones used in our model as explanatory variables. All these variables capture most determinants of consumption and therefore households' tax liabilities – demographics, geography and income.

To ensure proper matching, two conditions must be met when selecting variables. Firstly, the variables must have similar distributions in both surveys. Secondly, the variables must be significant in explaining the variations in the target variable, *i.e* the household's indirect tax liabilities.

AGE OF THE REFERENCE PERSON

In the HBS the household reference person (HRP) is clearly defined and identified, however, in the EU-SILC there is no household reference person as such. In the literature, the HRP is usually

the person who owns the household accommodation, the eldest household, or the one with the higher income. We defined HRP according to the latter definition.

SUBJECTIVE RENTS

Since we are working with annual variables, $rents_h$ was transformed from a monthly variable to an annual one in EU-SILC.

3.3.2 The Model

The model used to match both datasets is the following:

$$T_h = \beta_1 + \beta_2 degurba_{h,1} + \beta_3 degurba_{h,2} + \beta_4 dimension_h + \beta_5 income_h + \beta_6 rents_h + e_h, \quad (7)$$

where $degurba_h$ is a categorical variable that represents the degree of urbanization of the area where households live. It has three categories: cities (densely populated areas), towns and suburbs (intermediate density areas), and rural areas (sparsely populated areas).
In regression analysis, categorical variables that lack rank orders require the use of dummy variables to be included. It is essential to remember that the values without their dummy variable form the reference category, and those dummy variables indicate the differences in relation to the reference category. In this sense, the indices $degurba_h$ identify each dummy variable, respectively.

Furthermore, $dimension_h$ corresponds to the household dimension, *i,e* the number of individuals living together in a single dwelling and sharing common resources such as food, utilities, and living expenses, regardless of their relationship or legal status.

Finally, $income_h$ is the household's disposable income for the reference year and $rents_h$ represents the household's subjective rent, which refers to the potential monthly market rent of owner-occupiers' unfurnished dwelling. [3] This variable works as a proxy for the household's real wealth.

Table 1 presents the estimated results for regression (7):

[3] Payments such as for electricity, heating etc., should be excluded.

Table 1. Regression Model: Household's Indirect Tax Liabilities function

2015	
	Dependent Variable: Indirect Taxes Paid
Constant	379.139***
	(10.67)
Degree of Urbanisation	
Intermediate	-135.046***
	(-4.45)
Sparsely populated	-122.998***
	(-4.14)
Household Dimension	195.0697***
	(14.55)
Disposable Income	0.065***
	(26.94)
Subjective Rents	0.146***
	(20.86)
Observations	28,982
R-squared	0.498

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

The variable $degurba_h$ is a categorical variable that captures the regionalisation effects. All the categories are significant. *Coeteris paribus*, as expected, the amount paid in indirect taxes is greater in the omitted category, *i.e* cities (densely populated areas), when compared to less urbanized and rural areas. Logically, on average, a larger household consumes more and therefore pays a more considerable amount of indirect taxes. Furthermore, we observe that the amount paid in indirect taxes increases as households become richer. Finally, we find that owners with more valuable properties spend slightly more than the ones with less valuable properties (lower subjective rent).

Overall, our model presents a good fit because our independent variables collectively explain around 50 per cent of the variance in the dependent variable - the annual amount of the household's tax liabilities.

Figure 1 compares the model-based estimated values of indirect taxes in the different deciles of the income distribution with the observed values in the HBS. We conclude that they are very close.

Equipped with equation (7) we can compute the level of indirect taxes paid by each household in the EU-SILC database. Figure 1 and table 2 show the results of this exercise, by income deciles.

For the whole income distribution, the average value of observed indirect taxes in HBS is the same as the estimated one, €2929.03. The estimated value in EU-SILC is very close to these: €2915.109.

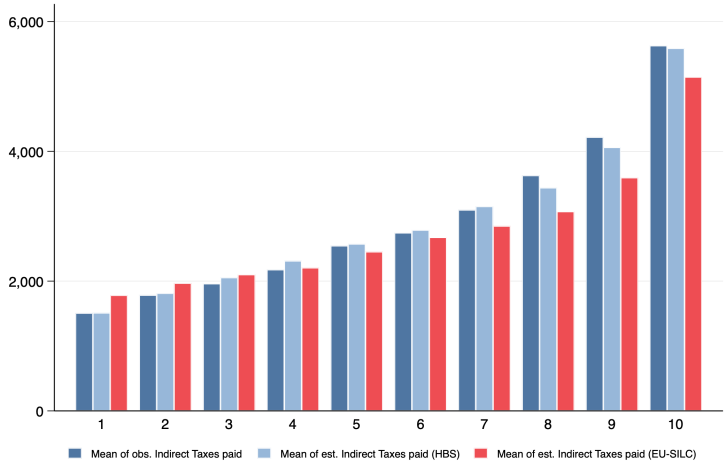


Figure 1. Comparison between the observed Indirect Taxes paid (IDEF) and estimated (IDEF and EU-SILC), 2015. Author's construction. Source: HBS and EU-SILC.

Table 2. Comparative Descriptive Statistics of Indirect Taxes by income decile, 2015

Equivalised Disposable Income Decile	Mean of Observed Indirect Taxes (HBS)	Mean of Estimated Indirect Taxes (HBS)	Mean of Estimated Indirect Taxes (EU-SILC)
1	1,507.130	1,511.057	1,782.767
2	1,784.387	1,813.004	1,968.407
3	1,961.310	2,054.917	2,101.066
4	2,176.743	2,312.813	2,206.045
5	2,544.997	2,573.904	2,452.635
6	2,744.753	2,786.777	2,675.110
7	3,097.527	3,151.152	2,849.553
8	3,628.396	3,439.171	3,072.620
9	4,218.907	4,061.877	3,594.964
10	5,629.057	5,588.261	5,145.488

Author's calculations. This table presents descriptive statistics of annual equivalised variables. Source: HBS and EU-SILC.

In section C of the Appendix we show the results of a model which includes *NUTS II*, *Age* and *Square of Disposable Income* variables. This model (equation (16)) would be presumably better, but it was not possible to implement due to three reasons. First, the variable *NUTS II* is absent in the EU-SILC for the years 2010 and 2015. Yet, this is an important variable to control for the different VAT rates existing in autonomous regions. Second, the variable *Square of Disposable Income* was not significant. And third, the age groups were not significantly different from each other. For these reasons, the three variables were not included in our model. From model (16) to model (7) we do not lose explanatory power, while all the regressors are significant and the coefficients are similar.

Also, by analysing figure 1, we verify that the observed mean value of indirect taxes paid in both HBS is close, on average, to its estimation in the same dataset. The corresponding estimated mean values (imputed) in EU-SILC follow the same pattern but for the low-income deciles we have an overestimation and for the high-income deciles an underestimation.

Limitation: Statistical Matching Problem

The accuracy of the model plays a crucial role in determining the reliability of parametric methods. There is a possibility of encountering the issue of regression towards the mean when using this approach. Our model assumes conditional independence (CIA), that is, the relationship between the variables being matched should not depend on any other variables. This assumption is important because it allows for the accurate matching of datasets and reduces the potential for bias in the resulting estimates.

There is a large body of literature on the uncertainty associated with statistical matching, particularly in parametric models. Kadane (2001), Moriarity and Scheuren (2001a), and Rässler (2002) are some examples. These studies have revealed that multivariate distributions are often not completely observed, which makes it challenging to identify the statistical model that generates the data. As a result, there is a degree of uncertainty surrounding the model, which can be quantified by calculating the range of an association parameter, such as the correlation coefficient in the normal bivariate case.

3.4 Descriptive Statistics of the Final Dataset

Table 3. Descriptive Statistics for 2015

Equivalised Variable	Mean	Observations	Households
Original Income	12,536.480	21,965	8,740
Cash Benefits	683.579	21,965	8,740
Gross Income	13,219.150	21,965	8,740
PIT + SSC	3,107.563	21,965	8,740
Wealth Taxes	115.632	21,965	8,740
Disposable Income	9,995.953	21,965	8,740
Indirect Taxes	1458.138	21,965	8,740
Post-Tax Income	8,537.814	21,965	8,740

Author's calculations. This table presents summary statistics of annual variables by the equivalent person. Original income has 678 negative observations that were transformed into 0.00001 to be taken into account in progressivity analysis using *progres* module. Source: HBS and EU-SILC.

Table 3 presents the summary and descriptive statistics for the variables of interest by the equivalent person of our final dataset. The total annual amount of direct taxes is given by the sum of PIT, SSC and wealth taxes, which is approximately €3,339. Furthermore, the total yearly amount of indirect taxes is €1,458. This means that the share of indirect taxes in total taxes is around 30%. Cash benefits represent around 14.6% of the total amount paid in taxes by households.

The mean annual gross income in 2014 (our reference year) is €13,219.15 and the annual disposable income is €9,995.953, which compares to €12,128 according to PORDATA. This underestimation of income is a standard feature in survey data.

In terms of average taxes, we find that the cash benefits rate is approximately 5.5%; and for direct [4] and indirect taxes the rates are approximately 24.4% and 15%, respectively. Indeed, the average income tax rate (as a percentage of gross income) recorded in 2015, added to the employees' SSC, was around 19.3%. [5]

[4] Note that we consider direct tax as the sum of PIT, SSC and Wealth taxes.
[5] Source: OECD Statistics.

4 Results for 2015

All the theoretical framework presented in this section is adapted from Gastwirth (1972) and Kakwani (1977a,b, 1984). In particular, Kakwani (1977b) proposes a suitable measure of tax progressivity that can be derived by comparing the Lorenz curve of income and the concentration curve of taxes.

This section is organized as follows. We start by describing the concepts of concentration and Lorenz curves and Gini index and measures of redistribution, followed by their respective theoretical frameworks. Those elements are crucial to understanding the progressivity and redistributive effects analyses we present afterwards.

4.1 Concentration Curves

Concentration curves are the extension and generalized version of Lorenz curves, which were originally used to analyse the distribution of income. Thus, Lorenz curves are just a particular case of such curves viz, the concentration curve for income.

Let $g(x)$ be a continuous function of x such that its first derivative exists and $g(x) \geq 0$, for $x \geq 0$. Assuming $E[g(x)]$ exists, then it can be defined as:

$$E[g(x)] = \int_0^{\infty} g(x) f(x)\, dx. \tag{8}$$

Thus,

$$F_1[g(x)] = \frac{1}{E[g(x)]} \int_0^{x} g(x) f(x)\, dx, \tag{9}$$

so that $F_1[g(x)]$ is monotonic increasing and $F_1[g(0)] = 0$ and $F_1[g(\infty)] = 1$. The concentration curve of $g(x)$ is the relationship between $F_1[g(x)]$ and $F(x)$. It can be shown that the Lorenz curve of income x is a special case of the concentration curve for $g(x)$ when $g(x) = x$.

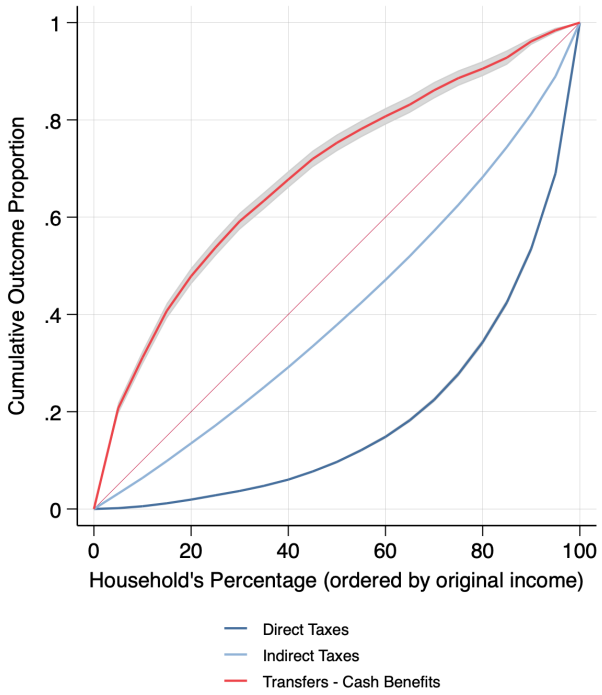


Figure 2. Concentration curves of equivalised taxes and benefits, ordered by equivalised original income, 2015. Author's calculations. Source: HBS and EU-SILC.

We use concentration curves to illustrate how transfers and direct and indirect taxes are distributed across families, ranking them by their equivalised original income (figure 2) or equivalised disposable income (figure 3). We observe that the concentration curve of transfers lies above the equality line, which means that transfers benefit low-income families. Approximately 60% of families in the bottom equivalised original income distribution receive 80% of the total amount of cash benefits.

On the other hand, in terms of tax paying, both the amounts paid in direct and indirect taxes seem to be skewed towards high-income families, which means the more income a family receives, the more taxes it pays. In the case of direct taxes, data shows that approximately 50% of families in the bottom equivalised original income distribution pay nearly 15% of the total amount. As for indirect taxes, that same 50% of families pay a considerably higher portion – almost 40% of the total amount (figure 2).

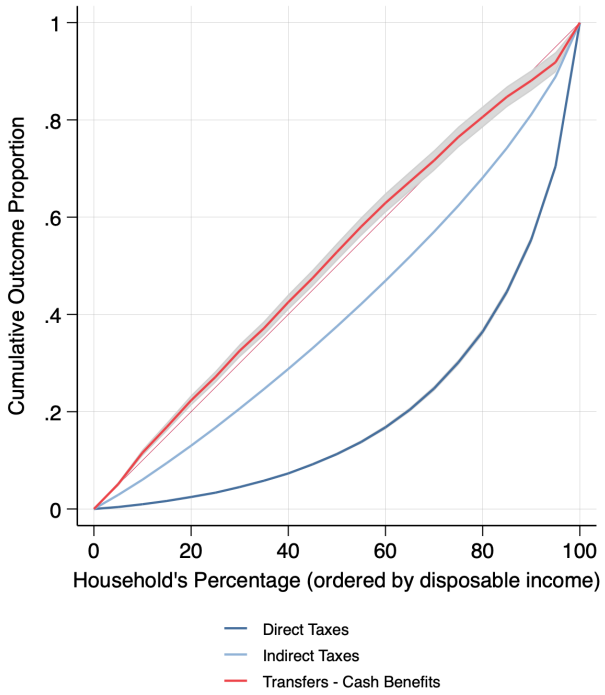


Figure 3. Concentration curves of equivalised taxes and benefits, ordered by equivalised disposable income, 2015. Author's calculations. Source: HBS and EU-SILC.

Ordering families by equivalised disposable income (figure 3), we now observe that the concentration curve of transfers lies close equality line, meaning that families get proportionally the same amount of cash benefits. Furthermore, the concentration curves for taxes are still skewed towards high-income families.

4.2 Lorenz Curves

The Lorenz curve is a graphical tool that provides complete information on the whole distribution of incomes relative to the mean. This tool has been conveniently used in the literature to represent the size distribution of income and wealth. In our case, we use it to assess income distribution.

Let the random variable x be the income of a given family, with a cumulative distribution function $F(x)$ *i.e.*, $F(x)$ represents the proportion of the population receiving incomes less than or equal to x.

$$F(x) = p = \int_0^x f(x)\,dx. \tag{10}$$

Considering $F^{-1}(t)$ as the inverse function of $F(x)$, and assuming that the mean μ of $F(x)$ exists, the Lorenz curve is defined as follows:

$$L(p) = \frac{\int_0^p F^{-1}(t)\,dt}{\int_0^1 F^{-1}(t)\,dt} = \frac{1}{\mu}\int_0^p F^{-1}(t)\,dt, \quad 0 \leq p \leq 1, \tag{11}$$

where the numerator $\int_0^p F^{-1}(t)\,dt$ sums the incomes of the bottom proportion, p (the poorest $100p\%$) of families. The denominator $\mu = \int_0^1 F^{-1}(t)\,dt$ sums all the incomes. $L(p)$ thus compares the cumulative proportion of total income received by a household arranged in ascending order, *i.e.*, from those with the least income to those with the most.

Simple summary measures of inequality can readily be obtained from the graph of a Lorenz curve. The share in the total income of the bottom p proportion of the population is given by $L(p)$; the greater that share, the more equal the distribution of income. Analogously, the share in the total income of the richest p proportion of the population is given by $1 - L(p)$; the greater that share, the more unequal the distribution of income.

Figure 4 shows the Lorenz curves of equivalised income variables for 2015.

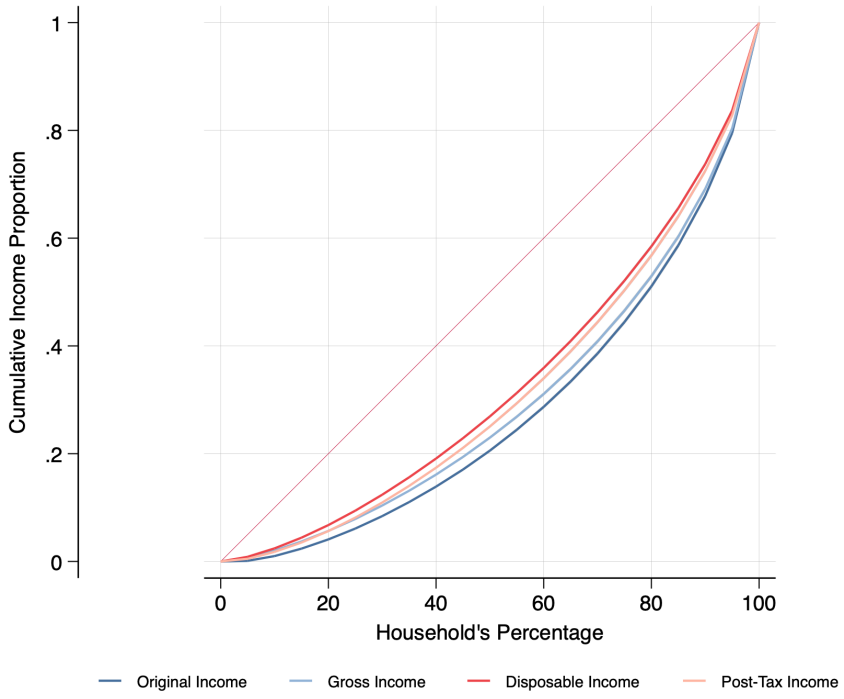


Figure 4. Lorenz Curves of equivalised income variables for 2015. Author's calculations. Source: HBS and EU-SILC.

As expected, we observe that the original income distribution is the most unequally distributed. Around half of the families receive nearly 25% of the total original income in 2015 in Portugal.

On the other hand, disposable income is the one that is closer to the equality line *i.e.*, it is the more equally distributed type of income. Approximately half of the families receive around 30% of the total disposable income in Portugal.

Additionally, figure 4 shows that indirect taxes are regressive given that the disposable income is more equally distributed across families than the post-tax income. Applying the same reasoning, both cash benefits and direct taxes are progressive since original income is less equally distributed than gross income and in turn, gross income is less equally distributed than disposable income.

Atkinson (1970) reintroduced the use of the Lorenz curve technique and demonstrated a connection between the Lorenz curve and a measure of social welfare – the social welfare function. The author demonstrated that the way income is distributed in a society, as represented by the Lorenz curve, can be used to determine the overall well-being of that society, regardless of how individuals' welfare is measured, as long as the Lorenz curves do not intersect. In this sense and vis-à-vis our data, we can conclude that from original income to disposable income distributions, society improves its overall well-being.

4.3 Gini Index and Measures of Redistribution

If everybody had the same income, the cumulative percentage of total income held by any bottom proportion p of families would also be p. The Lorenz curve would then be $L(p) = p$ *i.e.*, the share of the population and the share of total income would be the same. A proper enlightening content of a Lorenz curve is thus its distance, $p - L(p)$, from the line of perfect income equality - the egalitarian line. Compared to perfect equality, inequality removes a proportion $p - L(p)$ of total income from the bottom 100p% of the population. The larger the difference, the larger the inequality of income. The aggregate difference between the share of families and the shares of income across all families gives half of the Gini index (G), as formulated in equation (12).

$$G = 1 - 2\int_0^{\infty} p - L(p)\, dp. \tag{12}$$

The Gini coefficient thus measures the extent to which the distribution of income deviates from a perfectly equal distribution *i.e.*, the Gini index is equal to twice the area between the Lorenz curve and the egalitarian line. A coefficient of 0 expresses perfect equality, where every household has the same income, while a coefficient of 100 expresses full inequality where only one household has all the income.

The redistributive effect, R, of taxes and benefits can be summarized by the difference between the pre- and post-tax (transfer) Gini coefficients, as shown in equation (13):

$$R = G_{pre} - G_{post}. \tag{13}$$

CASH BENEFITS

The redistributive effect of cash benefits in 2015 is the difference between the Gini coefficients for original income ($G_o = 43.4\%$) and for gross income ($G_g = 40.4\%$), which means that inequality is reduced by 3 Gini percentage points (percentage points) from the original to gross income.

DIRECT TAXES

The redistributive effect of direct taxes (PIT, SSC and Wealth taxes) in 2015 is the difference between the Gini coefficients for gross income ($G_g = 40.4\%$) and for disposable income ($G_d = 34\%$), which means that inequality is reduced by 6.4 Gini percentage points from gross to disposable income. Caselli et al. (2016) use the EU-SILC dataset to compute several inequality measures for several EU countries and find that Portugal stands as a particularly high-inequality country both pre- and post-fisc – the EU Gini index pre-tax is 39% and after-tax it is 32%, whereas for Portugal it is 43% and 36%, respectively.

INDIRECT TAXES

The redistributive effect of indirect taxes (VAT and Excise Duties) is the difference between the Gini coefficients for disposable income ($G_d = 34\%$) and for post-tax income ($G_f = 36.6\%$), which means that inequality is increased by 2.6 Gini percentage points from disposable to post-tax income.

4.4 Overall Progressivity and Redistribution Effects

We use *progres* module for Stata, developed in Peichl and Van Kerm (2007), for calculations. This module does not include negative values when considering reference (pre-tax) incomes. However, the number of such negative values is small or zero, even when considering the original income as the (pre-fisc) income reference distribution (see notes on table 3).

In line with recent literature (see Hérault and Azpitarte, 2015; Verbist and Figari, 2014; Jenkins and Hérault, 2021), we take different reference distributions for different instruments when evaluating the separate effects: pre-benefit (original) income distribution for cash benefits; pre-fisc (gross) income distribution for taxes; and disposable income for indirect taxes. On the other hand, to evaluate combined effects accurately, we use original income as the reference distribution.

We employ the Kakwani Index (K) to measure progressivity, which is derived from the comparison between the Lorenz curve of income and the concentration curve of taxes. K is positive (negative) if the tax elasticity is greater (less) than unity for all incomes and it assumes a value of zero when the tax elasticity is one for all incomes. This means that a positive (negative) value of K for taxes indicates that taxes are progressive (regressive) and that the average tax rates increase

(decrease) with income. Conversely, a positive (negative) value of K for transfers (or benefits) suggests that benefits are regressive (progressive) and average benefit rates rise with income.

Kakwani showed that the redistribution effect can be decomposed in three elements, as described in equation (14). In this equation, the three elements are the average tax or benefit rate, t and b, respectively, the Kakwani Index, K, and the reranking measure, D. We adopt the usual practice in the literature and swap the two components on the right-hand side of equation (13) to ensure that progressive fiscal policy is consistently associated with a positive sign.

$$R_t = \left(\frac{t}{1-t}\right) K_t + D_{post}, \tag{14}$$

where K_t is the Kakwani index for taxes and D_{post} is a measure of the reranking of individuals between the pre- and post-fisc distributions and can be defined as the difference between the concentration [6] and Gini indices for post-fisc incomes.

The first term on the right-hand side of equation 14 is the vertical redistribution of taxes (V_t), which depends on the average tax rate (t) and on K_t that summarizes the distribution of taxes. When everyone pays the same proportion of their pre-fisc income, the vertical redistribution is zero. Yet, when taxes become more progressive, meaning the deviation from proportionality is greater, the vertical distribution increases. This can happen through either increasing the average tax rate or making the tax system more progressive, or both. Essentially, the more unequal the distribution of taxes, the greater the vertical redistribution.

Analogously, we can do the same decomposition for transfers or benefits:

$$R_b = \left(\frac{b}{1+b}\right) K_b + D_{post}, \tag{15}$$

where b is the average benefit rate.

[6] The Gini index is defined as $C = \sum_{i=1}^{n} \left(\frac{t_i}{t}\right) C_i$, where t_i is the average tax rate of the i^{th} tax and C_i is the concentration index of the i^{th} tax.

Table 4 summarizes our progressivity and redistribution results, as well as for each instrument:

Table 4. Tax-transfer progressivity, 2015

	2015			
	Redistribution Effect	**Average Tax or Benefit Rate**	**Kakwani Index**	**Reranking**
Overall Tax-Transfer System	-6.8	31.9%	0.1856	0.0177
Cash Benefits	-3	5.5%	-0.7923	0.0092
Direct Taxes	-6.4	24.4%	0.2083	0.0034
Indirect Taxes	2.6	14,9%	-0.1549	0.0003

Author's calculations. This table presents progressivity and redistribution results for different instruments.

Note: Cash benefits are the difference between gross income and original income. Direct taxes are the difference between disposable income and original income. Indirect taxes are the difference between post-tax income and disposable income. Redistribution Effect is in Gini percentage points (ppt).

CASH BENEFITS

The absolute value for the Kakwani progressivity index is 0.7923, meaning that cash benefits are progressive, which is in line with Jenkins and Hérault (2021), where the absolute value for cash benefits is never smaller than around 0.8. This shows that cash benefits are strongly targeted against income disparities (figure 5).

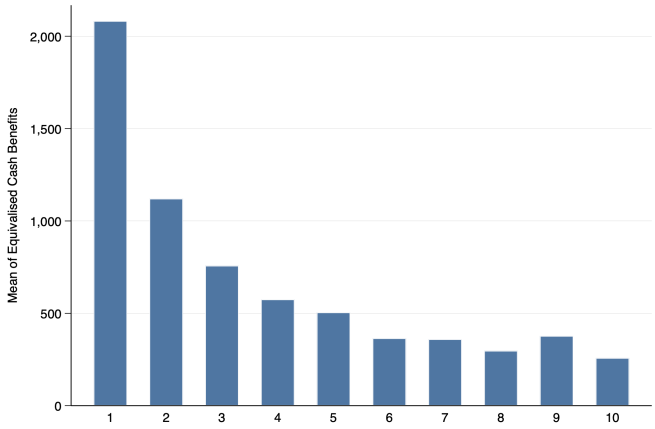


Figure 5. Mean of equivalised cash benefits by original income decile, 2015. Author's calculations. Source: EU-SILC.

In contrast, according to OECD (2019), cash social benefits are not always tightly targeted to the poorest. In 2016, on average, only less than 15% of cash benefits received by working-age individuals went to households in the bottom 20% of the income distribution, in Mediterranean European countries, including Portugal, where most benefits are related to past earnings. Indeed, these shares may vary across countries, since more than 40% of cash benefits go to the poorest 20% in Australia, Finland and New Zealand, countries with various income-tested benefits. In the OECD average, 23% of public cash transfers received by working-age individuals go to poorest quintile of the income distribution, while 19% go to households in the richest quintile of the income distribution.

Regarding average taxes, one should be mindful when comparing them, since different instruments use different income concepts for the income total. The average cash benefits rate is 5.35% (as a share of original income), a relatively lower amount when compared to Jenkins and Hérault (2021) (it ranges between 13% and 19%).

Direct Taxes

The Kakwani progressivity index is 0.2083, which means that direct taxes are progressive, corroborating the existing literature. By comparison, direct taxes are substantially less progressive than cash benefits. The average direct tax rate of 24.38% (as a share of gross income).

Figure 6 illustrates the average annual amount of direct taxes that households pay, by gross income deciles. The highest-income earners (10^{th} decile) pay more than the double than households in the immediate previous decile. According to Blanchet et al. (2022), from 1980 to 2017, in Western Europe, Germany is the country where the top 10 per cent share grew the most (9 percentage points), followed by Portugal and Italy (8 percentage points).

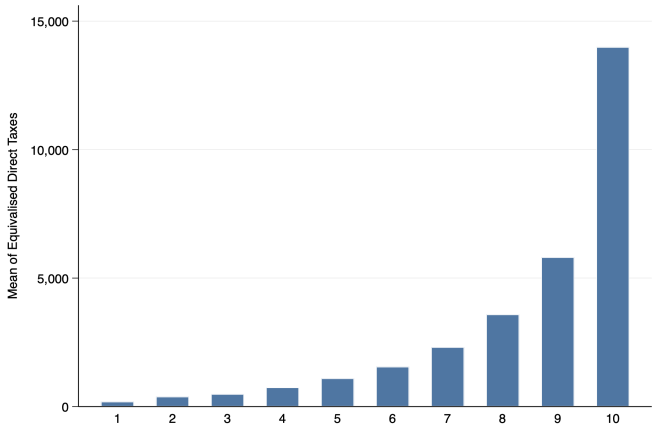


Figure 6. Mean of equivalised direct taxes by gross income decile, 2015. Author's construction. Source: EU-SILC.

Indirect Taxes

The Kakwani progressivity index for indirect taxes is negative (-0.1549), and the regressivity of indirect taxes is slightly lower than the progressivity of direct taxes. Araújo (2019), who estimated the regressivity of indirect taxes in Portugal, computed an absolute value of 0.1529 for the value of the Kakwani index of indirect taxes in 2015. The average indirect tax rate (as a share of disposable income) is around 15%.

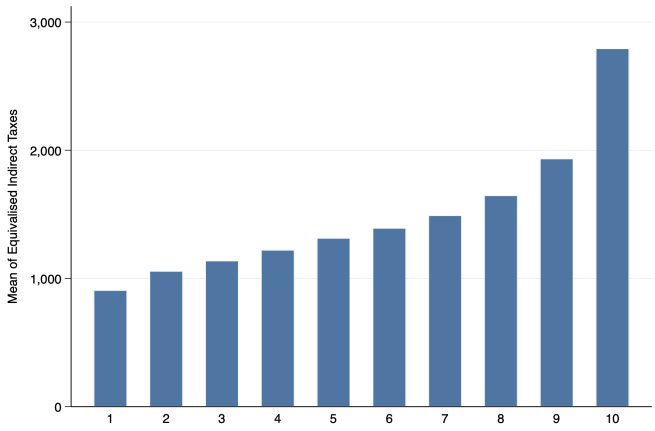


Figure 7. Mean of equivalised indirect taxes by disposable income decile, 2015. Author's calculations. Source: EU-SILC and HBS.

Similarly to the previous ones, figure 7 illustrates the average annual amount of indirect taxes paid by the households in 2015, by disposable income deciles.

Overall Progressivity

In terms of overall progressivity, we estimate that the Portuguese tax-transfer system in 2015 is progressive corresponding to a Kakwani index of 0.1856, using original income as our pre-fisc income distribution reference. The overall redistributive effects translate into a decrease of 6.8 Gini percentage points, from our reference income distribution to post-tax income distribution. The average net tax rate (*i.e.*, the average rate of the whole system that accounts for cash benefits, direct taxes and indirect taxes) is 31.9%.

5 Intertemporal Comparison of Results: 2010 and 2015

We performed an analogous analysis for the year 2010. The analysis can be found in section D of the Appendix and progressivity results are in table 5.

Table 5. Tax-transfer progressivity, 2010

	2010			
	Redistribution Effect	**Average Tax or Benefit Rate**	**Kakwani Index**	**Reranking**
Overall Tax-Transfer System	-5.9	26.4%	0.1964	0.0115
Cash Benefits	-3.9	5.8%	-0.8175	0.0058
Direct Taxes	-4.2	19,8%	0.1832	0.0029
Indirect Taxes	2.2	13.3%	-0.1442	0.0004

Author's calculations. This table presents progressivity and redistribution results for different instruments.

Note: Cash benefits are the difference between gross income and original income. Direct taxes are the difference between disposable income and original income. Indirect taxes are the difference between post-tax income and disposable income. The redistribution Effect is in Gini percentage points (ppt).

The cash benefits rate is roughly the same. However, the direct tax rate increased by almost 5 percentage points from 2010 to 2015. The mean gross income increased by around €200, but the mean value of direct taxes also increased in a considerably greater share (€523). This increase was a consequence of the decisions taken during the Economic and Financial Assistance Programme.

The average indirect tax rate presents a higher percentage in 2015 (14.9%) when compared to 2010 (13.3%). This increase is justified by the increase observed in VAT rates. [7]

In 2010, the Gini index is initially recorded at 42.1%. However, after accounting for cash benefits received and taxes paid, the index drops to 36.2%, indicating a reduction in inequality by 5.9 Gini percentage points. These findings suggest that from 2010 to 2015, the redistributive effects of social policies became more pronounced, due to a substantial increase in the redistributive power of direct taxes, in absolute terms (6.4 percentage points $>$ 4.2 percentage points). On the other hand, both cash benefits and indirect taxes contribute negatively to this trend. On the one hand, cash benefits have a smaller redistributive effect in 2015 (3 percentage points $<$ 3.9 percentage points); on the other hand, the evolution of indirect taxes contributes to increasing inequality from 2010 to 2015 (2.6 percentage points $>$ 2.2 percentage points).

[7] See table B1 of the Appendix.

Decomposing the effects as in equation (14), we conclude that the increase in redistribution between 2010 and 2015 was due to a sizeable increase in the net tax rate (from 26.4% to 31.9%), which more than compensated for the lower degree of progressivity of the overall tax-transfer system (with the Kakwani index declining from 0.1964 to 0.1856).

This overall outcome combines the following results for each instrument: (i) an increase in regressivity of indirect taxes, coupled with a higher indirect tax rate; (ii) a decrease in the progressivity of cash benefits, coupled with a decline in the average benefit rate; (iii) an increase in the progressivity of direct taxes, coupled with an increase in the respective tax rate. Thus, in 2015, direct taxes contributed more significantly to the overall redistributive effect, with cash benefits and indirect taxes basically compensating one another in terms of the overall redistribution impact.

6 Discussion and Conclusions

DISCUSSION

An advantage of conducting an overall analysis of the tax-transfer system is that it allows us to better understand the big picture: the progressivity and redistribution of each instrument and their respective interactions. Indeed, indirect taxes are attractive tax instruments due to their potential for generating significant government revenue and intrinsic self-enforcement capacity. Regardless, we observed that indirect taxes are regressive, which makes post-tax income more unequally distributed across families than disposable income, meaning, these taxes redistribute backwards. This is one of the most relevant criticisms made of indirect taxes: they treat unequal individuals (or families) equally. Both low- and high-income families pay exactly the same price – and, therefore, the same VAT amount – for the same good. This means that poor families spend a higher portion of their income, in comparison to high-income families. In this sense, the former will bear a higher burden of the amount paid in consumption taxes *vis-à-vis* their disposable income.

One can argue that reducing VAT rates is a solution to this unfair redistribution problem. This measure is widely seen as one that favours progressivity. Recently, we have witnessed its application in some European countries. Indeed, as a response to the COVID-19 Pandemic, Portugal reduced VAT rates on restaurants and gyms. Yet, few consumers have noticed reductions in the prices of meals or monthly fees immediately after this reform. On the contrary, prices have remained nearly the same, which means that firms increased their margins. Nonetheless, it can be arduous to discern the impact of VAT reductions during periods of significant market strain - as the one we are facing now - since prices comprise numerous determinants.

An alternative measure to correct for backward redistribution is to combine an increase in VAT with increases in transfers. Directly refunding the amount paid in indirect taxes by lower-income families would effectively help to correct indirect taxes' backward redistribution. Applying this measure proposes many challenges *inter-alia* for tax administrators since it can easily undermine VAT operations (Ebrill et al., 2001). However, in well-developed countries such as Portugal, the process would be rather simple: the consumer, when buying the goods, would just have to give their taxpayer number.

Correia (2010) shows that, for the United States, replacing the current direct taxes (PIT and wealth taxes) with a flat consumption tax, complemented by lump-sum transfers to every household improves efficiency and reduces inequality. The paper's findings suggest that policymakers should carefully consider the distributional implications of any changes to the consumption tax system. Specifically, the paper emphasizes the importance of designing tax systems that take into account the heterogeneous distribution of income and the potential regressive effects of consumption taxes.

CONCLUSION

The present research project proposes the assessment of the overall progressivity of the Portuguese tax-transfer system. We contribute to the existing literature through the use of parametric statistical matching methods to create a new database that imputes the annual indirect taxes paid by households from HBS into EU-SILC. To the best of our knowledge, this is the first attempt at doing so in Portugal. We compute the results for both 2010 and 2015.

The main conclusions are the following. First, the total redistributive effect of the tax-transfer system as a whole implies a decline in the Gini index from 43,4% to 36,4%. The global net tax rate is 31,9% and the system is clearly progressive. In 2015, the inequality reduction is mainly driven by direct taxes, corresponding to an inequality reduction of 6.4 Gini percentage points, while cash benefits reduce inequality by 3 Gini percentage points and indirect taxes redistribute backwards, contributing to an increase of 2.6 Gini percentage points.

Second, using the Kakwani Index, we also find that in 2015 indirect taxes are regressive and direct taxes and cash benefits are progressive. In absolute terms, the progressivity of direct taxes is greater than the regressivity of indirect taxes, thus direct taxes have a preeminent role in the progressivity of the Portuguese tax-transfer system. Cash benefits are even more progressive than direct taxes.

Third, from 2010 to 2015, the Portuguese tax-transfer system became marginally less progressive. The Kakwani progressivity value for the whole system is 0.1964 in 2010 and it is 0.1856 in 2015, taking the original income as our pre-fisc income distribution reference. Nonetheless, the overall redistribution effect is greater in 2015, given the increase in the global net tax rate.

The overall progressivity features follow the same pattern, however from 2010 to 2015 indirect taxes became more regressive, direct taxes became more progressive and cash benefits less progressive.

FUTURE RESEARCH

It would be interesting to implement our analysis in the more recent version of HBS (that is currently being conducted by Statistics Portugal), to contribute to the understanding of our current economic situation and its implications on Portuguese families in the aftermath of all the recent shocks that the economy is undergoing. Another way to improve this analysis would be to follow the example of Blanchet et al. (2022) who contribute to the distributional accounts literature with a new ”top-down” approach that combines several sources of data, besides household surveys, such as tax data and national accounts. These are important research ideas for the near future.

Bibliography

ALVES, N. (2012): "Uma Perspectiva Sobre a Redistribuição do Rendimento em Portugal e na União Europeia," *Boletim Económico, Banco de Portugal*, 44, 41–58.

ARAÚJO, A. (2019): "The Distributional Effects of Indirect Taxes in Portugal," Master's thesis, Nova School of Business and Economics.

ATKINSON, A. B. (1970): "On the measurement of inequality," *Journal of Economic Theory*, 2, 244–263.

BLANCHET, T., L. CHANCEL, AND A. GETHIN (2022): "Why Is Europe More Equal than the United States?" *American Economic Journal: Applied Economics*, 14, 480–518.

BRAZ, C. AND J. CORREIA DA CUNHA (2009): "The Redistributive Effects of VAT in Portugal," *Boletim Económico, Banco de Portugal*, 71–86.

CASELLI, F., M. CENTENO, A. NOVO, AND J. TAVARES (2016): "The Challenge of European Inequality," *After the Crisis. Reform, Recovery, and Growth in Europe*, 171–193.

CORREIA, I. (2010): "Consumption Taxes and Redistribution," *American Economic Review*, 100, 1673–94.

DALTON, H. (1922): *Principles of Public Finance*, Allied Publishers.

DECOSTER, A. (2005): "How progressive are indirect taxes in Russia?" *The Economics of Transition*, 13, 705–729.

EBRILL, M. L. P., M. M. KEEN, AND M. V. P. PERRY (2001): *The modern VAT*, International Monetary Fund.

FELLMAN, J. (1976): "The Effect of Transformations on Lorenz Curves," *Econometrica*, 44, 823–24.

GASTWIRTH, J. L. (1972): "The Estimation of the Lorenz Curve and Gini Index," *The Review of Economics and Statistics*, 54, 306–316.

HÉRAULT, N. AND F. AZPITARTE (2015): "Recent Trends in Income Redistribution in Australia: Can Changes in the Tax-Benefit System Account for the Decline in Redistribution?" *The Economic Record*, 91, 38–53.

JAKOBSSON, U. (1976): "On the measurement of the degree of progression," *Journal of Public Economics*, 5, 161–168.

JENKINS, S. AND N. HÉRAULT (2021): "Redistributive effect and the progressivity of taxes and benefits: evidence for the UK, 1977–2018," Working paper, Melbourne Institute Paper No. 23/21.

KADANE, J. B. (2001): "Some statistical problems in merging data files," *Journal of Official Statistics*, 17, 423.

KAKWANI, N. (1977a): "Applications of Lorenz curves in economic analysis," *Econometrica*, 45, 719–727.

——— (1977b): "Measurement of progressivity: an international comparison," *The Economic Journal*, 87, 71–80.

——— (1984): "On the measurement of tax progressivity and redistributive effect of taxes with applications to horizontal and vertical equity," *R. L. Basmann and G. F. Rhodes (eds), Advances in Econometrics*, 3, 149–168.

KIEFER, D. W. (1984): "Distributional Tax Progressivity Indexes," *National Tax Journal*, 37, 497–513.

MARTINS, A. (2016): "Progressivity in the Portuguese Personal Income Tax System," Master's thesis, Nova School of Business and Economics.

MATOS, M. (2017): "How Progressive are Portugal's Taxes?" Master's thesis, Nova School of Business and Economics.

MORIARITY, C. AND F. SCHEUREN (2001a): "Statistical matching: A paradigm for assessing the uncertainty in the procedure," *Journal of Official Statistics*, 17, 407–422.

MUSGRAVE, R. A. AND T. THIN (1948): "Income Tax Progression, 1929-48," *Journal of Political Economy*, 56.

OECD (2019): *Society at a Glance 2019*.

ON HOUSEHOLD INCOME STATISTICS, U. N. C. G. AND U. N. E. C. FOR EUROPE (2011): *Canberra Group Handbook on Household Income Statistics*, ECE/CES, UN.

PEICHL, A. AND P. VAN KERM (2007): "PROGRES: Stata module to measure distributive effects of an income tax," .

PIKETTY, T. AND E. SAEZ (2006): "The Evolution of Top Incomes: A Historical and International Perspective," *American Economic Review*, 96, 200–205.

PODGER, A., J. RAYMOND, AND W. JACKSON (1980): "The Relationship Between the Australian Social Security and Personal Income Taxation Systems - A Practical Examination," .

REIS, H. M. G. D. (2021): "Da gestão do IRS à progressividade fiscal : evidência empírica para Portugal," Ph.D. thesis.

REYNOLDS, M. AND E. SMOLENSKY (1977): *Public Expenditures, Taxes, and the Distribution of Income: The United States, 1950, 1961, 1970*, New York: Academic Press.

RITZEN, J. AND M. WOOLCOCK (2000): "Social cohesion, public policy, and economic growth: implications for countries in transition," .

RÄSSLER, S. (2002): *Statistical matching: a frequentist theory, practical applications, and alternative Bayesian Approaches*, vol. 168 of *Lecture Notes on Statistics*, Springer-Verlag.

SERAFINO, P. AND R. TONKIN (2017): *Statistical matching of European Union statistics on income and living conditions (EU-SILC) and the household budget survey*.

SUITS, D. B. (1977): "Measurement of Tax Progressivity," *The American Economic Review*, 67, 747–752.

VERBIST, G. AND F. FIGARI (2014): "The Redistributive Effect and Progressivity of Taxes Revisited: An International Comparison across the European Union," *FinanzArchiv: Public Finance Analysis*, 70, 405–429.

A COICOP Classification

Appendix Table A1. Commodity Aggregation by COICOP.

COICOP DIVISION	COICOP SUB-DIVISION
01. Food and non-alcoholic beverages	Food
	Non-Alcoholic beverages
02. Alcoholic beverages and tobacco	Alcohol
	Tobacco
03. Clothing And Footwear	Clothing
	Footwear
04. Housing, water, gas, electricity and other fuels	Actual rentals for housing
	Regular maintenance and repair of the dwelling
	Other services relating to the dwelling
	Electricity, gas and other fuels
05. Furnishings and household services	Furniture, furnishings and decorations, carpets and other floor coverings and repairs
	Household textiles
	Household appliances
	Glassware, tableware and household utensils
	Tools and equipment for house and garden
	Goods and services for routine household maintenance
06. Health	Medical products, appliances and equipment
	Outpatient services
	Hospital services
07. Transport	Purchase of vehicles
	Operation of personal transport equipment
	Transport services
08. Communications	Postal services
	Telephone and telefax equipment
	Telephone and telefax services
09. Recreation And Culture	Audio-visual, photographic and information processing equipment
	Other major durables for recreation and culture
	Other recreational items and equipment, gardens and pets
	Recreational and cultural services
	Newspapers, books and stationery
	Package holidays
10. Education	Early childhood education and elementary education
	Middle and secondary education
	Post-secondary education
	Higher education
	Other types of education
11. Restaurants And Hotels	Catering services
	Accommodation services
12. Miscellaneous Goods And Services	Personal care
	Personal effects n.e.c.
	Social protection
	Insurance
	Financial services n.e.c.
	Other services n.e.c.

Source: Eurostat.

B Changes in Value-added Tax (VAT)

Appendix Table B1. VAT Rates

		Mainland Portugal	A. R. of Azores	A. R. of Madeira
2009	**Reduced Rate**	5%	4%	4%
	Intermediate Rate	12%	8%	8%
	Standard Rate	20%	14%	14%
2014	**Reduced Rate**	6%	5%	5%
	Intermediate Rate	13%	10%	12%
	Standard Rate	23%	18%	22%

Source: Tax and Customs Authority, 2022.

C A more complete model (2015)

Regression (16) gives us a possible relationship between the household's indirect tax liabilities (T_h) and some determinant characteristics of the households:

$$\begin{aligned} T_h = \beta_1 + \beta_2 nutsII_{1,h} + \beta_3 nutsII_{2,h} + \beta_4 nutsII_{3,h} + \beta_5 nutsII_{4,h} + \\ + \beta_6 nutsII_{5,h} + \beta_7 nutsII_{6,h} + \beta_8 degurba_{1,h} + \beta_9 degurba_{2,h} + \beta_{10} age_h + \\ + \beta_{11} dimension_h + \beta_{12} income_h + \beta_{13} income_h^2 + \beta_{14} rents_h + e_h, \quad (16) \end{aligned}$$

where $nutsII_h$ is a categorical variable which represents NUTS II (Nomenclature of Territorial Units for Statistics) for Portugal, whose categories are *Norte*, *Centro*, *Área Metropolitana de Lisboa*, *Alentejo*, *Algarve* and the Autonomous Regions of *Açores* and *Madeira*. NUTS are an important tool with statistical information at a sub-national level, for highlighting specific regional and territorial aspects. It helps in analysing changing consumption patterns and the impact that fiscal policy decisions can have on our daily life (Eurostat - Data: regions and cities);

age_h denotes the age of the reference person in the Household, which we will explain later in this section. This variable assumes values from sixteen to eighty-five years old;

$income_h^2$ is the square of the household's disposable income. This variable was thought to be included for several reasons, *inter-alia* non-linearity between income and indirect taxes paid; diminishing marginal utility of income, *i,e* an additional euro of income has less impact on the indirect taxes paid at higher income levels; and *heteroscedasticity*, since the variance of the dependent variable is usually not constant across all levels of income.

Table C1 presents the regression results for A more complete model in 2015 (1) and for another model that does not include the categorical variable $nutsII_h$ (2):

Appendix Table C1. Ideal A more complete model: Household's Indirect Tax Liabilities function

	2015	
	Dependent Variable: Indirect Taxes Paid	
Explanatory Variables	(1)	(2)
Constant	981.5503***	946.892***
	(4.53)	(4.43)
NUTS II		
Centro	-93.468***	
	(-2.68)	
Área Metropolitana de Lisboa	-179.717***	
	(-5.13)	
Alentejo	-214.618***	
	(-5.81)	
Algarve	71.010*	
	(1.74)	
Açores	-900.513***	
	(-23.43)	
Madeira	-444.184***	
	(-10.88)	
Degree of Urbanisation		
Intermediate	-141.242***	-134.177***
	(-4.54)	(-4.43)
Sparsely populated	-92.587**	-115.515***
	(-2.91)	(-3.92)
Age of reference person		
20-24	-543.097**	-599.761***
	(-2.33)	(-2.59)
25-29	-534.403**	-572.181**
	(-2.35)	(-2.55)
30-34	-712.051***	-770.239***
	(-2.33)	(-3.65)
35-39	-577.827***	-649.671***
	(-2.70)	(-3.08)

Continued on next page

Appendix Table C1 – continued from previous page

	Dependent Variable: Indirect Taxes Paid	
Explanatory Variables	(1)	(2)
40-44	-676.410***	-745.881
	(-3.18)	(-3.56)
45-49	-501.471**	-566.881
	(-2.36)	(-2.70)
50-54	-523.170**	-585.270
	(-2.47)	(-2.80)
55-59	-549.778***	-612.804
	(-2.57)	(-2.90)
60-64	-667.831***	-724.820
	(-3.14)	(-3.45)
65-69	-387.850*	-443.590
	(-1.80)	(-2.08)
70-74	-593.715***	-659.387
	(-2.76)	(-3.11)
75-79	-447.671**	-503.706
	(-2.06)	(-2.34)
80-84	-358.379	-422.237
	(-1.59)	(-1.89)
85+	-534.056**	-602.145
	(-2.37)	(-2.71)
Household Dimension	179.570***	179.666***
	(12.57)	(12.68)
Disposable Income	0.073***	0.0734***
	(16.49)	(16.61)
$DisposableIncome^2$	-6.19e-08	-6.45e-08
	(-1.37)	(-1.43)
Subjective Rents	0.140***	0.135***
	(20.81)	(20.34)
Observations	28,982	28,982
R-squared	0.507	0.502

t-statistics in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

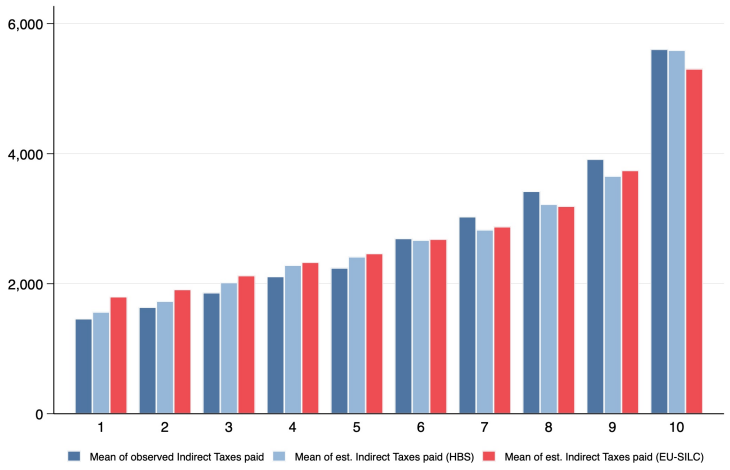
D 2010

D.1 Parametric Statistical Matching - Regression Results

Appendix Table D1. Regression Model: Household's Tax Liabilities function

2010	
	Dependent Variable: Indirect Taxes Paid
Constant	315.580***
	(7.77)
Degree of Urbanisation	
Intermediate	12.079
	(0.37)
Sparsely populated	-285.200***
	(-8.19)
Household Dimension	198.776***
	(14.14)
Disposable Income	0.060***
	(35.50)
Subjective Rents	0.148***
	(19.00)
Observations	24,323
R-squared	0.4752

t-statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.



Appendix Figure D1. Comparison between the observed Indirect Taxes paid (IDEF) and estimated (IDEF and EU-SILC), 2010. Author's construction. Source: HBS and EU-SILC.

Appendix Table D2. Comparative Descriptive Statistics of Indirect Taxes by income decile, 2010

Equivalised Disposable Income Decile	Mean of Observed Indirect Taxes (HBS)	Mean of Estimated Indirect Taxes (HBS)	Mean of Estimated Indirect Taxes (EU-SILC)
1	1,461.565	1,565.64	1,725.243
2	1,638.920	1,731.428	1,837.402
3	1,861.028	2,018.564	2,068.174
4	2,111.012	2,284.778	2,274.907
5	2,242.206	2,414.174	2,372.809
6	2,696.037	2,670.605	2,599.045
7	3,030.632	2,828.796	2,753.725
8	3,422.097	3,222.957	3,107.455
9	3,915.221	3,655.295	3,680.469
10	5,605.862	5,591.614	5,208.722

Author's calculations. This table presents descriptive statistics of annual equivalised variables. Source: HBS and EU-SILC.

D.2 Descriptive Statistics

Appendix Table D3. Descriptive Statistics for 2010

Equivalised Variable	Mean	Observations	Households
Original Income	12,314.710	13,368	5,182
Cash Benefits	714.871	13,368	5,182
Gross Income	13,029.580	13,368	5,182
PIT + SSC	2,519.461	13,368	5,182
Wealth Taxes	65.210	13,368	5,182
Disposable Income	10,444.910	13,368	5,182
Indirect Taxes	1,385.854	13,368	5,182
Post-Tax Income	9,059.052	13,368	5,182

Author's calculations. This table presents summary statistics of annual variables by the equivalent person. Original income has 317 negative observations that were transformed into 0.00001 to be taken into account in progressivity analysis using *progres* module. Source: HBS and EU-SILC.