Consumption Taxes and Income InequalityAn International Perspective with Microsimulation

Julien Blasco, Elvire Guillaud, Michaël Zemmour

LIS
Working Paper Series

No. 785

Consumption Taxes and Income Inequality
An International Perspective with Microsimulation

Julien Blasco, Elvire Guillaud, Michaël Zemmour

February 2020

CROSS-NATIONAL
DATA CENTER
in Luxembourg

Luxembourg Income Study (LIS), asbl

Consumption taxes and income inequality
An international perspective with microsimulation

Julien Blasco* Elvire Guillaud† Michaël Zemmour‡

February 27, 2020

Abstract

Consumption taxes are often considered as the most anti-redistributive component of the tax system. Yet, very few estimates, and fewer international comparisons of the redistributive impact of consumption taxes exist in the literature, due to scarce data on household expenditures. We use household budget surveys and microsimulation to provide consistent estimates of the regressivity of consumption taxes for a large panel of countries and years. We propose a new method for imputing consumption expenditure across countries, using widely available data on income and socio-demographic characteristics of households. We show that including the distribution of housing rents, when data is available, to impute households' consumption greatly improves the prediction of the model. Our results are threefold. First, there is a 1 to 2 ratio between the propensity to consume of the top decile (around 50% of their income) and that of the bottom decile (100% of income). Second, consumption taxes entail a significant rise in the Gini coefficient of income (between 0.01 and 0.04 point), yet of much smaller magnitude than the positive redistribution operated by direct taxes and transfers. Third, cross-country differences in the distributive effect of consumption taxes are mainly explained by variations in the tax rate (from 7 to 24% in our sample), rather than variations in the distribution of consumption, since everywhere the propensity to consume declines sharply with income.

Keywords: Indirect taxes; Redistributive Effect; Consumption; Income; Microsimulation; Luxembourg Income Study

---

This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" programme within the framework of the LIEPP centre of excellence (ANR-11-LABX-0091, ANR-11-IDEX-0005-02).

*INSEE, Université Cergy-Pontoise, THEMA, and Sciences Po, LIEPP, julien.blasco@insee.fr.

†University Paris 1 Panthéon-Sorbonne, CES, and Sciences Po, LIEPP, elvire.guillaud@univ-paris1.fr.

‡University Paris 1 Panthéon-Sorbonne, CES, and Sciences Po, LIEPP, michael.zemmour@univ-paris1.fr.

1 Introduction

In the comparative study of redistribution systems, the distributive effect of consumption taxes is the blind spot. Consumption taxes globally account for 30% of government revenue in developed economies, and evidence shows a positive cross-country correlation between the level of consumption taxes and the size of the welfare state (Kato, 2003). But consumption taxes are also considered an unfair tax, due to the fact that they are a flat tax applied on consumption expenditure, and that the share of income spent in consumption decreases with income (Warren 2008). The motivation of this paper is to measure the magnitude of the impact of consumption taxes on inequality, in order to assess the extent to which they counteract the redistribution effected by other socio-fiscal tools.

Very little research has been conducted on the subject, due to the scarcity of combined data on household consumption and household income. Therefore, existing studies either use micro-data but for a limited number of countries (for instance, Decoster et al. 2010 consider five European countries), or rely on aggregate imputations of consumption taxes (Garfinkel, Rainwater, and Smeeding, 2006). No study has undertaken micro-data analysis for 22 countries and 17 years, as we do here.

We propose a method to measure the magnitude of the distributional effect of consumption taxes in international comparison, even where microeconomic information on household consumption is not available. We apply it over a wide range of countries-years using readily available micro-data on household incomes. We calibrate our data to national accounts data, as is done in the studies conducted within the DINA project. But unlike these studies, we propose a simulation hypothesis that we validate beforehand on existing data, and thus avoid relying on an unfounded hypothesis of tax distribution that is proportional to income (Piketty and Saez 2007; Piketty, Saez, and Zucman 2018 on the United States). According to the WID guidelines, consumption proportional to income should be applied by default, and then possibly refined by applying theoretical savings rates. The problem with this method is that it has no empirical basis, and potentially produces biased results on the distribution of the consumption tax across the population.

The method we present is innovative and surprisingly simple. Starting from micro-data on expenditure and income provided by household budget surveys and gathered by the Luxembourg Income Study (LIS), we construct a simulation model of household consumption by income level, which allows us to obtain predictions of the distribution of the propensity to consume for all countries. Once this distribution is obtained, we apply a homothetic transformation of the data to match the macroeconomic aggregates. After testing the robustness of our model, we show that the composition of the household consumption basket is not a necessary piece of information to correctly impute the VAT actually paid by households. On the other hand, it is necessary to retrieve household income and composition, which are information systematically provided in international micro-data sets. When available, housing rents (including imputed rents) clearly increase the predictive power of the model. With this, the consumption and consumption tax data obtained are robust and allow accurate estimates of their distribution.

Using this method, we show that (1) the effective VAT rate relative to the standard of living of individuals in the bottom decile is twice as high as that of households in the richest decile; (2) the anti-redistributive effect of VAT is significant, and reduces the positive effect of direct redistribution (direct taxes and transfers excluding pensions) by one third on average in our sample; (3) to understand the (large) variation in the anti-redistributive effect among countries, the average consumption tax rate implemented within each country is the decisive explanatory factor.

2 Literature

As large as the redistribution made possible by consumption taxes revenue may be, the magnitude of their anti-redistributive effect still needs to be assessed. This is determined by the tax rate structure, households consumption patterns, and their average propensities to consume (Figari and Paulus, 2015). This distributive effect is likely to differ strongly from one country to another, since both the level of consumption taxes and the consumption behaviour of households vary a lot across countries (O'donoghue, Baldini, and Mantovani, 2004; Decoster et al., 2010; Savage, 2017).

2.1 Why should consumption taxes be regressive? Theoretical arguments

The amount of consumption tax paid by each household only depends on the goods and services the household has chosen to consume. Therefore, the distribution of tax rates across households will be determined by the share of their income spent by households –namely their propensities to consume, and their choices of goods and products, which will determine the rate applied on the total expenditure of the household.

It is widely acknowledged that the propensity to consume is decreasing with personal income. If this is true, then it means that for a given consumption tax rate, the amount of consumption tax paid by the household represents a decreasing share of the households' income. This is the main reason why consumption taxes are considered regressive.

For a given propensity to consume, the effective tax rate will then depend on the bundle of goods and services the household chooses to purchase. There is no clear evidence on the direction in which this "bundle effect" will affect the distribution of tax rates across levels of income.

Some empirical studies have been done on the subject, making use of detailed household budget surveys (Dauvergne, 2012; Boutchenik, 2015). In the case of France, evidence shows that the bundle effect is not correlated with income: the mean effective tax rate on consumption is almost constant. Indeed, the main reason why the average consumption tax rate is lower for lowest income households compared to high income ones is the share of expenditure that is allocated to housing rentals, which are not subject to VAT. Since low income households tend to be less home owner than high income households, and since their effort rate on housing (the share of their income dedicated to housing expenditure) is larger, the non-taxation of rent as a progressive effect on taxation. If we take only non housing consumption, it appears that the consumption tax rate paid by households is stable among the different deciles of income.

In the present study, housing rentals are removed from the taxable consumption when data is available. Therefore, one averaged effective tax rate is applied on all taxable consumption. Indeed, results from the literature underline that the bundle effect is clearly a third order effect, after the decreasing propensity to consume and the share of the rent (Figari, Paulus, and Sutherland (2015), Decoster et al. (2010)).

2.2 Issues in measuring consumption taxes

Unlike payroll or income taxes that can be measured at the individual level with administrative data, consumption taxes such as sales taxes or value-added taxes are generally not registered at the individual level. Therefore, it is not straightforward to count how much consumption taxes a household has paid. The most common way to measure this is using consumption data and microsimulation techniques. Indeed, with information on the household's consumption and the tax system of the country, one can derive the amount of consumption taxes paid by the household.

Three main issues arise with this technique: the first one is the design of the tax rate on consumption that has to be applied to the said consumption expenditure. Second, it is useful to ask, in the context of comparing the redistribution of fiscal systems in a cross-country fashion, whether micro-data from different national surveys can be compared directly or if they have to harmonized with national accounts. Lastly, as consumption data are costly to gather and can be missing for some countries, one can ask if imputation methods can be used.

Measuring the tax rate on goods and services. Two competing strategies exist in order to compute effective tax rates on consumption on a cross-country perspective. The first one consists in using legal statutory tax rates. This method presents the benefit of being an exact method, provided that we have a decomposition of the household's bundle of goods, so that we can apply the right tax to each good purchased. This method is unfeasible in practice, as it requires to go through the legislation of every country in the study, for every year of interest. Moreover, in order to apply different tax rates according to the nature of every good or service purchased, one would need the consumption data to be broken down into fine categories. Databases on consumption rarely match this level of precision.

The second method is the computation of implicit tax rates These are computed through national accounts data on households consumption and tax revenue, and yield averaged tax rates for every country-year. For each country-year, we compute the ratio

$$\tau = \frac{consumption\ taxes\ revenue}{taxable\ consumption}$$

which is the effective tax rate on consumption paid by households.

With this method, since the implicit tax rate is aggregated over all types of goods and services, the "bundle effect" cannot be simulated. Based on the discussion in subsection 2.1, we make the assumption that this effect is of second order when compared to the effect of differences in propensities to consume among different levels of income.

The need for recalibration for international micro-data comparisons. There is always a gap between micro-data from surveys and aggregated data from national accounts. In this case, as we use individual income and consumption data in order to estimate the impact of consumption taxes, we want to make sure that the amounts can be compared from one country to another. National accounts, as they are standardized, are more fit for international comparisons. Indeed, propensities to consume computed at the national level vary significantly between countries, as measured with national accounts. These differences, however, do not always appear in micro data.

Therefore, we choose to combine micro and macro-level data in order to produce distributed information on income and consumption that are comparable at the international level.

Imputation of consumption. Data on consumption is more rare than data on income, and they are even more rarely combined in micro datasets. In order to produce estimates of the distributive impact of consumption taxes for a broad range of countries, we develop an imputation model for consumption expenditure, based on households' income and socio-economic characteristics.

Our database contains information on different countries for multiple periods of time. It may be the case that, for some countries, data on consumption be available for some specific years only. For some countries, however, no annual dataset includes consumption data. The challenge is then to design a model that can be calibrated on some datasets where consumption is available, and used to impute consumption on some datasets where consumption data is entirely missing. This model must ben independent on the country or the year it is applied on.

3 Methodology and data

Starting from cross-country micro-level databases on income and consumption, we estimate the amount of consumption taxes paid by each household. This allows to define the tax-to-income ratio of a household as the ratio of consumption taxes paid to their income. In this section, we present the methodology and data used in order to produce consistent estimates of the distribution of tax-to-income ratios in different country-years.

3.1 Data

We use micro-level data from surveys on income and consumption in order to compute households' propensities to consume. Implicit tax rates are computed via national accounts data on consumption tax revenue and consumption. National accounts data on consumption and income are also used in order to scale micro data.

3.1.1 Household surveys

The Luxembourg Income Study (LIS) is a cross-national data center that gathers survey and administrative data on households' income, wealth, consumption, and other socio-economic characteristics. Our dataset includes more than 200 country-years, spanning more than 30 countries and years 1967 to 2016. The country coverage includes Austria, Belgium, Denmark, Estonia, Finland, France, Georgia, Germany, Greece, Hungary, Iceland, Ireland, Italy, Lithuania, Luxembourg, the Netherlands, Norway, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, and the United Kingdom.

In most countries, this data comes from national household budget surveys that have been conducted by national statistics institutes. This data is then collected by the Luxembourg Income Study in order to be harmonized[1].

Observed data for household consumption is only available for 77 datasets, spread across 23 countries, going from 1973 to 2013 (see appendix D). When it is not available, we use an imputation method, as described in section 3.2.3. The datasets that are used in order to estimate the imputation model are that of France, Germany, Italy, Spain, United Kingdom, Poland, Switzerland, South Korea, Estonia, Slovenia, Taiwan. Imputed consumption is thus used to produce estimations on those countries as well as Australia, Austria, Belgium, Brazil, Colombia, Czech Republic, Denmark, Finland, Greece, Hungary, Iceland, Ireland, Mexico, the Netherlands, Norway, Russia, Slovenia, South Africa, Sweden, Switzerland and the United States.

3.1.2 National accounts

Survey data is complemented by national accounts data for each country-year. They will be used for two main purposes: the first one is to scale micro-data on consumption and expenditure for it to be consistent with national accounts, and thus comparable from one country to another. The second use of national accounts data is the computation of consumption tax rates, based on tax revenue and total consumption of households.

These data, collected on OECD Statistics[2], are produced by national statistics institutes. They cover not only OECD members but also a number of other cooperating countries.

[1] For more information, see http://www.lisdatacenter.org/about-lis/.

[2] Data available at http://stats.oecd.org

3.2 Methodology

We use microsimulation to produce estimates of consumption taxes paid by households: this requires information on households' consumption expenditures as well as on taxes on consumption. Having computed the amount of consumption taxes paid, we can define the tax-to-income ratio for a household i in country c at year y:

$$TIR_{i,cy} = \frac{\tau_{cy} \cdot consumption_i}{disposable\ income_i} = \tau_{cy} \cdot prop_i \tag{1}$$

where τ_{cy} is the effective tax rate on consumption in country c at year y and $prop_i$ is the propensity to consume of household i, that is the share of disposable income which is actually spent.

The distribution of this measure relatively to income determines the regressivity of consumption taxes in the country-year. The more decreasing the TIR is with income, the more regressive the consumption tax.

Similarly, we can define a post-tax disposable income as a measure of disposable income once consumption taxes have been accounted for:

$$\text{post-tax income}_i = (1 - TIR_{i,cy}) \cdot \text{disposable income}_i \tag{2}$$

In this section, we present in 3.2.1 the method used to compute propensities to consume and to make them consistent across country-years, then we define in 3.2.2 the implicit tax rates on consumption used to simulate consumption taxes. Eventually, we describe the imputation strategy when consumption data is missing in 3.2.3.

Similarly, we define aggregate propensities to consume for each country c at year y, based on aggregate values of consumption and income in national accounts:

$$P_{c,y} = \frac{C_{c,y}}{I_{c,y}}$$

3.2.1 Definition of propensity to consume

Households' propensities to consume, which are the household-level term in eq. (1), represent the share of income that is effectively spent on goods and services. It is computed at the household level with budget survey data and is defined as the ratio of taxable consumption over disposable income. Thus, for every household i:

$$prop_i = \frac{taxable\ consumption_i}{disposable\ income_i}$$

Taxable consumption includes all monetary expenditure, excluding rents, which are not subject to consumption taxes[3]. Moreover, loan repayments are considered as savings, not

[3]Rents represent a higher share of income at the lower end of the income distribution. Therefore, failing to subtract them from taxable consumption, as it is done in most international comparisons, produces a slight overestimation of the regressive effect of consumption taxes (see appendix B.2).

consumption. Disposable income is the amount of money that households have available for spending after direct taxes, social security contributions, and monetary transfers have been accounted for.

Similarly, we define an average propensity to consume for each country c at year y, based on aggregate values of consumption and income in national accounts:

$$P_{c,y} = \frac{C_{c,y}}{I_{c,y}}$$

In order to make the distribution of propensities to consume consistent with national accounts, micro-data on consumption and income is thus scaled according to these aggregates. After this homothetic transformation, we then have:

$$\sum_{households\ i} taxable\ consumption_i = C_{c,y}$$
$$\sum_{households\ i} disposable\ income_i = I_{c,y}$$

The combinations of both kinds of data allows to use micro-data to estimate the distribution of consumption with income, while the relative average levels of income and consumption are put in accordance with national accounts[4]. See appendix A.1 for details on the economic aggregates used for the scaling.

3.2.2 Implicit effective tax rates

Consumption taxes rate include value-added-taxes (VAT), excise taxes, and other taxes on goods and services. In order to account for all of these taxes and to reflect their average respective weights in consumption, we compute an implicit tax rate based on national tax revenue and domestic consumption. That is, we do not rely on statutory rates but rather on tax revenue and consumption data from national accounts.

For each country-year, we compute the ratio between consumption taxes revenue and the amount of taxable consumption. This defines the average effective tax rate on consumption paid by households. For each country c at year y, the effective tax rate on consumption is

$$\tau_{c,y} = \frac{consumption\ taxes\ revenue}{domestic\ monetary\ consumption}$$

This definition of implicit tax rates is based on previous research (see Mendoza, Razin, and Tesar (1994) and O'donoghue, Baldini, and Mantovani (2004)), with small improvements regarding the evaluation of domestic monetary consumption: we take into account the fact that housing rentals and some part of public consumption do not generate consumption tax revenue. See appendix B.1 for details on the method of computation of the tax rates.

[4]The method of homothetic scaling relies on the assumption that the undereporting of consumption is at first order, general phenomenon, independent of individual characteristics.

This implicit tax rate, which averages all the different rates on specific products, is thus the same for all households of a country-year. Based on the discussion in section 2.1, we argue that the effect of different bundles of goods and intermediate VAT rates is of second order compared to the decreasing propensity to consume and decreasing share of rent in income (see appendix B.1 for robustness checks on the effect of different effective tax rates along the income distribution).

3.2.3 Imputation of consumption

We design a regression model in order to impute consumption data according to households characteristics, when consumption data is not available in the dataset. The key issue here is that the distribution of consumption data will have to be imputed on whole datasets for which there is no observation on expenditure regarding the specific country-year. Therefore, it is necessary to design a model which is generic enough to be calibrated on some countries and used on others.

In order to achieve that, we apply medianization to all monetary variables (including income, consumption and value of housing). That is, all monetary variables are expressed in proportion of their median values in the country c at year y. The **medianized disposable income** of household i is thus:

$$\widehat{income}_i = \frac{income_i}{\text{median}_{c,y}(income_i)}$$

The underlying assumption is that the relations between medianized consumption and medianized income are similar in every country, account taken of other individuals determinants. We apply a generalized linear model with logarithmic link, where medianized consumption depends on the medianized disposable income of the household, the medianized value of its housing and other socio-demographic variables X_i.

$$\widehat{income}_i + \widehat{housing}_i + X_i \;\rightarrow\; \widehat{consumption}_i \tag{3}$$

The value of housing is defined as the total cost of housing, including rents and utilities, as well as imputed rents for occupying owners. Socio-demographic determinants include the number of members of household, occupancy status of household, etc. The cost of housing is a relatively common information in income databases (we have it for 60% of our datasets), and this is a very good proxy of the standard of living in addition with income[5]. While we use this independent variable in this article, models without housing expenditure yield quite satisfactory results as well (see appendix B.4). For detailed explanations of the model, see appendix A.3.

With this regression model, we impute medianized values of the households' monetary consumption. They are then scaled with national accounts data in order to be comparable with observed values, according to section 3.2.1.

[5]Moreover, while income can be subject to important transitory shocks that do not entail similar shocks in consumption, housing expenditure is smoothed, as consumption is expected to be

4 Results

4.1 Accuracy of imputation of consumption

The model we have defined is used to impute consumption on every dataset where income and other relevant socio-economic determinants are available. We find that the curves of propensity to consume related to income obtained with imputed consumption data are very close to that on actual observations of consumption data, for various countries and years. *In fine*, the same model applies to a wide range of country-years, and the shape and downward slopes of the curves are well adjusted (see fig. 1). In this figure, each pane is done in cross-validation: the nine panes present results from nine imputation models which were estimated with pooled data that were excluding respectively Australia, Estonia, France, etc. For example, the imputed consumption compared to observed consumption in Australia 2010 has been obtained with a regression model that was calibrated on every country but Australia.

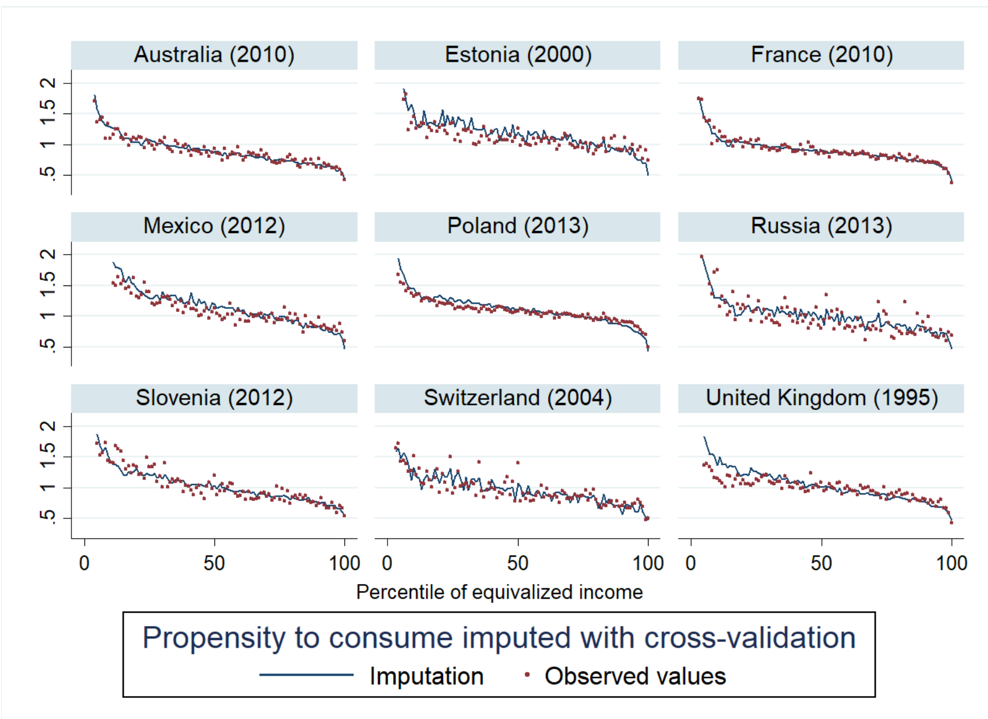

Figure 1: Estimated and actual propensities to consume (using cross-validation[6])

We show that an innovation of the model, adding the total cost of housing[7] as an independent variable greatly increases the quality of the imputation model. On average, a

[6]For this figure, the imputation model has been estimated 9 times, each time removing from the training sample the country the model is compared to. For example, the imputed consumption compared to observed consumption in Australia 2010 has been obtained with a regression model that was calibrated on every country but Australia.

[7]This includes monetary and non-monetary expenditure (e.g. imputed rents).

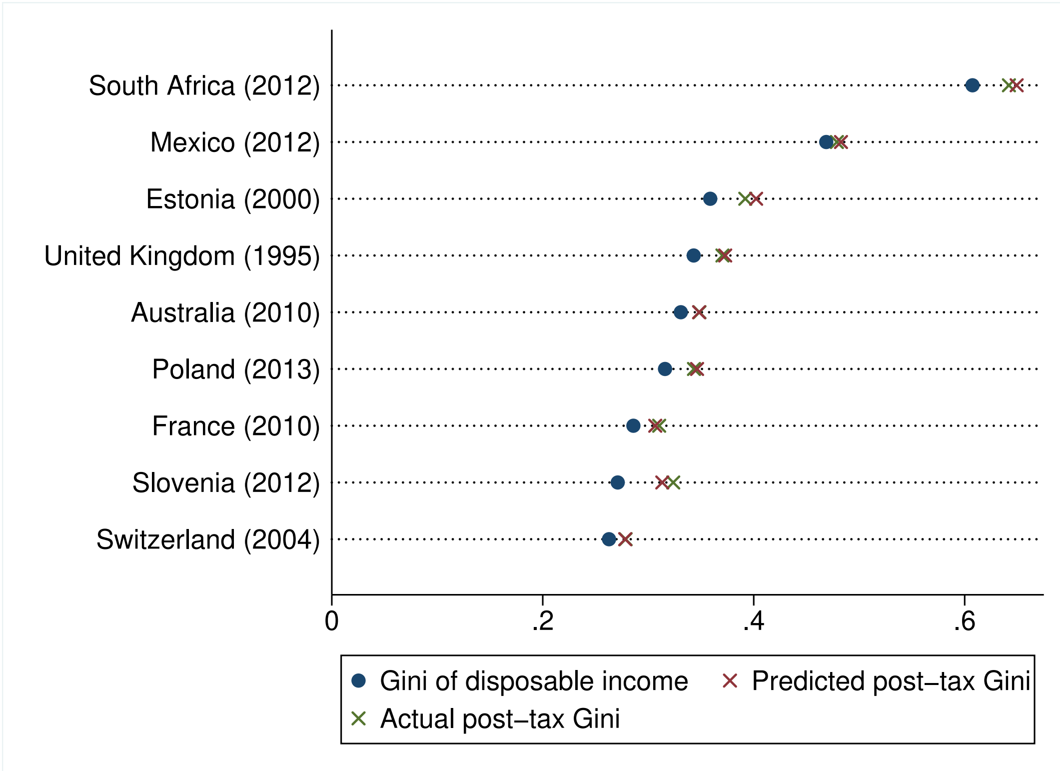

Figure 2: Actual and predicted Gini coefficient of post-consumption-tax income

model that includes this independent variable explains a 33% higher share of the variance in consumption than a model based on the sole income (see appendix B.4). The expenditure in housing of the household seems to bring valuable information about its total monetary consumption.

In the following parts, unless stated otherwise, results using our imputation model will be used in international comparisons (based on (mod. 2) specification from appendix A.3).

Likewise, the imputation model produces reliable estimates of post-tax Gini coefficients (see fig. 2).

4.2 Results on consumption taxes

4.2.1 The tax-to-income ratio of the 10% richest is less than half of that of the 10% less affluent

The first results that we present here are the global tax-to-income ratios for each percentile of income. We observe that they are decreasing with income in every country-year, and that our imputation model predicts relatively well the evolution of consumption relatively to income.

We find that in all countries and years, propensities to consume are decreasing with the percentile of income. In general, the amount of consumption is higher than the disposable income for the first percentiles, indicating significant dissavings. On the contrary, households in the highest percentiles of equivalized income consume about 50-60% of their income.

Tax-to-income ratios follow the same downward trends (fig. 3). Consumption taxes are therefore significantly regressive: in France in 2010, poorest households spend more than 20% than their disposable income in consumption taxes, while the tax-to-income ratio is lower than 10% for richest households. Germany presents a very similar curve, both in level and in slope. The estimated regressivity is slightly lower in France, as consumption propensity curves cross at the middle of the income distribution. The slope is similar for other countries, even if the levels are different: in Denmark, the tax-to-income ratio is higher than 30% for half of households, while it is lower than 10% for the majority of households in the United States.

In most countries, the tax-to-income ratio of the 10% richest is less than half of that of the 10% less affluent.

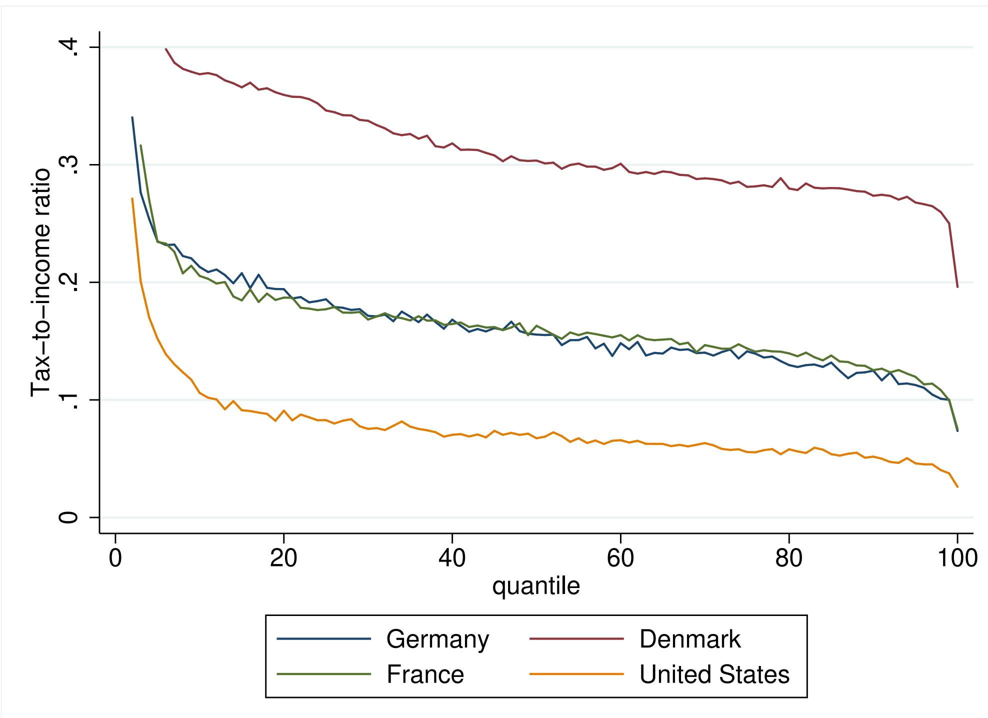

Figure 3: Tax-to-income ratio. Years for Germany, Denmark, France, United States respectively 2013, 2004, 2010 and 2013

4.2.2 The anti-redistributive effect of consumption taxes is between 1/5 and 1/3 of direct redistribution

In this section, we measure the distributive effect of consumption taxes with synthetic measures of income inequality and progressivity. First, we show that the effect of consumption taxes on income inequality is significant, but of much lower magnitude than direct socio-fiscal redistribution. Second, we decompose this distributive effect into one horizontal and one vertical component, and show that the largest part of between-country differences are due to the differences in consumption tax rates.

By defining a post-tax disposable income, we can measure the distributive impact of consumption taxes. We compare in fig. 4 the Gini coefficients for three concepts of income: factor income, which is income from labour and capital; disposable income, which is factor income after redistribution operated by direct taxes and transfer; and post-tax income, which is the disposable income of which we subtract the amount of consumption taxes.

$$\begin{aligned}\text{post-tax income} &= \text{disposable income} - \text{consumption taxes}\\ &= \text{factor income} + \text{transfers} - \text{direct taxes} - \text{consumption taxes}\end{aligned}$$

Inequality is lower for post-tax income than for factor income, and higher than for disposable income. Indeed, consumption taxes entail a rise in inequality, but not close to the reduction performed by the rest of the socio-fiscal system. For all countries [8], the Gini coefficient of post-tax income is much closer to that of disposable income than that of factor income.

We define the redistributive effect of consumption taxes as the difference in income inequality between the disposable income distribution and the post-tax disposable income distribution. We thus use the following index of effective redistribution:

$$\Delta G = G_{dhi} - G_{dhi-tax}$$

where G_{dhi} (resp. $G_{dhi-tax}$) is the Gini index of the pre-tax (resp. post-tax) disposable income. This measure is positive for a progressive tax, and negative for a regressive tax. For consumption taxes, we are expecting negative values, meaning an increase in the level of income inequality.

Figure 5 shows that the anti-redistributive effect lies between 0.01 and 0.05 Gini points, while most countries staying in a range between 0.015 and 0.035.

The anti-redistributive effect is thus significant, and is large enough to change the ranking in terms of income inequality between countries which present similar levels of disposable income inequality, but different distributions of propensities to consume and different consumption tax rate. For example, the United States have higher disposable income inequality than Greece, but post-tax income inequality is higher in Greece than in the USA (fig. 4). This is mainly due to the high VAT rate in Greece (24%), while sales taxes in the USA are much lower (about 7%).

[8] except for South Africa, which has both low redistribution through taxes and transfers, and highly regressive consumption taxes

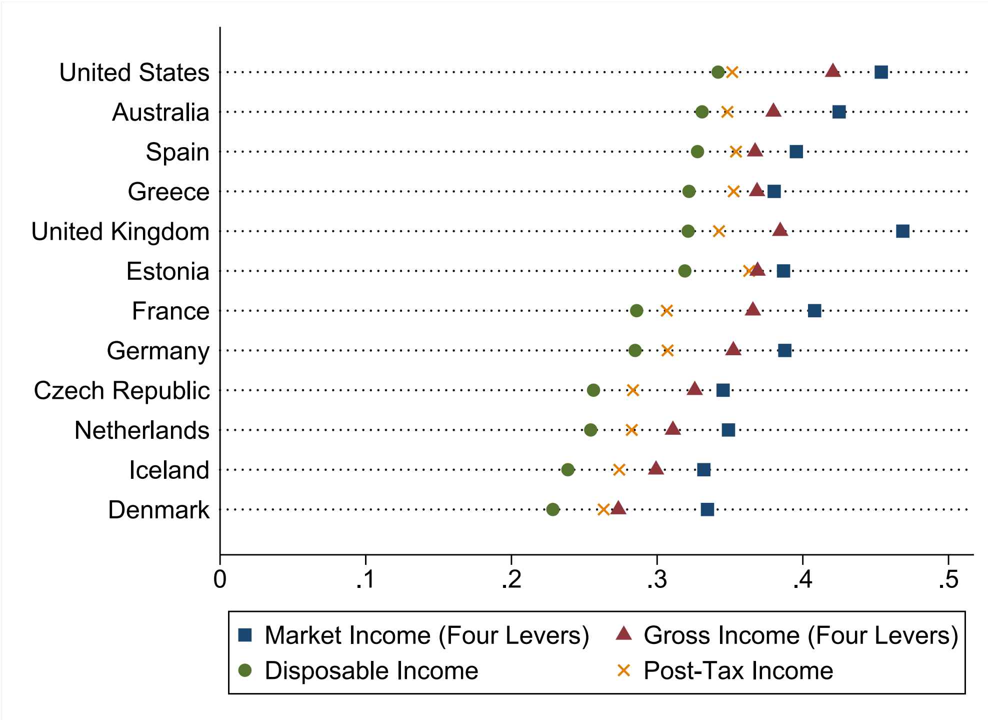

Figure 4: Gini coefficients of market, gross, disposable and post-tax income. Market income and gross income are from Guillaud, Olckers, and Zemmour (2019)

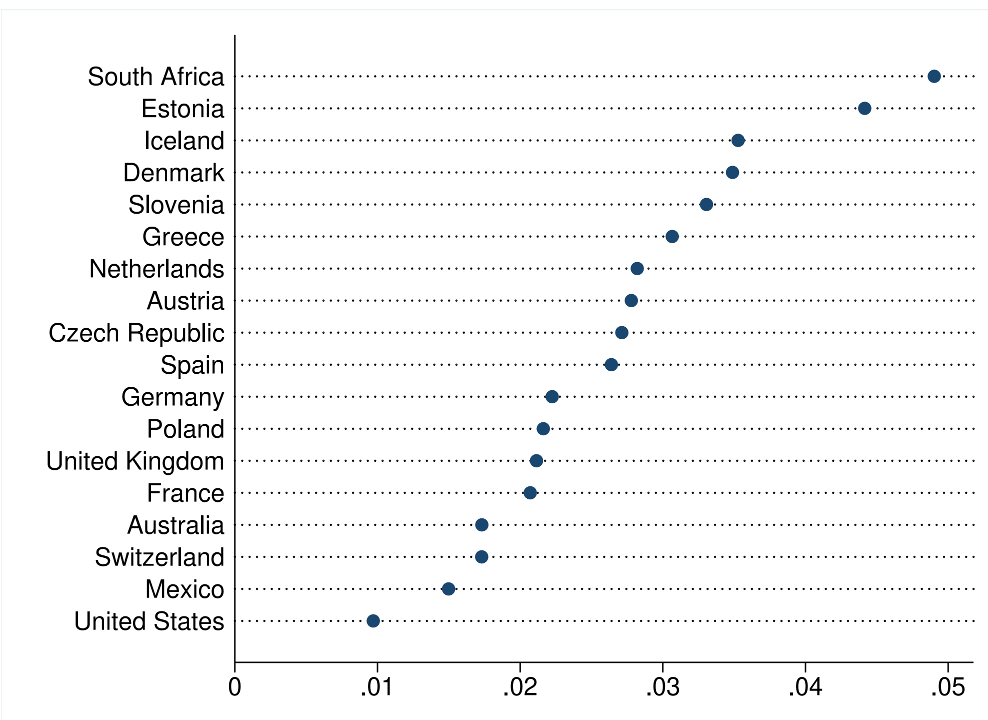

Figure 5: Estimated rise in Gini index due to consumption taxes in 2010 (year UK=2013 and year DK=2004)

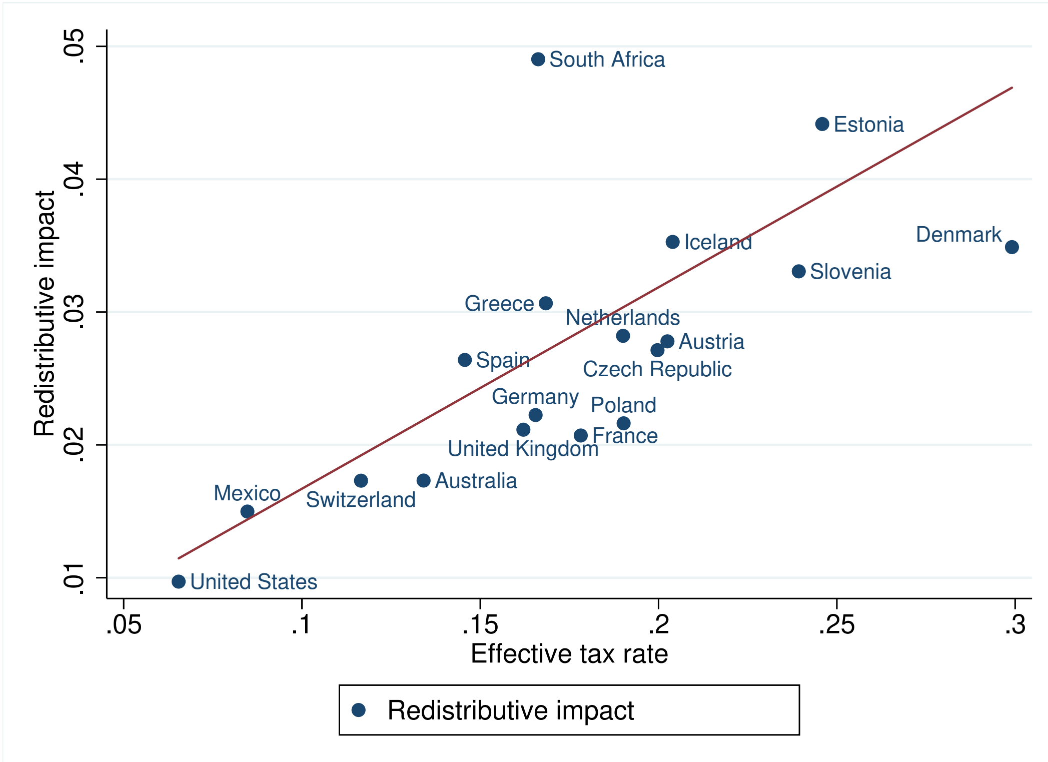

Figure 6: Redistributive impact is mainly driven by the tax rate
(Years UK=2013 and DK=2004)

4.2.3 The redistributive impact is mainly driven by the tax rate

In this part, we investigate what are the drivers of the differences in the redistributive effect of consumption taxes between countries: do all the differences between countries can be explained by the average tax rate, as in the example of Greece versus the USA? Or can the different magnitudes of the anti-redistributive effects be also explained by more or less unequal distributions of propensities to consume? In order to answer that question, it is useful to decompose the indicator of redistributive effect.

Effective redistribution can be decomposed into vertical redistribution, measured by the Reynolds-Smolensky index RS, and horizontal redistribution, measured by the reranking index Re (see appendix C.1 for detailed calculation):

$$\Delta G = RS - Re$$

While the former is the amount of redistribution that is linked to the regressivity of the tax, the latter is orthogonal to the income distribution: it represents a redistribution operated between households of same disposable income. In practice, vertical redistribution constitutes most of the redistributive effect of consumption taxes (see fig. C.1.a).

As shown in Kakwani (1977), the RS index is the product of two terms, that are respectively linked to the regressive pattern of the tax, and to its average rate:

$$RS = K \cdot \frac{TIR}{1 - TIR} \tag{4}$$

Here TIR is the aggregated tax-to-income ratio, defined as the average consumption taxes paid over the average disposable income, and K is the Kakwani index, a measure of the regressivity of consumption taxes (see appendix C.2). This result shows that vertical redistribution can be decomposed into one distributional parameter, and one macro-level parameter. The Kakwani index is determined by the pattern of consumption and income of the population, and is not a policy parameter. On the other hand, the tax rate can be tuned by the policy maker.

We see in fig. 6 that the tax rate explains most of the redistributive effect between countries. Differences in redistributive effect between countries of same average rate can be explained by different levels of tax regressivity. On this graph, it can be seen that the very high level of anti-redistributive impact of consumption taxes in South Africa can be explained by the regressive pattern of the tax. Indeed, while the effective tax rate is of same order than in other developed countries, its very high income inequality entails high inequality in consumption and saving rates.

In practice, we see that there is little variation in the value of the Kakwani index of regressivity. We compute the Kakwani indices for all the datasets where consumption data is available (i.e. 77 country-years). Approximately half of Kakwani indices lie between -0.10 and -0.15, while almost all of them lie between -0.05 and -0.20 (see fig. C.2.a). We see that vertical redistribution is mainly driven by the tax rate, since the Kakwani indices lie more or less in the same region (see fig. 7).

Our imputation model produces a similar range of Kakwani indices: most of them lie between -0.10 and -0.15. It is accurate enough in order to distinguish between the very low, medium and high Kakwani indices. The absolute difference between the Kakwani index computed from imputed data and that from observed data is less than 0.055 in 9 out of 10 country-years, and this situation arises only in high income inequality countries such as South Africa, India or Mexico.

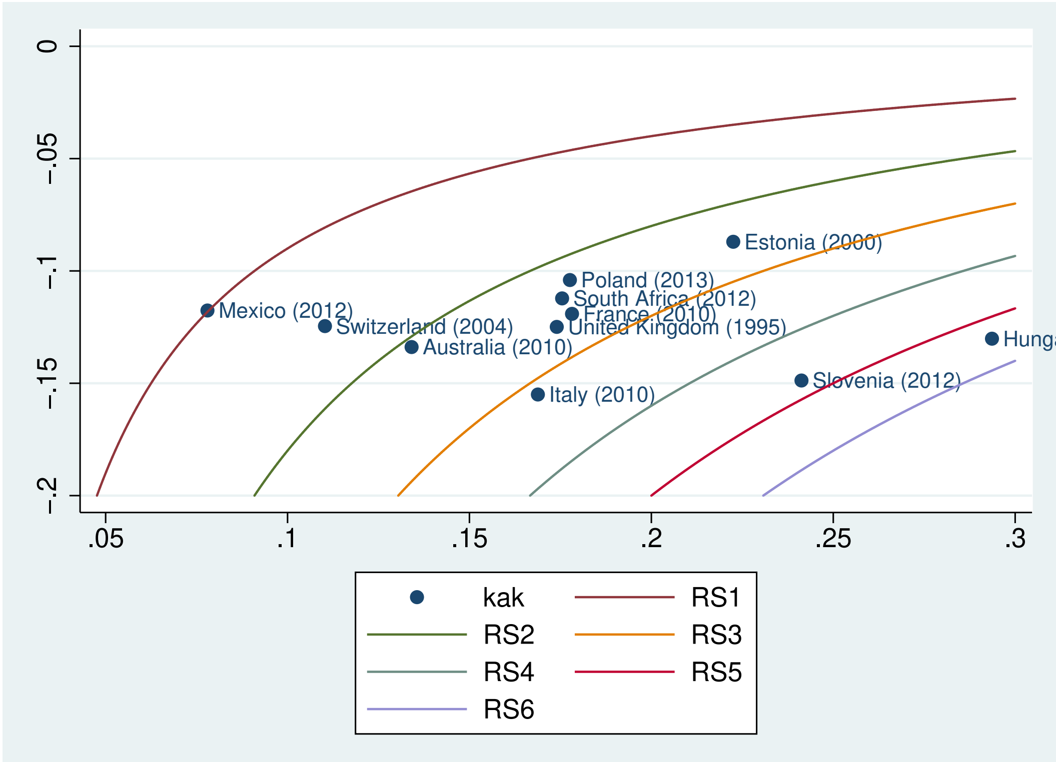

Figure 7: Kakwani indices (y-axis) depending and effective tax rates (x-axis)
Curves are isolines of Reynolds-Smolensky index.

5 Discussion

Consumption tax raises a lot of resources, but it is an unfair tax. Curiously, the distributive effect of consumption taxes is the blind spot in international comparisons of redistribution systems. No studies have undertaken micro-data work with 22 countries and 17 years, as we are doing here. In this study, we have measured the impact of consumption taxes on inequalities to see to what extent this tax may counter the redistribution operated by other social and fiscal tools.

This analysis is a follow up of a previous research decomposing the distributive impact of direct taxes and transfers across developed economies, and showing that both tax progressivity and the average rate of taxation have large impacts on redistribution (Guillaud, Olckers, and Zemmour, 2019). We thus add a block to the analysis of tax and transfer systems, by taking indirect taxation into consideration, and measuring its effect on the distribution of net disposable income.

The impact of consumption taxes on inequality can be broken down into two effects: a behavioral parameter, which is the propensity to consume (although it is found to be declining in all countries, but the international variation is not so great), and a more political parameter, which is the average effective rate (whose international variation is relatively greater in our sample, in the range of 1 to 3). Our study shows that the redistributive effect of consumption taxes greatly varies from one country to another, because of the political parameter of the average rate. Thus, countries with significant anti-redistributive effects are those that have chosen to have high consumption taxes rates (e.g. DK).

We have been able to compare the (small) anti-redistributive impact of indirect taxation (consumption tax, including value-added-taxes, sales taxes, excise taxes, and taxes on specific services) to the (big) redistributive effect of direct taxes (personal income tax, social security contributions) and cash transfers across OECD countries. Our results lead to the conclusion that since consumption taxes makes it possible to finance public goods, the moderate anti-redistributive effect found is not sufficient to justify a reduction in taxation.

References

Boutchenik, Béatrice. 2015. "Les effets redistributifs de la taxe sur la valeur ajoutée." Tech. rep., Conseil des prélèvements obligatoires.

Carey, David and Harry Tchilinguirian. 2000. "Average effective tax rates on capital, labour and consumption." *OECD Economics Department Working Papers* 258.

Dauvergne, Roy. 2012. *Qui paye les taxes indirectes en France ?* Master's thesis, Paris School of Economics.

Decoster, André, Jason Loughrey, Cathal O'Donoghue, and Dirk Verwerft. 2010. "How regressive are indirect taxes? A microsimulation analysis for five European countries." *Journal of Policy analysis and Management* 29 (2):326–350.

Eurostat. 2016. *Taxation Trends in the European Union*. Publications Office of the European Union.

Figari, Francesco and Alari Paulus. 2015. "The Distributional Effects of Taxes and Transfers Under Alternative Income Concepts: The Importance of Three "I" s." *Public Finance Review* 43 (3):347–372.

Figari, Francesco, Alari Paulus, and Holly Sutherland. 2015. "Microsimulation and policy analysis." In *Handbook of income distribution*, vol. 2. Elsevier, 2141–2221.

Garfinkel, Irwin, Lee Rainwater, and Timothy M Smeeding. 2006. "A re-examination of welfare states and inequality in rich nations: How in-kind transfers and indirect taxes change the story." *Journal of Policy Analysis and Management* 25 (4):897–919.

Guillaud, Elvire, Matthew Olckers, and Michaël Zemmour. 2019. "Four Levers of Redistribution: The Impact of Tax and Transfer Systems on Inequality Reduction." *Review of Income and Wealth* .

Kakwani, Nanak C. 1977. "Measurement of tax progressivity: an international comparison." *The Economic Journal* 87 (345):71–80.

Kato, Junko. 2003. *Regressive taxation and the welfare state: path dependence and policy diffusion*. Cambridge University Press.

LIS. 2019. "Luxembourg Income Study (LIS) Database." http://www.lisdatacenter.org. (multiple countries; accessed May to August 2017 and June to December 2019). Luxembourg: LIS.

Mendoza, Enrique G, Assaf Razin, and Linda L Tesar. 1994. "Effective tax rates in macroeconomics: Cross-country estimates of tax rates on factor incomes and consumption." *Journal of Monetary Economics* 34 (3):297–323.

O'donoghue, Cathal, Massimo Baldini, and Daniela Mantovani. 2004. "Modelling the redistributive impact of indirect taxes in Europe: an application of EUROMOD." Tech. rep., EUROMOD Working Paper No. EM7/01.

Piketty, Thomas and Emmanuel Saez. 2007. “How progressive is the US federal tax system? A historical and international perspective.” *Journal of Economic perspectives* 21 (1):3–24.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. “Distributional national accounts: methods and estimates for the United States.” *The Quarterly Journal of Economics* 133 (2):553–609.

Savage, Michael. 2017. “Integrated Modelling of the Impact of Direct and Indirect Taxes Using Complementary Datasets.” *The Economic and Social Review* 48 (2):171–205.

Appendix

A Methodology

A.1 Scaling with national accounts

Households' propensities to consume are the first term in eq. (1). Taxable consumption includes all monetary expenditure, excluding rents, which are not subject to consumption taxes. Moreover, loan repayments are considered as savings, not consumption.

$$prop_i = \frac{hmc_i - rents_i}{dhi_i}$$

where hmc is household monetary consumption, $rents$ is household expenditure in rentals, dhi is disposable income of household.

Disposable income is the amount of money that households have available for spending after income taxes, social security contributions, and transfers have been accounted for.

After some preprocessing on the data (equivalization, bottom-coding), propensities to consume are scaled according to national accounts data. In order to do a international comparison, micro-data is transformed to reflect the amounts of consumption and income at the national level. Conceptually, this means that micro-data are used to get the shape of consumption (its relative distribution with income), while the total levels of income and consumption are put in accordance with national accounts.

We apply a scaling factor on propensities to consume.

$$scaled\ prop_{c,y,i} = scaling_{c,y} \cdot prop_{c,y,i}$$

$$scaling_{c,y} = \frac{CH - R}{I - R_{imputed}} \cdot \frac{\sum_i dhi_i}{\sum_i hmc_i - hmchous_i}$$

where:

- CH is final consumption expenditure of households[9]
- I is gross disposable income of households

In national accounts, services of housing that owners produce for themselves are included in both the consumption expenditure and the income of households. We thus remove those terms from the scaling factor of propensities to consume. For some countries however, information on imputed rentals was missing, so this term was omitted.

[9] When this figure was not available, we used expenditure and income of households and non-profit institutions serving households.

These national accounts ratios are stored in a Stata data file, and scaling factors are computed after the preprocessing.

We are now able to compute the tax ratio on each household i in country c at year y, that is

$$TIR_i = \tau_{c,y} \cdot scaled\ prop_{c,y,i} = \underbrace{\tau_{c,y} \cdot \frac{CH - R}{I - R_{imputed}}}_{macro\ data} \cdot \underbrace{\frac{\sum_i dhi_i}{\sum_i hmc_i - hmchous_i} \cdot prop_i}_{micro\ data}$$

A.2 Definition of effective tax rate

We define tax rates implicitly: instead of using official statutory rates that do not reflect the average rate at which households' consumption is taxed, we use tax revenue and consumption data from national accounts. For each country-year, we compute the ratio between consumption taxes revenue and the amount of taxable consumption. This defines the average effective tax rate on consumption paid by households.

We rely on previous research on implicit tax rates (see appendix B.1), and try to improve existing definitions by defining the effective tax rate as follows:

$$\tau_{c,y} = \frac{consumption\ taxes\ revenue}{C - CGW - R}$$

where *consumption taxes revenue* includes all revenue from consumption taxes, including value-added-taxes (or sales taxes if applicable), excise taxes, taxes on specific services, etc. C is the total final consumption expenditure (private consumption and consumption of general government). CGW is the amount of wages of employees paid by general government, and $R = R_{actual} + R_{imputed} =$ are actual and imputed rentals for housing.

It is important to note that the "final consumption expenditure of households" includes actual rentals that tenants are paying and imputed rentals that occupying owners are paying to themselves. We need to remove this from the denominator, as it generates no tax revenue.

The government does not purchase its whole consumption. In order to account for the share of its consumption that is self-produced (and thus does not generate tax revenue), we remove the term CGW from the denominator, as per Mendoza, Razin, and Tesar (1994).

For each country-year, we compute tax rates according to the three available definitions. When some numbers are missing, we impute rates with a regression model based on other tax rates (see appendix B.1).

A.3 Consumption imputation model

We use a generalized linear model in order to impute consumption data according to households characteristics. The key issue here is that consumption data will have to be imputed on whole datasets for which there is no observation on expenditure. Therefore, it is necessary to design a model that can be calibrated on some countries and used on others.

In order to achieve that, we apply medianization to all monetary variables (including income, consumption and value of housing). That is, all monetary variables are expressed in proportion of their median values in the dataset. The **medianized disposable income** is thus:

$$\widehat{income}_i = \frac{income_i}{\text{median}_{c,y}(income_i)}$$

The underlying assumption is that the relations between medianized consumption and medianized income are similar in every country, account taken of other individuals determinants. In order to reduce the heterogeneity of the different countries, we remove from the training sample a few countries whose income distributions are the most extreme. That is, countries with a high degree of income inequality or with very low median income, compared to the rest of the sample. The list of countries used in the regression are listed in appendix D. We nonetheless apply the imputation model to every country, whether they have been used to calibrate the model or not.

We design two nested models, depending on the information available in the dataset. The first one is a generalized linear model where dependent variables are the medianized disposable income of the household, a few socio-demographic determinants (number of members in household and conjugal status of head), and a binary variable indicating whether the household is under the monetary poverty threshold. This threshold is defined as 60% of the median equivalized income. Therefore, a household is under the poverty threshold if its medianized disposable income is under 0.6. This indicator has been added in order to account for the fact that, for lowest incomes, levels of consumption tend to be less correlated (even not correlated) with income [référence à ajouter]. Therefore, this indicator allows to account for different consumption patterns of poor households.

$$\log(\widehat{hmc}) = \alpha_0 + \beta_0 \log(\widehat{dhi}) + \mathbb{1}_{pov} \cdot \Big[\alpha_1 + \beta_1 \log(\widehat{dhi})\Big] + \Gamma^\intercal X + \epsilon \tag{mod. 1}$$

In the second nested model, we add some socio-demographic determinants (number of people aged 65 or older, number of wage income-earners in household, housing ownership status). Moreover, we add another monetary variable that is the total imputed or effective cost of housing. This amount may correspond to actual money spent by the household in rentals, or to non-monetary consumption for housing services (e.g. imputed rents for occupying owners). This variable is much more widely available in households surveys than total consumption, and is a good proxy of the standard of living of the household. Therefore, it carries valuable information on the amount of consumption of the household. The second model is thus:

$$\log(\widehat{hmc}) = \alpha_0 + \beta_0 \log(\widehat{dhi}) + \mathbb{1}_{pov} \cdot \Big[\alpha_1 + \beta_1 \log(\widehat{dhi})\Big] + \delta \log(\widehat{housing}) + \Theta^\intercal Y + \epsilon \tag{mod. 2}$$

With this regression model, we impute medianized values of the households' monetary consumption. It is then scaled with national accounts data in order to be comparable with observed values, according to section 3.2.1.

B Robustness checks

B.1 Different definitions of effective tax rate

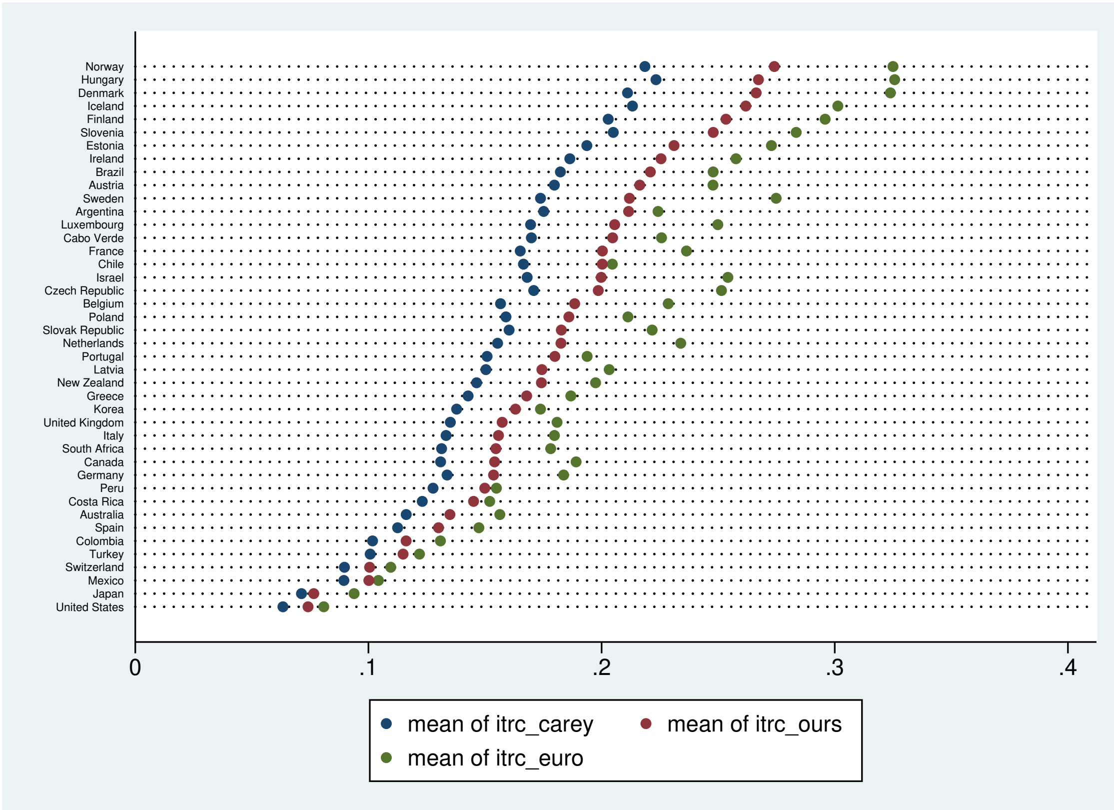

Figure B.1.a: Mean of implicit tax rates on consumption for each country.

There are three main definitions for computing implicit effective tax rates on consumption in the economic literature, as described in Eurostat (2016); Mendoza, Razin, and Tesar (1994); Carey and Tchilinguirian (2000). We draw on those works in order to propose the following definition:

$$\tau_{c,y} = \frac{consumption\ taxes\ revenue}{C - CGW - R}$$

where *consumption taxes revenue* includes all revenue from consumption taxes, including value-added-taxes (or sales taxes if applicable), excise taxes, taxes on specific services, etc. $C = CP + CG$ is the total final consumption expenditure (private consumption and consumption of general government). CGW is the amount of wages of employees paid by general government, and $R = R_{actual} + R_{imputed} =$ are actual and imputed rentals for housing.

The different possible definitions of implicit tax rates rely on different definitions of the taxable consumption. For example, the definition in Eurostat (2016) relies on a narrower taxable basis, constituted only of private consumption

$$\tau_{c,y} = \frac{consumption\ taxes\ revenue}{CP} \tag{5}$$

while the definition in Carey and Tchilinguirian (2000) considers a broader definition, by taking all consumption

$$\tau_{c,y} = \frac{consumption\ taxes\ revenue}{C} \tag{6}$$

The choice of removing or not rents from the denominator depends on the definition of taxable consumption in micro-data. Since we account for the fact that rents are not subject to consumption taxes by removing rents from the micro-data on consumption, we subtract rents from the denominator of the implicit tax rate. If we do the same for the two alternative definitions described earlier, our definition of implicit tax rates on consumption is thus structurally bounded above by the tax rate under definition (5) and below by that under definition (6) (see fig. B.1.a). These alternative definitions can be used to produce robustness checks.

B.2 Estimated regressivity is mitigated when taking rents into account

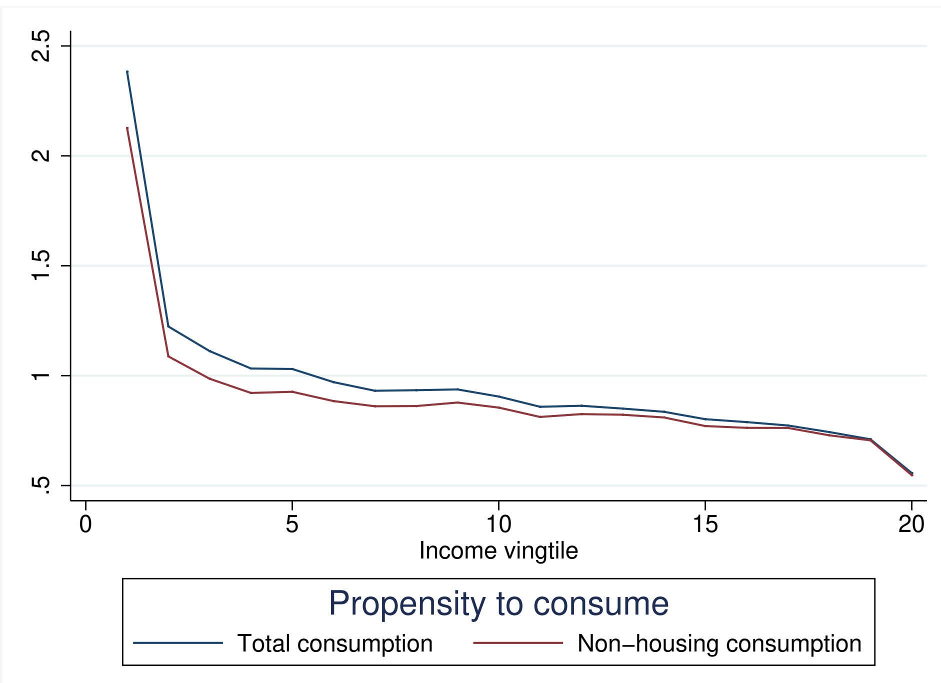

Figure B.2.a: Rents represent a higher share of consumption at the bottom of the income distribution (France 2010)

Our method allows to account for the fact that housing rentals are not subject to consumption taxes. They are an important part of households' consumption, and they represent a higher share of consumption for poorer households (fig. B.2.a). As a result, the downward slope of propensities to consume are less pronounced when rents are removed from the total amount of consumption. Therefore, we can conclude that micro-simulation methods

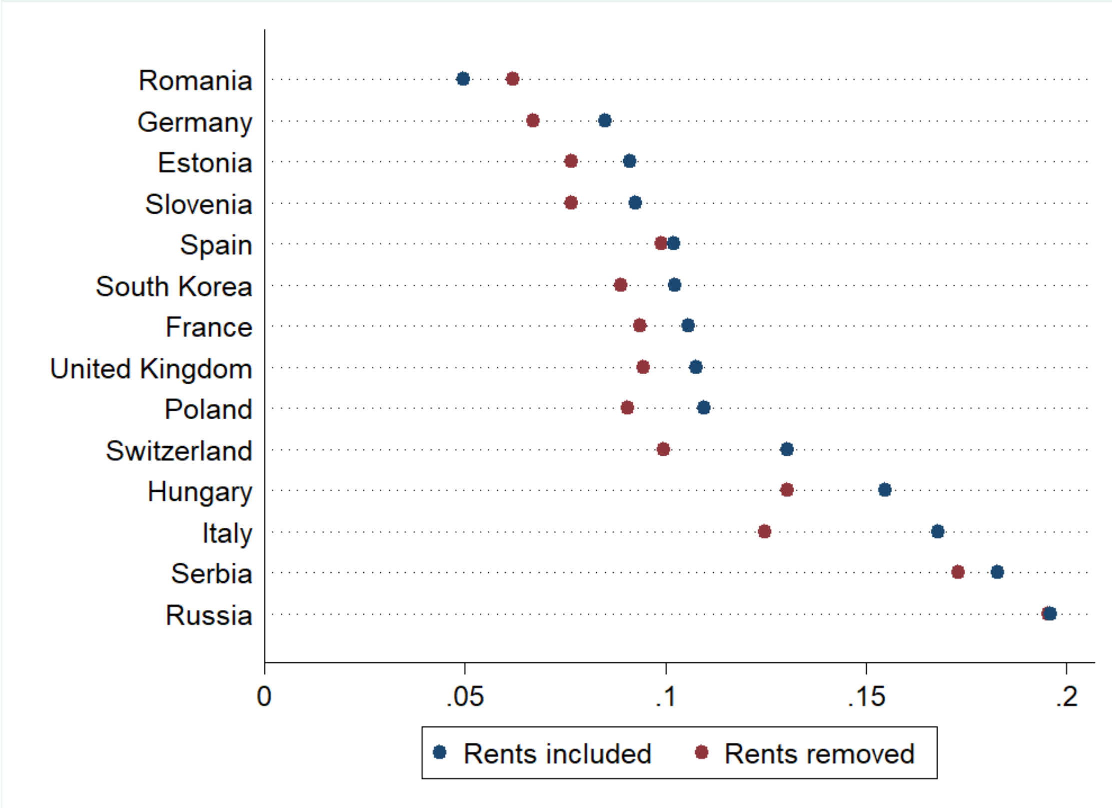

Figure B.2.b: Mean value of Kakwani index whether taxable consumption includes rentals

which apply tax rates on the whole consumption (rent or not) are slightly overestimating the regressivity of consumption taxes.

In order to maximize our coverage of countries and years, we also define another version of the effective tax rate, where actual rentals are not removed from private consumption at the denominator. This definition will be used when micro-data on consumption is not separable between rentals and the rest of the consumption. This smaller rate will be applied on a bigger amount of consumption.

$$\tau_{wr} = \frac{consumption\ taxes\ revenue}{CP - R_{imputed} + CG - CGW}$$

As shown in fig. B.2.b, estimated regressivity is reduced when rents are taken into account, and removed from the amount of consumption: the absolute value of the Kakwani index of regressivity can be reduced by up to 20% for some countries.

B.3 Progressive consumption tax rates: test of the bundle effect

Many countries enforce reduced VAT rates for some goods, either in order to boost some economic sectors or to lighten the amount of consumption taxes paid by least affluent households. On the other hand, some goods are more heavily taxed, such as oil or alcohol.

These variations on statutory tax rates can affect the overall progressivity of the tax, as the baskets of goods are different at one point or another of the income distribution. As a robustness check, we design two scenarios, where the effective tax rate is increasing with income. These situations would occur if tax rates had been specifically designed to make consumption taxes more progressive, as it is generally the case. These scenarios introduce a deviation from the median effective tax rate that depends linearly on the percentile of income. In the intermediate scenario, the first percentile of income enjoys an 0.5 percentage points lower effective tax rate than the median household, while the last percentile faces a 0.5 points higher tax rate, so that there is a 1 percentage point difference between the effective tax rate paid by poorest households and richest households. We also provide an extreme scenario where the gap between effective consumption tax rates at both ends is 2 percentage points. Those gaps are in accordance with national case studies that compute effective tax rates depending on income.

We show in fig. B.3.a the distribution of tax-to-income ratios by percentiles of income. This allows to compare different situations, from the least progressive to the more progressive. In the first situation, we assume that all goods and services (including housing) would be taxed at the same effective tax rates for all households. The regressive pattern then comes exclusively from the propensity to consume. In the second curve, consumption taxes are applied only on non-housing consumption, which mitigates the regressive pattern (as seen in appendix B.2). The third and fourth curves introduce progressive consumption taxes according to the intermediate and extreme scenarios.

The curves are very close to one another: this confirms that the ‘ propensity to consume” effect is the first order effect. Indeed, simulating an aggregated amount of consumption for each household allows to capture most of the regressivity of consumption taxes. When added the fact that housing consumption is not uniformly distributed, a uniform effective tax rate is even closer to the two scenarios of progressive tax rates.

Eventually, the first and last curves yield upper and lower bounds for the regressivity of consumption taxes. Indeed, we know that using the same effective tax rate for every household, and not taking into account the fact that housing is not subject to consumption taxes actually yields an overestimation of the regressivity of consumption taxes. On the other hand, when we take housing rentals into account and we apply a strong progressivity to effective consumption tax rates, we know that this yields a lower bound of the regressive effect. We can observe that even in the extreme scenario, tax-to-income ratios still present a strong regressive pattern.

The same can be said when looking at measures of redistribution and inequality. Figure B.3.b shows that a constant effective tax rate on whole consumption produces the highest post-tax Gini index (overestimating the eventual income inequality). Adding the information on rent captures most of the difference between the latter and both progressive scenarios. In every case, we observe that all those measures of post-tax income inequality are much closer to one another than to the inequality of disposable income. Qualitatively, the anti-redistributive effect stays significant, and the first and simplest measure provides a quite tight upper bound.

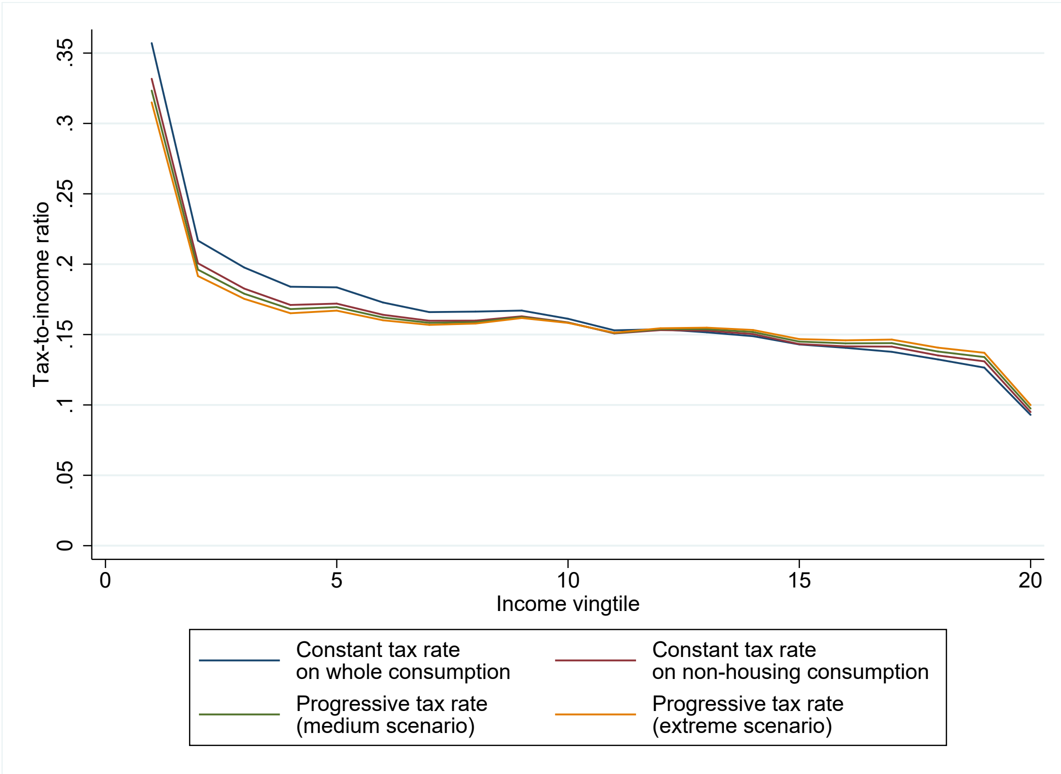

Figure B.3.a: Tax-to-income ratios for France 2010. Medium scenario: 1 pct point difference between effective consumption tax rates of the richest and the least affluent. Extreme scenario: 2 pct points difference

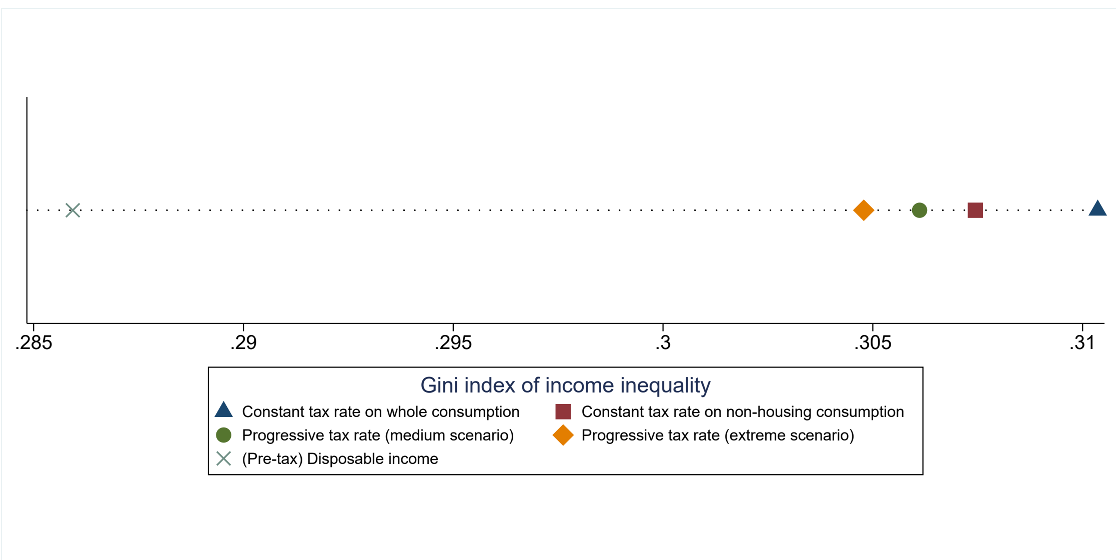

Figure B.3.b: Gini indices of income inequality for France 2010. Medium scenario: 1 pct point difference between effective consumption tax rates of the richest and the least affluent. Extreme scenario: 2 pct points difference

B.4 Imputation error on consumption depending on the model

Figure B.4.a shows the R2 coefficient for 57 countries-years where imputed and observed values of consumption can be compared. Out of the 57 country-years for which we impute consumption both with (mod. 1) and (mod. 2), we observe that model 2 increases the explained variance for 54 of them. On average, it is increased by 33%, as measured by the R2 coefficient. Overall, the R2 coefficient for model 2 is 0.45 on average, being larger than 0.36 for 75% of country-years, and larger than 0.56 for a quarter of our observations.

This shows that the independent variable "cost of housing", which is the main difference between the two models, provides significant additional information for the imputation of the households' consumption.

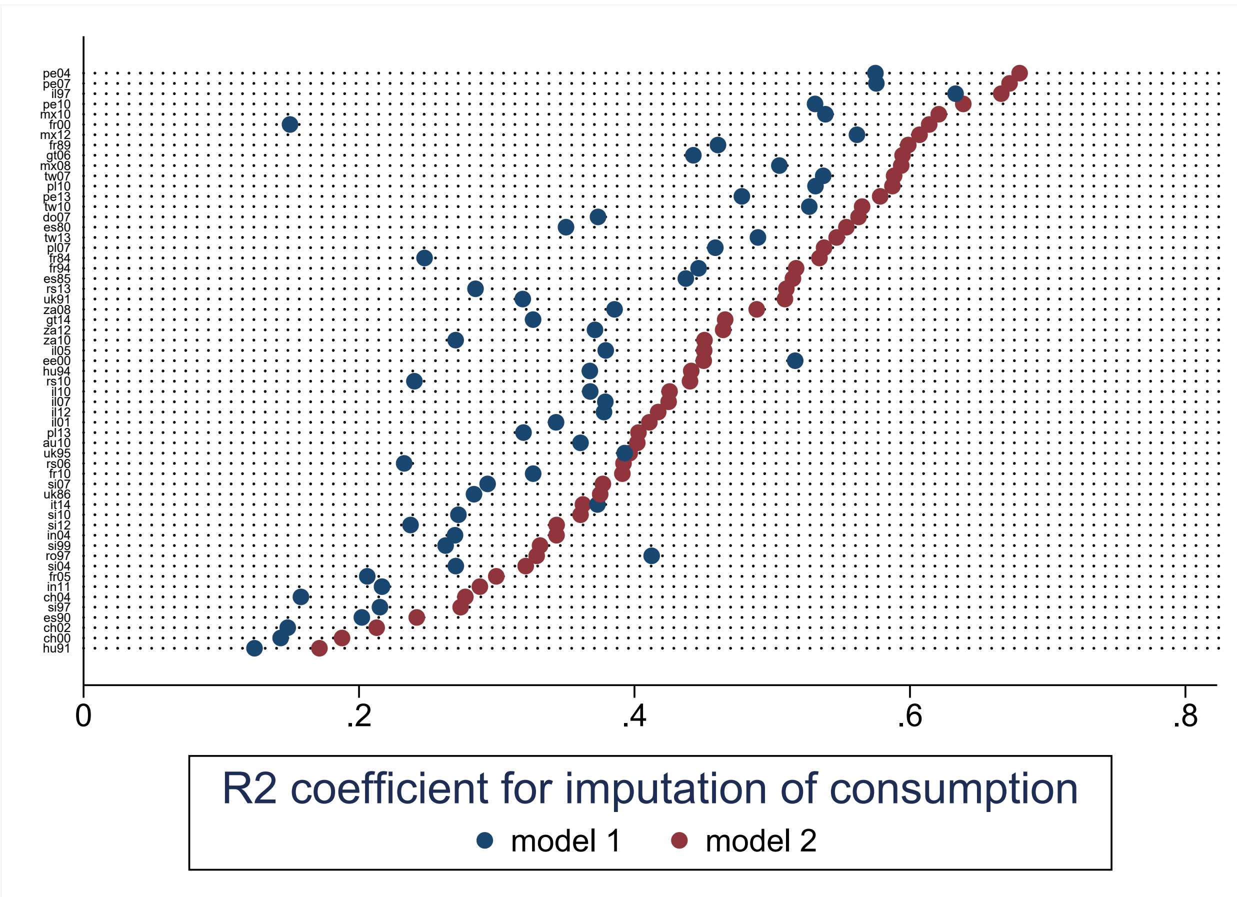

Figure B.4.a: Explained variance in the two models for various countries

C Decomposition of redistributive effect

C.1 Vertical and horizontal redistribution

Effective redistribution can be decomposed into vertical redistribution, measured by the Reynolds-Smolensky index (RS), and horizontal redistribution, measured by the reranking index (Re):

$$\Delta G = G_{dhi} - G_{dhi-tax} = RS - Re \tag{7}$$

Vertical redistribution relates to the amount of tax that is distributed in a progressive or regressive way related to income. A measure of vertical redistribution, the Reynolds-Smolensky index, is defined as follows [reference à ajouter]:

$$RS = G_{dhi} - C(dhi - tax, dhi)$$

where G_{dhi} is the Gini index of the pre-tax income, while $C(dhi-tax, dhi)$ is the concentration index of the post-tax income ordered on the pre-tax income. This term is thus relatively close to the Gini index of the post-tax income.

Horizontal redistribution is the amount of redistribution that is orthogonal to the distribution of income. The reranking index of horizontal redistribution is a measure of the amount of redistribution that is not due to the regressivity of the tax, but rather an inequality that is created between individuals in the same range of income. It is defined as follows:

$$Re = G_{dhi-tax} - C(dhi - tax, dhi)$$

By definition, the reranking Re is non-negative, so by eq. (7) the Reynolds-Smolensky index is an upper bound for effective redistribution, and it is a measure of the maximum reachable redistribution if no reranking was entailed by consumption taxes. In our case, if redistribution is negative, then the RS index is a lower bound for the anti-redistributive effect (in absolute value). The rise in income inequality due to taxes is thus the sum of the vertical anti-redistribution (due to the regressive pattern) and the reranking due to the variation in propensities to consume between households of same levels of income. In practice, the Reynolds-Smolensky index is close to the difference in Gini coefficients (see fig. C.1.a): the reranking generally accounts for less than 20% of the redistributive impact.

C.2 Kakwani indices of regressivity

We have seen in eq. (4) that the vertical redistribution operated by consumption taxes can be viewed as the product of two independent terms: the regressivity, a micro-level term linked to propensities to consume decreasing with income, and the rate of consumption taxes, a macro-level term.

We use the Kakwani index as an index of regressivity of consumption taxes. This indicator is a measure of how concentrated taxes are on one end of the income distribution or the other.

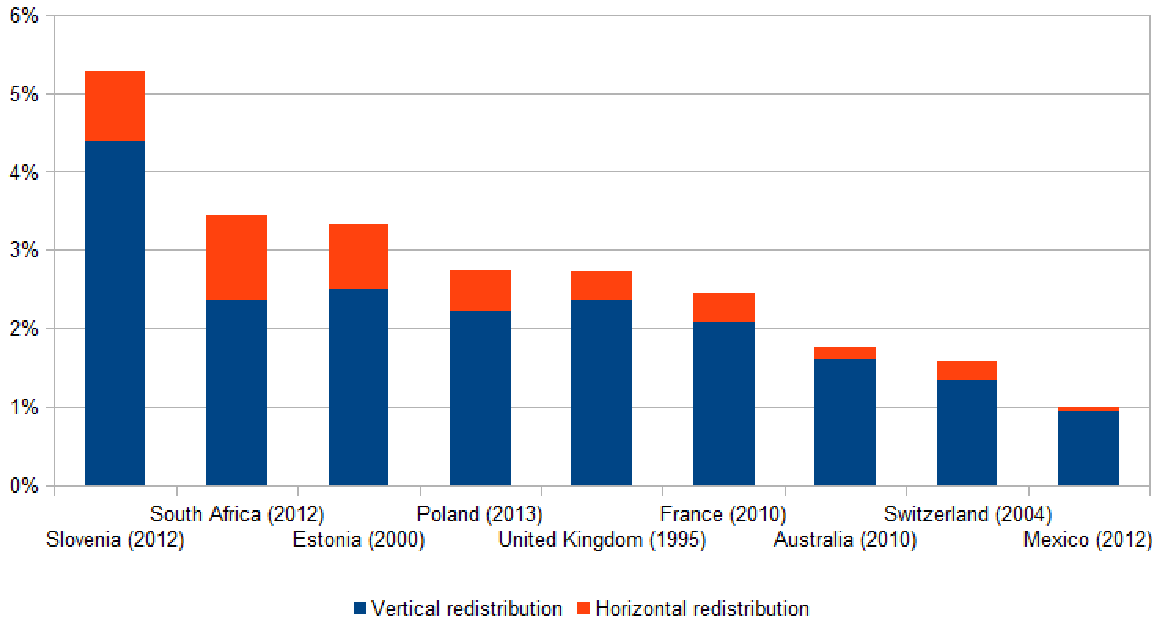

Figure C.1.a: Decomposition of redistributive effect

It is equal to the difference between the concentration index of the tax relatively to (pre-tax) income and the Gini coefficient of the income [référence à ajouter]. Namely:

$$Kakwani(tax, dhi) = C(tax, dhi) - Gini(dhi)$$

The concentration index $C(tax, dhi)$ is a measure of how much the distribution of the tax payments is skewed towards highest incomes. The range of its values is [-1;1], -1 indicating that all the tax payments are concentrated on the one poorest individual, while 1 indicates that all the tax payments are concentrated on the richest individual. The computation of this concentration index does not take into account the level of initial income inequality. By substracting the Gini index of income, the Kakwani index provides simple information, based on its sign: if the Kakwani index is positive, it means that the tax payments are more heavily concentrated towards the highest percentiles of income than income itself, meaning that the tax is progressive. On the contrary, if the Kakwani index is negative, then the distribution of tax payments is less skewed to the right than the distribution of income, meaning that the tax is regressive. We are thus expecting negative Kakwani indices.

For one fixed tax rate, we can make assumptions on the Kakwani index and thus have a range of possible RS index values. When the Kakwani indices are derived from imputed consumption values, this will be useful to provide lower and upper bounds on the possible RS values.

We compute the Kakwani index for all the datasets where consumption data is available (i.e. 77 country-years), the results are summed up in fig. C.2.a. Approximately half of Kakwani indices lie between -0.10 and -0.15, while almost all of them lie between -0.05 and

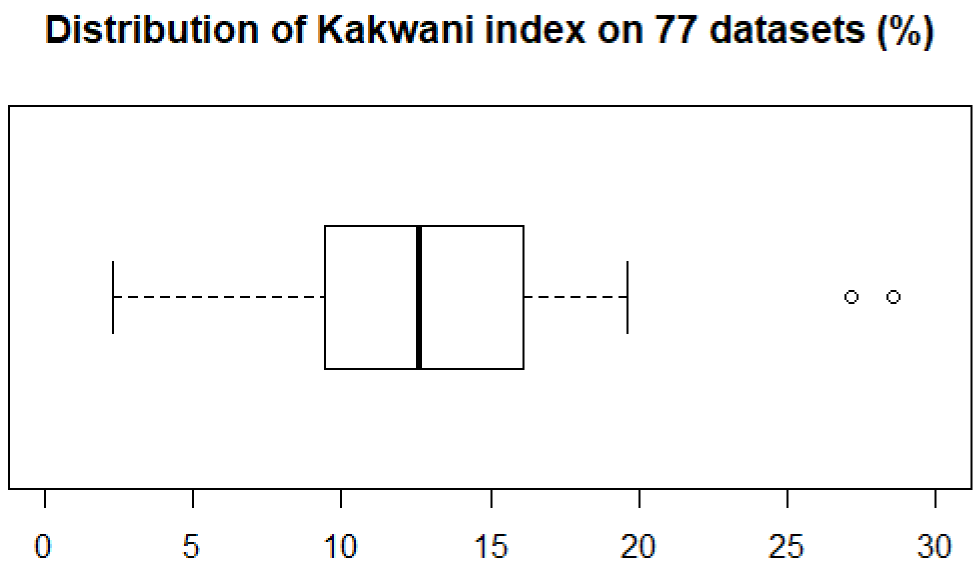

Figure C.2.a: Boxplot of the Kakwani indices

-0.20.

Based on the different tax rates that we have computed earlier, we are now able to frame the possible values of the RS index. As summed up in fig. C.2.b, most values for the Reynolds-Smolensky index will lie between -0.02 and -0.08.

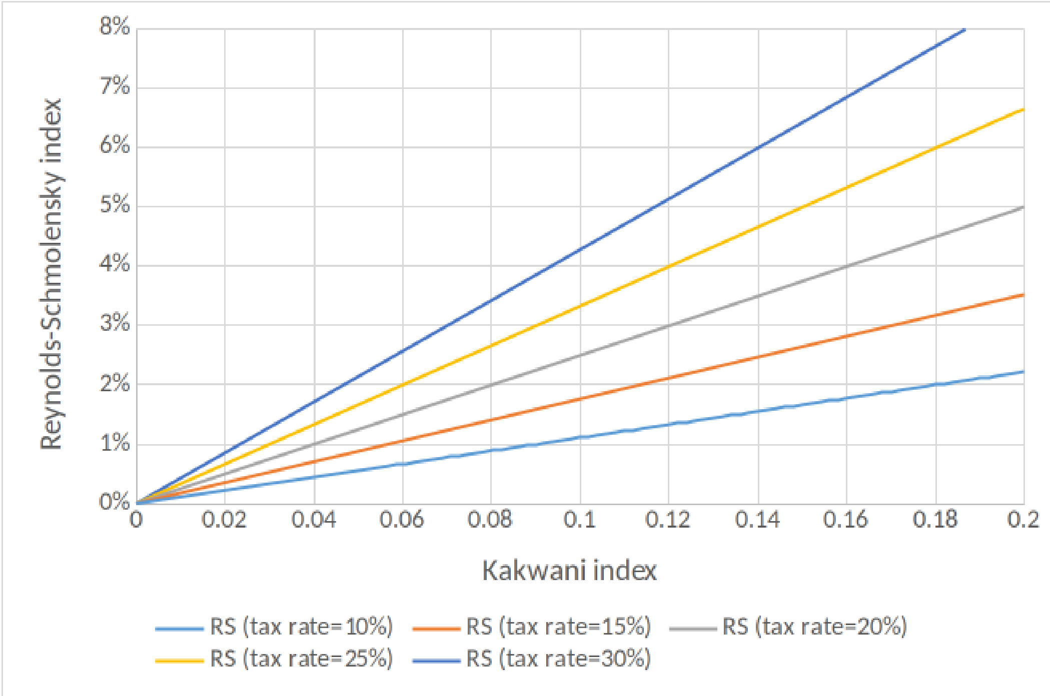

Figure C.2.b: Value of Reynolds-Smolensky index depending on tax rate and Kakwani index

D Country and years coverage

We select a total of **216** LIS datasets.

In table 1, countries marked with an **(R)** are used in the regression pool. For years marked with an *, information on rents is missing.

Table 1: Country and years used in the study

Country	Years with observed data	Years with imputed data
Australia **(R)**	2010*	1989, 1995, 2001, 2003, 2008
Austria		1994, 1995, 1997, 2000, 2004, 2007, 2010, 2013
Belgium		1992, 1995, 1997*, 2000
Brazil		2006, 2009, 2011, 2013
Colombia		2007, 2010, 2013
Czech Republic		2004, 2007, 2010, 2013
Denmark		1987*, 1992, 2000, 2004, 2007*, 2010*, 2013*
Dominican Republic	2007	

Estonia	2000	2004, 2007, 2010, 2013
Finland		1987*, 1991, 1995*, 2000*, 2004*, 2007*, 2010*, 2013*
France **(R)**	1984, 1989, 1994, 2000, 2005, 2010	
Germany		1984, 1989, 1994, 2000, 2004, 2007, 2010, 2013
Greece		2004, 2007, 2010, 2013
Guatemala	2006, 2014	2011
Hungary **(R)**	1991, 1994, 1999, 2005, 2007, 2009, 2012	
Iceland		2004, 2007, 2010
India	2004, 2011	
Ireland		1994, 1995, 1996, 2000, 2004, 2007, 2010
Israel	2001, 2005, 2007, 2010, 2012	1979*
Italy **(R)**	1995*, 1998*, 2000*, 2004*, 2008*, 2010*, 2014	1986*, 1987, 1989, 1991, 1993
Japan		2008
Luxembourg		1991, 1994, 1997, 2000, 2007, 2010, 2013
Mexico	2008, 2010, 2012	1984*, 1989*, 1992*, 1994*, 1996*, 1998*, 2000*, 2002*, 2004*
Netherlands		1983*, 1987, 2004, 2007, 2010, 2013
Norway		1979*
Panama		2010, 2013
Paraguay		2010, 2013
Peru	2004, 2007, 2010, 2013	
Poland **(R)**	2007, 2010, 2013	1986*, 1995*, 1999*, 2004*
Russia	2000, 2004, 2007, 2010, 2013	
Serbia	2006, 2010, 2013	

Slovakia		2004, 2007, 2010, 2013
Slovenia **(R)**	1997, 1999, 2004, 2007, 2010, 2012	
Spain **(R)**	1980, 1990	1995, 2000, 2004, 2007, 2010, 2013
Sweden		2000, 2005
Switzerland		1982*, 1992, 2007, 2010, 2013
Taiwan **(R)**	1981, 1986, 1991, 2007, 2010, 2013	1995*, 1997*, 2000*, 2005*
United Kingdom **(R)**	1986, 1991, 1995	1994, 1999, 2004, 2007, 2010, 2013
United States		1979, 1986, 1991, 1994, 1997, 2000, 2004, 2007, 2010, 2013
Uruguay		2004*, 2007, 2010, 2013