Evolution of fiscal systems: Convergence or divergence?

Paloma Péligry Xavier Ragot*

November 8, 2023

Abstract

We analyze the convergence or divergence of the diversity of fiscal systems after the financial crisis of 2007. Studying 29 countries, we first document the evolution of the taxation of households, firms, labour, consumption and capital. We identify three types of fiscal systems: liberal, intermediate and high-redistribution, which can be ranked in ascending order of tax rates, confirming known typologies in the diversity of capitalism literature. Only the tax rate on corporate profits shows signs of downward convergence over the period. The other tax rates show rather signs of divergence. Second, a divergence is observed among the liberal and high-redistribution group over the period. The European countries are converging towards the high-redistribution model, with the exception of Great Britain, which is moving towards the liberal model. Thus, the financial crisis seems to contribute not to the convergence, but to the divergence of fiscal systems.

Keywords: tax systems, globalization, capital taxation

JEL classification: H12, H6, P43

1. Introduction

Fiscal systems are the whole set of tax instruments that states are using to raise their resources. Their evolution is crucial, as it conditions the resources that the state raises for redistribution, for the financing of public goods and any other expenditures. In addition, the structure of fiscal systems is known to evolve rapidly after major economic crises (Martin, Mehrotra and Prasad (2009)). The evolution of fiscal systems after the 2007 crisis, which was the most severe crisis during the post-war period, may thus shed some light on the dynamics of welfare states, and the cause of their evolution.

* We thank Robert Boyer, Elvire Guillaud, Mathilde Viennot, Bruno Palier for useful discussions. We especially thank Elvire Guillaud and Mathilde Viennot for sharing their data. All erors are ours. Paloma Péligy, ENS Paris Saclay: paloma.peligry@ens-paris-saclay.fr. Xavier Ragot, Sciences Po, CNRS and OFCE: xavier.ragot@sciencespo.fr.

In particular, some scholars have documented a decrease in capital taxation (Genschel and Schwarz (2011)) sometimes referring to an emerging neoliberal state (Swank (2006), Hakelberg and Rixen (2021)). The goal of this paper is to assess more broadly the convergence of fiscal systems, which is important to anticipate the evolution of economic policies, but also to identify the social forces shaping the evolution of welfare states (Martin, Mehrotra and Prasad (2009)).

Fiscal systems are now studied from various perspectives, using economy, sociology and political science, providing a rich set of results. At least, two main views of the dynamic of fiscal systems can be found among scholars. The first one puts forward economic forces as the main driver of the structure of fiscal systems (Musgrave (1981)), either due to international competitions or due to the progressive use of the most efficient tools. This last view is mostly analyzed by economists in the public finance discipline and summarized by international institutions as best practices. A standard result is that the tax on mobile factor should converge to a very low rate due to international competition, whereas the tax on immobile factors should increase to compensate for the revenue loss. An important literature has studied the convergence of specific taxes and usually finds persistent differences in tax structures and the absence of "race to the bottom" even on taxes on mobile factors (Plümper, Troeger and Winner (2009), Genschel and Schwarz (2011) for a literature review).

In front of the persistent diversity of economic structures, a second approach has constructed typologies of welfare states or capitalisms, possibly explaining persistent differences in tax systems. Again, typologies have been provided by various disciplines and may differ in the key social aspects to construct the typology. To quote only a few important ones, Esping-Andersen (1990) provided a famous distinction between three types of welfare states: liberal, conservative and social-democratic welfare states. The French Regulation school (Amable and al. (1997) or Amable (2004)) have provided a typology of five types of capitalism where the state plays different roles. Comparative political economy often distinguishes between two types of capitalisms, coordinated market economies and liberal market economies (Hall and Soskice (2001)), where the state has a high or low role in redistribution.

In this article, a typology of fiscal systems and its evolution are provided for the period after the 2007 financial crisis. As mentioned above, our main research question is to identify if there is a convergence of fiscal systems after the crisis toward, as Peters (1991) finds in the seventies. For this investigation, we use OECD and LIS data to analyze the evolution of tax systems among OECD countries (See Section 2 and 3 and the Appendix for a description of the data and methodology) to construct a decomposition of total taxes in comparable eleven taxes rates for 29 countries over the period 2007 to 2019. Then, we use statistical procedures to identify a typology of fiscal systems across countries. We use the K-means method, which is now often used in machine learning, and which is described below in some details. The goal of this procedure is to identify consistent clusters using a

large set of variables, and it is thus more relevant than principal component analysis (PCA) for instance, which is used to reduce the dimensionality of a large set of variables.

Our first finding is that we indeed find a global convergence of corporate capital income tax to low levels, confirming results in a vast literature (Genschel and Schwarz (2011)). However, we find persistent and sometimes increasing differences in the taxation on labour, and a divergence in the evolution of the taxation on capital stocks. The taxation on the capital stock mainly concerns the taxation of real estate, which is a non-mobile factor. Thus, these trends show that the previous results on the convergence of taxes on mobile factors only are still valid, if the mobile factors are considered as being firms only. In addition, there is on average a shift from capital to labour taxation, confirming the trends identified by Carey and Tchilinguirian (2000) in the 1990s.

Then, the statistical approach to identify coherence confirms the distinction of three fiscal models, which can be easily identified: a high-redistribution model, composed of the large European countries; a liberal model, with the United States, the United Kingdom and Ireland in particular; and an "intermediate" model, composed of Japan and other European countries. This intermediate group is composed of countries roughly experiencing the same trends as the high redistribution group. Hence, this group may either catch-up with the high-redistribution group or it may compose a new group with an intermediate level of redistribution. This typology of countries, based solely on the taxation of countries, is consistent but a little bit different from the typology of welfare states of Esping-Andersen (1990), which considers other forms of state regulations. It is also consistent with the literature on comparative political economy, or the French Regulation School. As there is no recent specific typology of fiscal systems identifying the relevant tax rates since Peters (1991) to the best of our knowledge, we present our typology in some details.

Finally, and importantly, considering the dynamics of fiscal systems, our main finding is that the two main models (liberal and high-redistribution) are gradually diverging from 2006 and 2019. Total taxes increase in high-redistribution states, generating more redistribution, whereas they decrease in liberal states. During this period, labour taxes increased more in high-redistribution states, whereas capital taxes decreased in both types of fiscal systems. The European Union countries we consider belong to different models (Germany, France, Ireland, Spain, UK). However, we observe a convergence toward a more similar tax system, probably due to some European regulation and incentives (Kemmerling (2011)). It is interesting to note the divergence of Great Britain from the high-redistribution model towards which European countries were converging, well before the Brexit vote.

The striking divergence in fiscal systems raises the questions of the social forces at play (Martin, Mehrotra and Prasad (2009)). Again, they can be more economic in nature, reacting to different institutional comparative advantages, which are studied in the comparative political

economy literature and among institutionalists. Alternatively, they can be political in nature, being the outcome of different powers of social groups competing within heterogeneous legal structures, to reduce their tax burden (Peters (1991)). Finally, they can be due to different social preferences or "cultures". It isn't possible to clearly identify the specific forces at play out of the specific event, which is the 2007 crisis. The pieces of evidence seem however to point to independent political factors shaping the evolution of fiscal systems, one being obviously the current form of the European construction. They are discussed in the text and in concluding remarks.

The presentation of the article is as follows. The second section presents the accounting breakdown of the state budget, in order to identify all the components and explain the series. The trends in the main tax rates are then described. The third section presents the comparison of the redistributive effects of fiscal systems on individual data. The fourth section presents the results of the clustering and typology of fiscal systems. The fifth section is the conclusion.

Other related literature

This article is part of work studying the diversity of social states and fiscal systems. A seminal contribution is the work of Peters (1991), who analyzes fiscal clusters in OECD, in 1965, then studies their dynamics, from 1965 to 1987. During this period, his statistical method of clustering identifies four groups, countries in our "intermediate" model being separated in two groups. He finds a global convergence toward a uniform model of taxation, which is a broad-based model of taxation. Our result about the divergence of groups after a crisis is thus different. Observing the difference between Peters (1991) and our results, two specific clusters with low and high taxes and redistribution seem to persist, with some convergence periods, discussed by Peters (1991), and some divergent periods, discussed in the current paper. The rich literature on the diversity of capitalisms or welfare states provides many independent results on the structure of fiscal systems. Indeed, studies on the variety (Hall and Soskice (2001), Schneider and Paunescu (2012), Jessop (2012)) or diversity of capitalisms, including Regulationist work (Amable (2004), Boyer (2007)), finds that the diversity of capitalisms corresponds to a diversity of social states and therefore a diversity of fiscal systems. A recent study compatible with these analyses is Hasse and al. (2020), which draws up a typology of growth regimes in relation to the reforms of social states in developed countries. As mentioned above, the diversity of social states, or forms of welfare states, has been studied by Esping-Andersen (1990), who produced a typology of three types of welfare states. Fiscal systems are a component of welfare states and show systematic differences. Thus, Prasad and Deng (2009) study the progressiveness of taxation according to Esping-Andersen's typology of welfare states, and the results show that the taxation system is rather progressive in liberal countries (with the exception of the United Kingdom) and regressive in high-redistribution countries (European countries). This echoes the paradox of

redistribution form Korpi and Palme (1998), they find that the more countries target the redistributive policies, the less redistribution is actually achieved (more recently studied by Guillaud and al. (2020)). The results from Prasad and Deng (2009) are explained by the importance of consumption taxes in high-redistribution countries. The relative high tax burden of consumption particularly concerns the European countries, result that we analyze in the rest of the paper. Scholars have argued that this provides a tax base less distortionary than progressive income taxes (Lindert (2004)) and allows a low burden on capital (Wilensky (2002), Kato (2003)). It is also a very stable tax base, relatively less sensitive to competition and globalization, which ensures a financial basis for the welfare state. This could have enabled the development of the European welfare states (Kato (2003)) and their resilience (Hays (2003)).

Finally, studies of fiscal systems focus on differences in progressivity by analyzing household taxation. Piketty's work was pioneering in building historical databases (Piketty (2014)), and more recent works study components of taxation by extending the levies studied, such as social contributions or public expenditures (Amoureux and al. (2018), Rousselon and Viennot (2020), Causa and Hermansen (2018)). Within this work, our contribution is to mobilize a large number of series to construct typologies of fiscal systems, to understand their evolution.

2. Structure of the state budget: data and methodology

We first decompose the government budget in order to compare the tax structure between countries. This approach is based on comparable international data (OECD, Taxation Trends report of the European Commission (2020) among others), derived from national accounts. Simple accounting equalities allow a disaggregation of the state budget, both in terms of resources and expenditures. Government resources come from tax revenues (T) and the budget deficit (noted D). Government expenditures are composed of government final consumption expenditures (health, education, civil servants' salaries, etc.), noted G, interest payments on government debt rB (where r is the apparent interest rate on government debt and B is the total amount of government debt) and transfers to households and firms, noted Tr. These transfers are measured net of taxes and other contributions.

Then, the tax rates T can be decomposed according to the contribution of different taxes. As mentioned in the introduction, many decompositions of the tax structure are possible. Following the recent literature, we decompose the fiscal system by distinguishing between taxes on the factors of production, i.e. capital and labour, and on consumption. This first traditional decomposition (Mendoza and al. (1994)) aggregates economic transactions into three main markets to analyze redistributive effects. Next, we decompose taxes by the payers, i.e. firms or households. This second

decomposition between consumers and producers is also common (e.g. Taxation Trends (2020)), because the impact of taxes could be different in case of differentiated financial constraints[1]. Finally, for the taxation of capital, we distinguish between the taxation of capital income (e.g. tax on dividends) and the taxation of capital stock (e.g. tax on real estate), because the incidence of these taxes may also be different. For example, real estate ownership is immobile, whereas capital income may move between countries depending on the tax residence of the household. Unfortunately, it is not possible to distinguish in our data the share of capital stock taxation paid by corporations from the one paid by households. Finally, consumption taxation concerns only households. We therefore consider six taxes: consumption tax (T_C), tax on labour paid by households (T_{L_hou}) or by corporations and the self-employed ($T_{L\_corp\_SE}$), tax on capital income paid by households ($T_{K\_inc\_hou}$) or by corporations and the self-employed ($T_{K\_inc\_corp\_SE}$)[2] and tax on capital stock (T_{K_sto}), all six taxes are represented in Figure 1.

The analysis is based on OECD Revenue Statistics data, which allows us to study 29 countries, including 7 non-EU members (Canada, USA, Japan, Switzerland, Korea, New Zealand, Norway), between 2006 and 2019. They were complemented by Eurostat data, in particular by using micro-data to allocate taxes on household income between labour and capital. In order to allocate taxes between the aggregates labour, capital, consumption and corporations and households, we use the method described in the European Commission's Taxation Trends (2020) report, derived from the article by Mendoza and al. (1994) and completed by Carey and Tchilinguirian (2000). The initial aim of this method is to associate the yields of taxes on capital, labour and consumption with their tax base, in order to calculate average rates, also known as implicit tax rates (ITR). More particularly, it integrates the variation in the tax base which makes possible a better representation and comparison of the tax burden on each factor. However, the methodology of the implicit tax rates has some limitations regarding the reliability and comparability of the tax bases (for the ITR on capital, see Carey and Tchilinguirian (2000) and the discussion in Appendix). Therefore, in Section 2, we mainly focus on total tax yields, and not on implicit tax rates. During the period, some countries implemented some tax credits. Following the literature and consistent with OECD data, we consider tax income net of tax credits (as long as the tax credit does not exceed the amount of the tax collected) therefore, tax credit on the corporate income tax will be accounted as a decrease in the yields of corporate income tax. Some simplifications cannot be avoided for international comparisons. For instance, France implemented reimbursement of a part of corporate social contribution through a corporate income tax credit. This is here accounted by a reduction on the corporate labour contribution, what seems the most frequent representation.

The origin of the data, the description of the OECD tax classification and their construction are described in more details in the Appendix. The next figure presents the structure of the six tax rates.

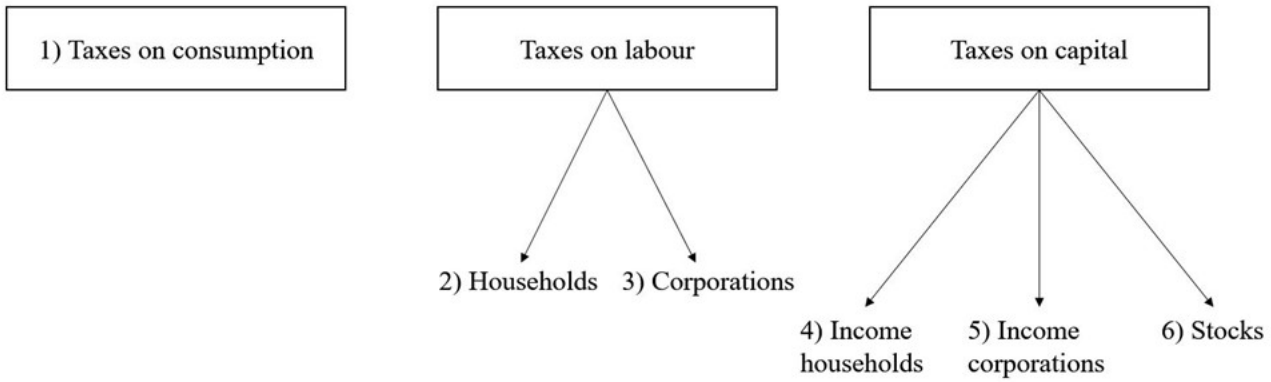

Figure 1: Decomposition of the six tax rates

2.1 The structure of the state budget: Evolution of some aggregates

International comparisons first allow us to identify the main differences between the ways in which governments finance themselves. The decomposition of the state budget for each country in 2006 and 2019 is presented in the Appendix to this article.

In Table 1, we present a global average of the main aggregates for the years 2006 and 2019. The first column is the year considered, either 2006 or 2019. The second column shows, for each year considered, the average of the variables and then the standard deviation of the aggregates considered. The column *rB* indicates the average interest burden on public debt relative to GDP across countries in 2019. The average is 1.5% in 2019 compared to 2.2% in 2006, which is due to the decrease in interest rates over the period. Column *G* represents government final consumption expenditure. Column *D* is the change in debt relative to GDP. This average turns out to be close to zero for the two years considered (but not over the whole period). Column *Tr* shows transfers to households over GDP. The *T* column is the overall tax rate. Finally, the last three columns break down the tax revenues into three components: the yield from the consumption tax (*C*), the labour tax (*L*) and the capital tax (*K*). Thus, the sum of the columns *rB, G* and *Tr* is equal to the sum of the columns *D* and *T*, and the sum of the last three columns is equal to the column *T*, for each country and each year. All variables are expressed in percentage of GDP.

The analysis of this table shows, first of all, the high average of the tax rate (T), around 35% of GDP, which increased slightly between 2006 and 2019. The variation in the tax rate between

countries is high, with a standard deviation of over 6%. A first reading of this table therefore indicates a relative stability of taxes and their structures between 2006 and 2019, as well as heterogeneity between countries.

	Year	rB	G	D	(=)	Tr	T	T C	T L	T K
Mean	2019	1.50	19.55	-0.19		14.92	35.78	10.84	18.72	6.22
	2006	2.19	18.9	-0.05		13.89	34.93	10.55	17.43	6.95
St. dev.	2019	1.02	3.39	2.44		4.09	6.10	2.99	4.58	2.16
	2006	1.25	3.14	4.81		4.53	6.34	2.68	4.83	2.65

Table 1: State budget and tax rates (mean and standard deviation)

However, this stability is illusory and the result of misleading aggregation. Before a systematic analysis, we present the results for two of the most different countries in our sample: the United States and France. The United States has the lowest tax rate in the sample over the period 2006-2019, and France the highest. To simplify the reading, Table 2 represents the same variables in columns. The composition of taxes is very different between the two countries. France, like the European countries, uses taxes on consumption (VAT), which are very low in the United States. Second, labour taxes are higher in France than in the United States. This is also the case, but to a lesser extent, for the taxation of capital.

The difference between France and the United States has increased over the period 2006-2019. The tax rate has increased in France while it has decreased in the United States. In both countries, however, the taxation of capital has decreased. The following analysis will show that these trends reveal divergences in fiscal systems, common to groups of countries.

2.2 Analysis of tax revenues

This section focuses on the structure of the tax burden presented in Table 3 and Figure 1. We provide the decomposition of taxes in 2006 and 2019 for all countries in Tables A1a and A2b in the Appendix[3]. As an introduction to the analysis, Figure 2 represents the structure of taxation for five significant countries and the 28-member European Union, in 2006 and 2019, decomposing the tax

revenues into the six taxes. The ordinates are percentages of annual GDP. The size of each tax thus represents the yield of the tax as a percentage of each country's GDP in that year.

Country	Year	rB	G	D	(=)	Tr	T	T C	T L	T K
FRA	2019	1.45	22.98	-3.06		23.87	45.23	12.12	24.76	8.36
	2006	2.60	22.76	-2.44		21.38	44.29	10.69	24.30	9.31
USA	2019	4.09	14.07	-6.66		13.68	25.19	4.09	14.82	6.27
	2006	3.98	15.00	-3.33		11.13	26.78	4.10	14.37	8.31

Table 2: State budget and tax rates (France and the United States)

First, some countries, such as Germany, France and Japan, have seen their tax rates increase, while the United States has seen its tax rates decrease. Countries with a high tax burden in 2006, such as France, have seen an increase in their tax rates, while countries with a low tax burden in 2006, such as the United States, have seen their tax rates decrease. We also observe a diversity of taxes used, as well as their variation.

These figure and tables contain a lot of information, which we break down into main results, focusing on the key trends in tax dynamics. We present the results in the form of stylized facts in order to compare them with existing literature.

	Year	T	C	L hou	L corp SE	K inc hou	K inc corp SE	K sto
Mean	2019	35.78	10.84	11.17	7.55	0.79	3.10	2.34
	2006	34.93	10.55	10.17	7.26	0.83	3.83	2.28
St. dev.	2019	6.10	2.99	3.04	3.39	0.50	1.21	1.21
	2006	6.34	2.68	3.15	3.33	0.58	1.93	1.14

Table 3: Six tax rates

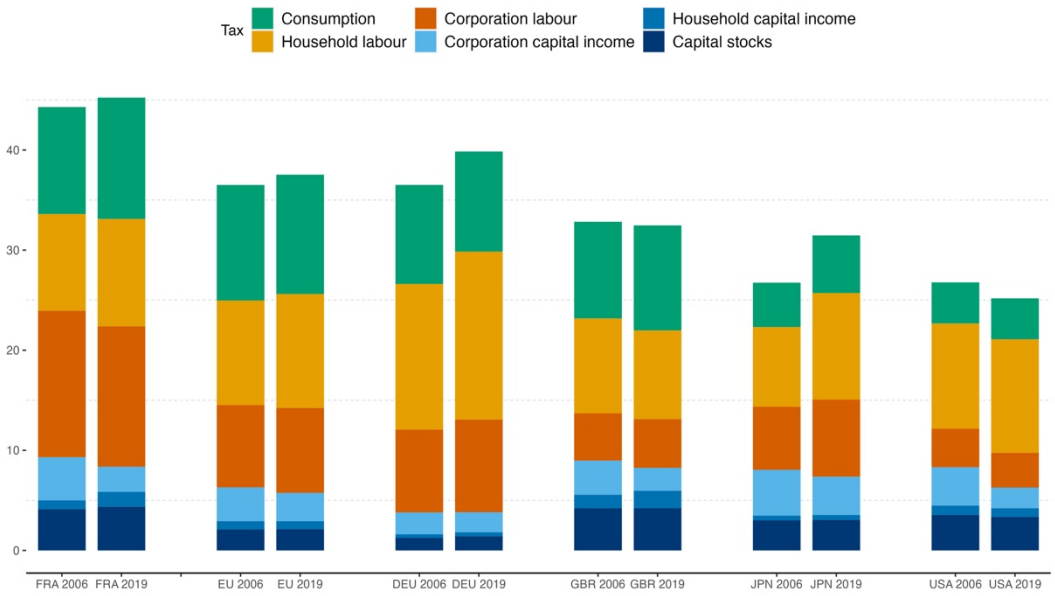

Figure 2: Taxation structure

Fact 1: Heterogeneity in the dynamics of tax revenues

The evolution of tax rates suggests persistent differences after the crisis from 2006 to 2019. The average tax revenues over GDP increased from 34.9% to 35.8% from 2006 to 2019, and heterogeneity, measured by standard deviation, decreased from 6.34 to 6.10 over the period. At the European scale, countries adopted different strategies between 2006 and 2010, when fiscal austerity resulted in lower public spending or higher taxes. Thus, the European tax rate decreased between 2006 and 2010 to reach 35.7%, but has been increasing almost constantly since, and in 2019 it was higher than in 2006, at 37.5% of GDP[4].

Fact 2: Importance and stability of labour tax share

A comparison of the structure of taxation across countries shows the importance and stability of taxes on labour in government revenues. It represents on average 49.9% and 52.3% of tax revenue respectively in 2006 and 2019. Moreover, this share of labour in the total taxes is stable in all the countries for the period: its normalized standard deviation is the lowest compared to all others tax shares (0.13 compared to 0.22 and 0.38 for respectively consumption and capital taxes).

Fact 3: Heterogeneity of distribution of the tax rates between households and corporations

Interestingly, countries differ in the fraction of taxes paid by households or by corporations. In our sample, we observe a slight shift in taxation from corporations, whose yields decrease by 0.5 percentage points from 2006 to 2019 to households, whose yields increase by 1.3 percentage points.

The increase in household taxation comes from consumption (+0.3 points) but especially from the taxation of household labour (+1.0 points).

2.3 Analysis of convergence

The relative heterogeneity of tax rate, however, hides convergence in some taxes. We now perform convergence tests on tax rates. We run both beta-convergence and sigma-convergence tests. Beta-convergence is a measure to evaluate the convergence over time of a variable. Beta-convergence is considered as a condition of sigma-convergence. The latter is captured by the downward trend, over time, of an indicator dispersion, here coefficient of variation and standard deviation. The detailed results are displayed in Tables A3 and A5 in Appendix. We here only provide the main results.

Fact 4: Convergence of corporate capital income taxation

Consistently with the literature, we observe a strong convergence in the taxation of corporate capital income, as shown in Figure 3. The coefficient of variation has decreased by 22 percent between 2006 and 2019.

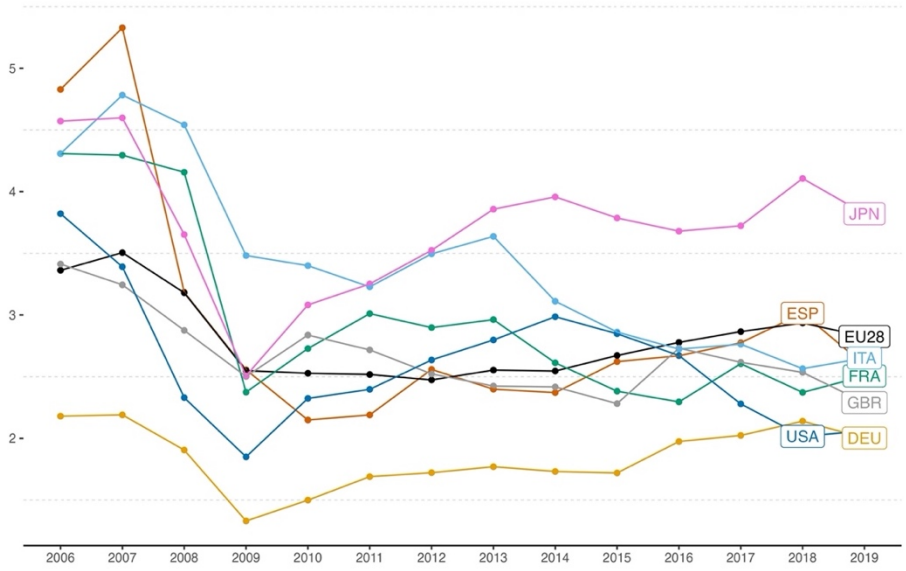

Figure 3: Taxes on corporate capital income

Except for Japan, which in 2019 was almost at its pre-crisis level, the rest of the countries have sharply reduced taxation on corporate capital income (such as taxation on profits).[5] Corporate capital income is indeed one of the most mobile tax bases. This convergence has already been investigated by many empirical studies, as reported in Leibrecht and Hochgatterer (2012). However,

Plümper and al. (2009) show that the government ability to reduce taxes corporate mobile capital is limited by budget constraints and societal fairness norms. It should be noted, however, that the amounts of taxation remain at a low level, amounting to 3.1% of GDP in 2019, or 8.2% of average tax revenues. The graphical representation of Figure 3 is confirmed by formal tests of beta and sigma convergence, performed in Table A3 and Table A5 in Appendix.

Fact 5: Persistence differences in capital stock taxation

The convergence of the taxation of capital income contrasts with the persistent difference in the taxation of the capital stock, as shown in Figure 4. The coefficient of variation increased by 0.02 percentage points within the period.

This difference indicates that the capital stock is perceived as a factor with little mobility and is therefore likely to be taxed differently across countries. Unfortunately, the data do not allow us to distinguish between taxation of household wealth (e.g. property taxes or wealth tax) and corporate wealth (taxes on production). Taxation of the capital stock is rather low in Germany (1.4% of GDP in 2019) and particularly high in France (4.4%). In this respect, France is close to Great Britain (4.0%)[6].

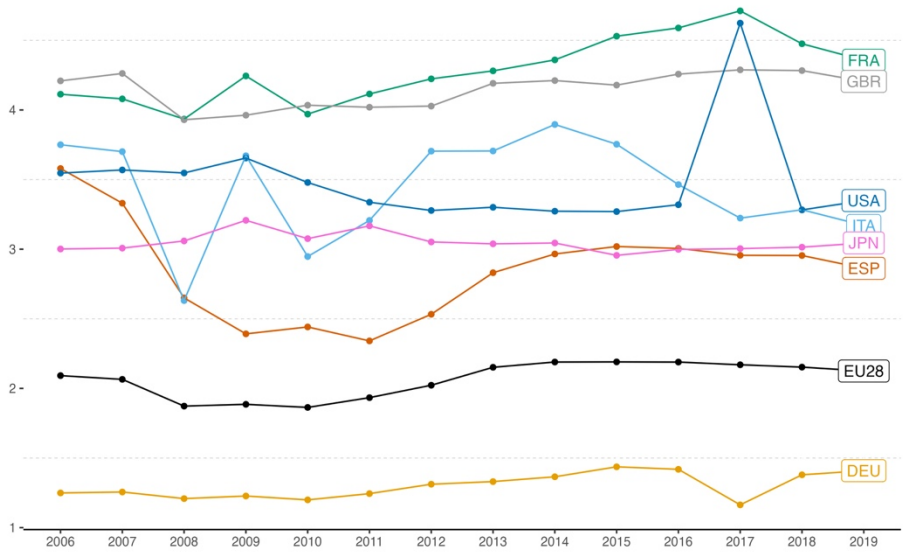

Figure 4: Taxes on capital stocks (corporations and households)

Fact 6: European convergence on consumption taxation

Finally, we find European convergence on consumption taxation. This is the result of a political choice to harmonize the definition of bases and then rates, and is part of the desire to build a single

market and avoid relative price differences for tax reasons alone. European Union has established harmonized rules on VAT, including a standard rate that cannot be less than 15%. This convergence is taking place towards relatively high levels of VAT in Europe, compared to the United States for example.

3. Redistribution of tax systems

The previous section analyzed the state budget and considered different taxes. This section focuses on households only in order to consider the effects of taxes and transfers on income inequality. This analysis captures some political economy aspects of fiscal systems which is not captured by aggregate series (Prasad and Deng (2009)). In this Section, to complete our analysis of fiscal systems, we focus on the redistributive effects of fiscal systems between households, as redistribution can be considered as one of the main objectives of tax policy (Peters (1991)). We consider the socio-fiscal distributivity, thus we include the effects of taxes and transfers on income inequality.

This Section is based on microeconomic data of a different nature from those provided by the national accounts in the previous section. The use of sometimes heterogeneous microdata reduces the number of countries, years and tax instruments that can be analyzed. We rely on the LIS (Luxembourg Income Study) microdata, collected by national sample and harmonized. Among the 29 countries studied in the previous section, the redistributive effects of 17 of them are measured, between 2006 and 2016, some of them over several years, which constitutes a total of 48 observations. The temporal evolution of redistributive effects cannot be studied exhaustively due to the lack of data. However, this database allows us to compare the redistributive effects on household disposable income (by standard of living percentile) of a set of levies and transfers. While this database is among the most exhaustive in international comparison, it suffers from weaknesses that must be identified in order to clarify the limits of the analysis.

First, household survey data provide few information about the very top of income distribution. Using LIS and tax authorities' data, Yonzan and al. (2021) show that there is a discrepancy between household survey data and tax data for the 1 percent top of the income distribution. Therefore, the first and last percentiles are excluded for the analysis (following Amoureux and al. (2018)). Then, in line with the literature, we compare income inequalities before and after taxes and transfers. The limitations of this methodology are well identified. First, it does not allow us to measure the redistributive effects of taxes on consumption or wealth (see Andre and Biotteau (2021), Basco and al. (2021), and Rousselon and Viennot (2020)). Moreover, the reduction

of inequalities does not involve only socio-fiscal redistribution, but also the final consumption of public administrations, or corporate tax policy.

3.1 Methodology

The effect of taxes and transfers is measured by the reduction in the Gini coefficient within the country between income before taxes and transfers (primary income) and income after taxes and transfers (disposable income)[7]. Our measure of primary household income before taxes and transfers is wages plus employee and employer social security contributions, because the issue is to measure the cost of labour generated by taxation. Primary income is therefore super gross (including employee and employer contributions). Employer and sometimes employee contributions are not available in LIS, so we use data from Amoureux and al. (2018), obtained by micro-simulation, to impute social contributions by income percentile[8]. Indeed, as Guillaud and al. (2020) show, the progressivity of taxes also comes from the progressivity induced by social contributions and exemptions on low wages.

Before presenting the results, it is worth focusing on the status of retired households. The difference in pension systems between countries changes the nature of pensions, which are sometimes considered as market income in funded systems or as transfers in pay-as-you-go systems. Since we do not wish to overestimate inequalities before taxes and transfers or to minimize the measurement of the redistributive effects of pensions, we have chosen to exclude the retired population from our samples.

3.2 Results: Analysis of redistributive effects

Table 4 presents the numerical results. The first column shows the country. The second column indicates the date of data collection. The third column presents the Gini coefficient of income before taxes and transfers (primary income). The fourth column is the Gini coefficient of income after transfers and before taxes (gross income). The fifth column is the Gini coefficient of income after taxes and transfers (disposable income). The sixth column presents the reduction in the Gini achieved by taxes and transfers. It corresponds to the percentage change between column (3) and column (5). Column (6) thus measures the reduction in inequality induced by the social and fiscal system.

There is no correlation between primary inequality and the reduction in income inequality. In other words, high primary inequality (before taxes and transfers) is not associated with a greater reduction in inequality. Similarly, we don't find a significant correlation between the amount of primary inequality and the size of the socio-fiscal system, as measured by the tax revenue over GDP. Thus, there is a strong dispersion of primary inequality (before taxes and transfers) and a weak correlation with the tax rate. For example, the Gini index reaches its minimum at 0.27 for Switzerland

and its maximum at 0.53 for Ireland in 2010, which have similar tax revenues over GDP: 24.8% and 27.7% of GDP respectively. The correlation between tax rates and the Gini index of primary income is −21.0%, but not very significant (p-value of 0.1516). We also perform linear regressions of the change in the Gini index on the initial level of inequality and other controls. As presented in Table A7 in Appendix, the coefficient on initial inequality level is rarely significant. Thus, a high reduction in inequality through the social-fiscal system is *not* a corrective instrument for high primary inequality, but rather the result of national preferences for a reduction in inequality. Considering the questions of the introduction, this result is a piece of evidence that social preferences should therefore be considered as an explanatory factor for the heterogeneity of fiscal systems. This is consistent with the political economy literature, where political groups are able to shape the fiscal system in a country specific way (Peters, 1991).

	Year	T	Primary inc. Gini (Before Tr and T)	Gross inc. Gini (After Tr before T)	Disp. Inc. Gini (After Tr and T)	Var. Gini (%)
(1)	(2)	(3)	(4)	(5)	(6)	(7)=(6)/(4)-1
AUT	2013	41.93	21.36	38.18	35.10	29.14
CHE	2013	25.14	13.55	29.35	26.90	25.91
CZE	2013	33.65	13.88	34.68	32.35	25.28
DEU	2013	38.17	16.67	37.54	33.07	26.16
DNK	2013	44.85	18.42	32.61	28.16	25.03
ESP	2013	33.10	17.13	44.87	41.60	37.14
EST	2013	31.67	12.77	40.91	39.12	33.85
FIN	2013	43.41	20.12	35.37	29.60	23.02
FIN	2013	43.41	20.12	35.37	29.60	23.02
FRA	2010	42.47	22.84	34.75	31.49	23.85
GBR	2013	31.99	14.35	43.07	36.46	31.13
GRC	2013	35.24	23.75	42.43	41.23	36.20
IRL	2010	27.72	38.33	53.31	42.50	34.48
ITA	2014	44.10	22.94	41.60	41.20	35.15
LUX	2013	36.18	18.48	39.43	35.67	31.14
NLD	2013	36.12	11.73	44.64	39.88	33.24
SVK	2013	30.75	13.65	35.07	32.75	28.39
USA	2016	25.88	13.17	41.48	39.18	34.80

Table 4: Redistributive effects (last year by country). *T* is for taxes over GDP.

Furthermore, the reduction in inequality induced by the socio-fiscal system is positively correlated with the tax revenue over GDP. The correlation between the latter and the variation in the Gini is 48.5% and is significant at the 0.1% level: the higher the tax rate is, the greater the reduction in the Gini. The following Figure 5 presents these results graphically. On the x-axis is the amount of tax revenues relative to GDP for each country. The ordinate shows the absolute change in the Gini (Column 6).

There is a clear positive correlation between the amount of tax revenues and the reduction of inequalities. We confirm this result with multivariate regressions, adding taxation variables, with the Gini variation as the dependent variable. The results are presented in Table A7 in Appendix. In all regressions, the coefficient associated to tax revenues is highly significant and negative.

The case of Ireland is singular, as noted in other studies (Guillaud and al. (2020), Rousselon and Viennot (2020)). Despite a relatively low tax rate of taxation of 27.7%, Ireland is the second-most inequality-reducing country, with a variation of −35.3% between pre and post tax and transfer Gini. This result is explained by the very high level of primary inequality in Ireland (interdecile ratio of primary income: 23.9, and 7.1 for the European median), by targeted transfers and taxation of income. The interdecile ratio of disposable income is then 3.2 for Ireland and 3.4 for the European median according to Rousselon and Viennot (2020). Thus, without taking Ireland into account in the sample, the correlation between the tax rate and the variation in the Gini reaches 58.2%.

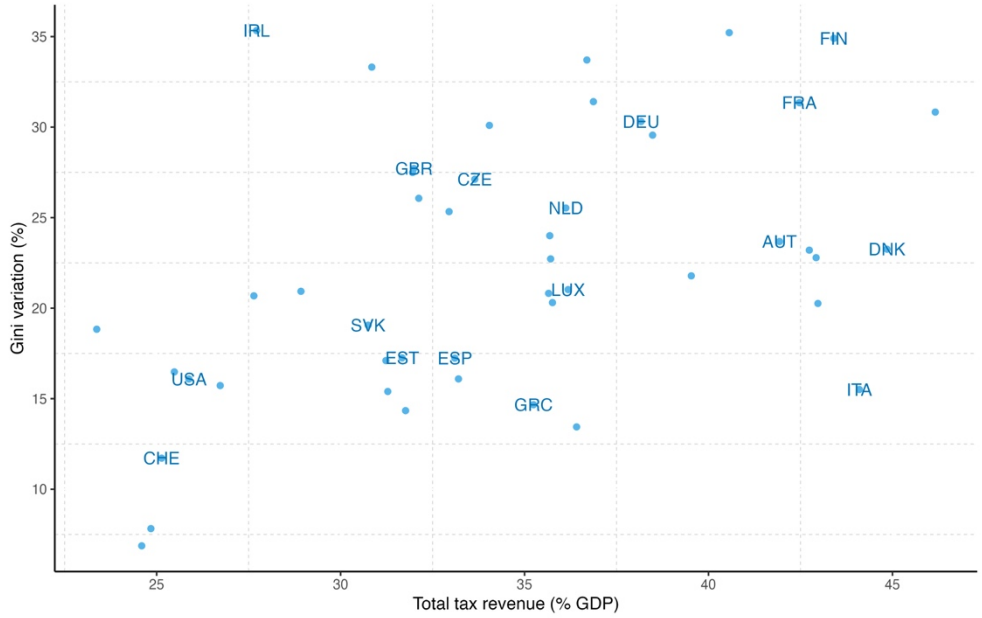

Figure 5: Tax revenues over GDP - Gini variation

4. Coherence of socio-fiscal systems

This Section now provides our identification of types of fiscal systems. As many others, we use statistical tools to identify coherences in an "agnostic" way and then to discuss the economic coherence of the types of fiscal systems identified. We use a machine learning algorithm, which is the $K-means$, explained below, that allows us to perform this type of clustering with a large number of variables[9]. While most works use principal component analysis (PCA) (Amable and al. (1997) and Amable (2004)), we run a clustering analysis, using the K-*means* method, which allows a more direct interpretation of the results. PCA and K-means are closely linked (see Ding and He (2004) for more details) but they have different objectives. Intuitively, PCA seeks to represent the data by reducing its dimensions, while K-means groups the observations around the cluster centroids, based on their similarity. Clustering with the $K-means$ method classifies observations (here countries) into k groups, that contribute in finding patterns and relationships between study variables (here their tax structure). We perform this exercise in 2006 and 2019, to study the convergence or divergence of fiscal systems, similarly to the work of Wang (2007). Moreover, this method provides information about differences between clusters and similarities within groups, which give us insights about the fiscal systems compositions and coherence.

More precisely, the $K-means$ algorithm is an unsupervised non-hierarchical clustering algorithm. It allows to group data (here countries) into a finite number of sets based on the provided characteristics (here variables describing the socio-fiscal systems). Intuitively, from a given number of clusters, the algorithm provides the partition of countries that minimize the distance within clusters (see Kassambara (2017) for a presentation of the algorithm and its implementation). The standard algorithm is the Hartigan-Wong algorithm (1979), which defines the total within-cluster variation as the sum of squared distances Euclidean distances between observations and the corresponding centroid. First, k countries are arbitrarily selected as the initial cluster center. Then, each observation is assigned to a cluster to which it is the closest, in terms of fiscal systems variables, based on its distance to the cluster mean. Finally, new initial cluster centers are again selected. This process is iterated until the updated cluster means are stable. For each exercise, we determine the number of clusters endogenously by using three different statistical methods. The description of these methods and their results are detailed in the Appendix.

We perform this exercise in 2006 and 2019, in order to identify clusters before the 2008 crisis and clusters in 2019. We also perform the exercise with several sets of variables. The main results discussed in this Section are realized with the ten variables presented in Section 2. These are the six taxes identified in the decomposition of taxes, as a share of total taxation, the tax revenues over GDP, the total amount of transfers over GDP and the implicit tax rates on labour and capital income. We

preferred to include the implicit rate on capital income, because it better reflects the reality of the average tax rate (see Appendix).

The use of these macroeconomic variables allows us to consider 29 countries. In Appendix, clusters in 2006 and 2019 are displayed using the PCA representation, to show that our results don't depend on the methodology. We also perform clustering by introducing distributional variables such as inequality reduction, using the LIS data discussed in Section 3. Data availability leads us to consider only 17 countries in different years. As the results are very similar[10], we use the variables that allow us to treat the largest number of countries and present the results with redistribution in Appendix[11]. We also run the cluster analysis for other sets of variables, including the decommodification index of Esping Andersen (1990). In particular, we control for the annual average growth rate over the 2006-2019 period to capture variations in GDP (see Figure A5), which does not change the repartition of countries between clusters.

4.1 Three socio-fiscal systems

The statistical analysis for 2006 and 2019 indicates the existence of two clusters in 2006 and three in 2019. Figure 6 presents the results by indicating the clusters in the same colour and representing the countries in a plane formed by the tax revenues over GDP on the x-axis and the implicit tax rate on labour on the y-axis. The choice of these two variables is made on a statistical basis, as they have the greatest weight in the formation of the groups. Some clusters may overlap when represented on these two dimensions in Figure 6. Indeed, other variables are used in the identification of the clusters.

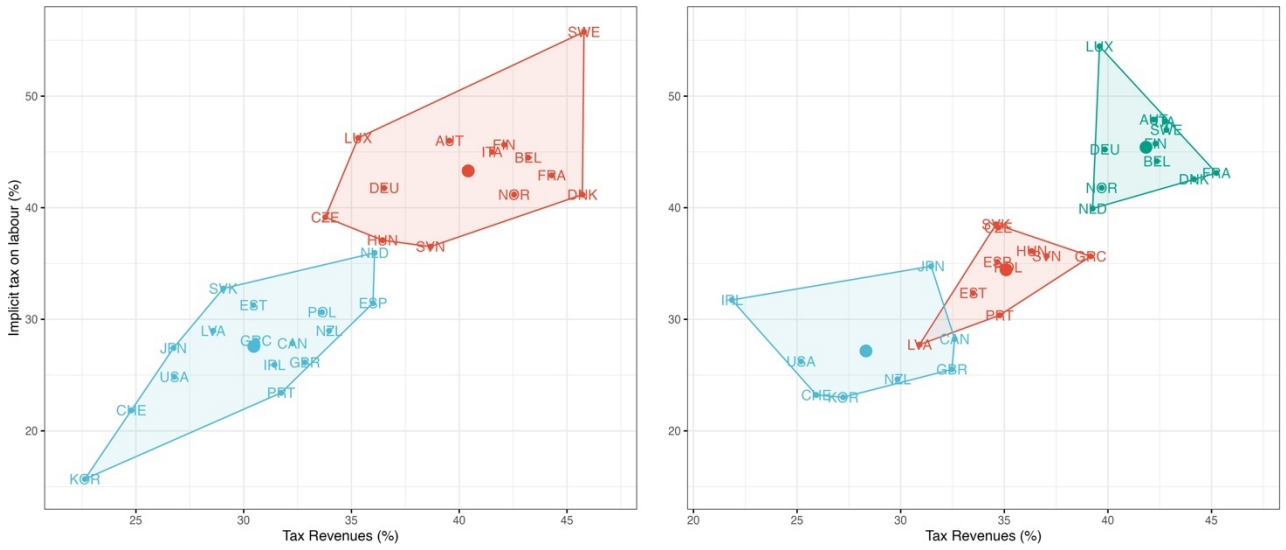

Figure 6: Clustering: Tax and transfer variables

This clustering is robust to the analysis of tax variables alone (without transfers) and to the inclusion of redistribution variables. We find that the main variables that contributed to the formation of clusters are the tax revenues and the implicit tax rate on labour (see details in Appendix). Therefore, the groups of countries are differentiated and represented by these two variables in Figure 6. In 2006, a first group of countries has low values for these two variables (Korea, United States, Canada among others). A second group of countries has high values for these two variables (Sweden, France, Germany, Austria among others). In 2019, a third intermediate group appears, clustering countries with intermediate values for these two variables (Spain, Portugal, Poland among others). To simplify the discussion, we refer to the three groups as "low" (USA, Canada), "high" (Sweden, France, Germany) and "intermediate" (Spain, Poland), designating their amount of tax revenues over GDP in 2019.

The consistency of the groups is greater in 2019 than in 2006 (measured by the percentage of variance explained). This result is due not only to the larger number of groups, but to the dynamics of the countries. Thus, a large number of countries are found with high and close taxation in 2019, like Sweden, France, Belgium and Germany in particular. Low tax countries are also more consistent in 2019, such as the US, UK, Canada and Korea. Similarly, "intermediate" tax countries in 2019 show strong consistency.

As explained in the introduction, the outcome can be compared to the literature on the variety of capitalisms that proposes classifications of countries or welfare state systems. First, our statistical partition partly overlaps with the typology of Esping-Andersen (1990) considering the liberal regime, the social-democratic regime, and the corporatist regime. The liberal regime corresponds to the "low-tax" group, the social-democratic group to the “high-redistribution” group, and the corporatist model to the “intermediate-tax” group (which includes Greece, Spain, Portugal and Poland). However, a more detailed comparison of the clustering leads to important nuances. For example, our statistical classification leads us to consider countries that Esping-Andersen (1990) defines as corporatist in the high-redistributive group, because we here only consider budgetary economic approach. In the literature on the variety of capitalisms, often two subgroups are distinguished within the high-redistribution group. For example, Amable (2004) distinguishes between the social-democratic model and the continental European model. Focusing on the fiscal and tax side only, we find a homogeneous high-redistribution group.

Based on this proximity to the literature on the diversity of social models, we have chosen to call countries with low taxation "liberal", countries with high taxation "high-redistribution" and countries with intermediate taxation "intermediate", with the understanding that these labels are not value judgments and are descriptive.

4.2 Identification of three models of socio-fiscal system

The statistical clustering method must now lead to an economic analysis to understand the economic meaning of the results obtained. In order to show the dynamics of the countries between 2006 and 2019, we use the 2019 clustering, which leads to end-of-period consistency. Beyond simple economic identification, we identify common economic trends.

In Appendix, Tables A8 to A11 report the variables that contributed the most in distinguishing clusters (mainly total tax revenues and implicit tax rate on labour, displayed in Table A8) and the variables that best describe the clusters (Tables A9, A10 and A11 for each cluster). This provides information in identifying clusters. Finally, Table 5 details the clusters' tax structure (all variables are expressed in percentage of GDP. Although the implicit tax rate on labour and tax revenues are highly correlated, it still captures differences between clusters. Indeed, except for Ireland, the liberal cluster have on average lower ITR on labor for a given tax revenues level. On the opposite, the ITR on labour of high-redistributive countries is mainly higher.

First, the three groups are distinguished by the amount of tax revenues over GDP. We now plot the average tax liability rates by country group in 2006 and 2019 (thus considering the 2019 cluster classification). Figure 7 presents the results by showing the tax liability rate by country group, by date (as well as the composition by tax instrument, for the six instruments identified). To our knowledge, there is no independent typology of fiscal systems since Peters (1991), for this reason it may be useful to summarize the main statistical regularities on recent data.

The high-redistribution model is characterized by a high rates of tax revenues, of taxation of labour and of transfers.

The group referred to as high-redistribution includes eleven countries: Norway and ten European countries. Among them Belgium, Denmark, Finland, France, Italy, Sweden, Luxembourg, Germany, the Netherlands and Austria. The characteristics of this group are based on a high rate of tax revenues, high taxation of labour and a high level of transfers.

The liberal model is characterized by relatively high capital taxation, a low tax rate and low labour taxation.

In contrast, the liberal model is characterized by low taxation, particularly on labour and consumption. The low taxation of consumption is explained in particular by the high proportion of non-European countries in this group: The United States, Switzerland, Japan, New Zealand, Canada and Korea. Only Ireland (and the UK) is a member of the European Union. These countries are distinguished by a high taxation of capital stocks as a proportion of total tax revenues, already noted by Amable (2004).

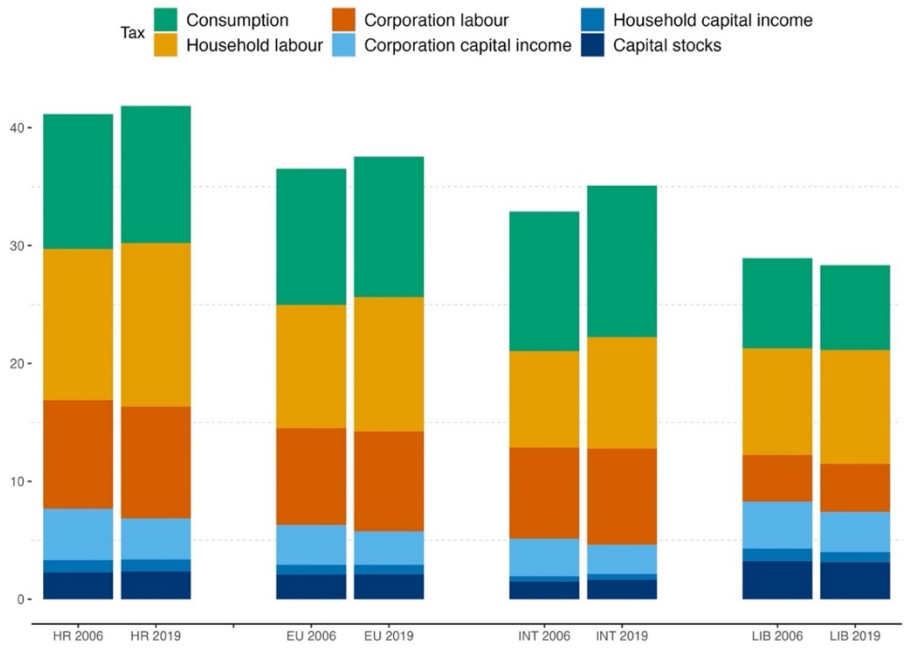

Figure 7: Tax structure by cluster and European Union

	Year	T	C	L hou	L corp SE	K inc hou	K inc corp SE	K sto
EU	2019	37.53	11.92	11.38	8.48	0.80	2.83	2.12
	2006	36.51	11.55	10.45	8.23	0.83	3.36	2.09
HR	2019	28.32	7.20	9.63	4.10	0.86	3.38	3.15
	2006	28.93	7.67	9.03	3.97	1.07	3.96	3.24
INT	2019	35.09	12.87	9.43	8.18	0.48	2.48	1.66
	2006	32.89	11.86	8.17	7.74	0.43	3.17	1.52
LIB	2019	41.84	11.63	13.87	9.50	1.02	3.45	2.36
	2006	41.15	11.45	12.81	9.23	1.03	4.34	2.28

Table 5: Tax variables by cluster and European Union

The intermediate model finds coherence, potentially from the influence of international competition in the tax structure.

The intermediate group is by nature more difficult to characterize as only two clusters are identified in 2006. Indeed, we performed the clustering exercise for each year and find that the optimal number of clusters switched from 2 to 3 for year 2010, with approximatively the same country composition

since. It is composed of ten countries: Estonia, Spain, Greece, Latvia, Poland, Portugal, Slovakia, as well as the Czech Republic, Hungary and Slovenia. The first seven are from the group with the lowest tax burden in 2006, they experienced an increase in tax rates over the period 2006-2019. The three remaining countries the rest belonged to the group with a higher tax burden. The most discriminating variable is the high share of taxes on consumption in the tax revenues. This is explained by the composition of only European countries in the group. Latvia, despite its relatively low implicit tax rate on labour and tax revenues belongs to the intermediate group, due to its taxation structure. This intermediate group includes countries that are relatively small in terms of population, which are more likely to adopt a strategy of tax competition.

4.3 Dynamics of tax systems in the crisis

We now study the temporal evolution of tax rates in order to identify the tax trends of groups. Figure 8 presents for each group and the European Union (28 countries) the dynamics of taxes on consumption, labour and capital (top three graphs) as well as the standard deviation of these three variables between countries and for each year within each group (bottom three graphs), in order to study the stability and dynamics of intra-group heterogeneity[12]. In Appendix, we also run beta-convergence and sigma-convergence tests to study the trends of the tax rates, for the three clusters and the European Union. We now present the main results.

Figure 8: Evolution, by cluster, of consumption, labour and capital taxation and heterogeneity

Stability of the high-redistribution model

The high-redistribution model appears to be a very stable group, with low relative dispersion in consumption and labour taxation. The heterogeneity of capital taxation (income and stocks) is relatively high. However, this group shows strong signs of convergence in corporate capital income taxation and of little divergence in the capital stocks taxation.

Dynamics of the intermediate group

In contrast, the intermediate group appears more heterogeneous than the high-redistribution model. Between 2006 and 2019, the tax ratio increased from 32.9% to 35.1% of GDP, and the standard deviation fell from 3.4 to 2.2. This increase is due to the rise in the tax on consumption (+1.0 points) and labour (+1.7 points), which again offsets the fall in the tax on capital (-0.5 points), which amounted to 4.6% in 2019. This low rate is a feature of the middle group, justifying the analysis of a group of small open economies. But the increases in labour and consumption taxation are leading the intermediate group to move closer to the high-redistribution model. Finally, the intermediate cluster observes a decrease in heterogeneity in the capital taxation, mostly driven by the taxation ofhousehold capital income and to some extent by the corporate capital income (on the implicit tax rate). To summarize, this intermediate group is composed of countries roughly experiencing the same trends as the high redistribution group, but at a smaller scale. Hence, this group may either catch-up with the high-redistribution group or it may compose a new group with an intermediate level of redistribution.

Heterogeneity but coherence of the liberal model

This group of countries is characterized by an overall decrease in average tax rates in both consumption and capital. Most of the countries in the group have either stagnated or lowered their tax rate slightly, with the exception of Ireland, which has lowered its tax revenues over GDP by nine points over the period[13]. Only the labour taxation has increased, but with a relatively high heterogeneity. Thus, while the countries did not move much closer together over the period, each of them (with the possible exception of Japan and Korea) moved towards a continuation of the logic of low taxation. Moreover, as pointed out in Fact 5 in the Section 2, the liberal group experienced a particularly large divergence on the capital stock taxation as the coefficient of variation has increased by 0.10.

A common trend: substitution of the tax from capital to labour?

It is worth noting that all groups have decreased their taxation of capital (-0.7 points), which comes mainly from the decrease in taxation on corporate capital income. However, this was more than

compensated by an increase in taxation on labour in the high-redistribution (+1.2 points) and intermediate (+1.7 points) models. In the liberal model, on the other hand, labour taxation increased slightly (+0.7 points). This recent trend confirms the finding of Carey and Tchilinguirian (2000) that there was a shift from capital to labour taxation in the 1990s.

4.4 The beginning of European convergence?

The European Union countries we consider belong to all three models. Great Britain and Ireland belong to the liberal model. Portugal, Greece and Spain belong to the intermediate model, while Germany, France and Austria correspond to the high-redistribution model. However, taken as a whole, Europe is closer to the high-redistribution model than to the liberal model, as can be seen in Figure 7. European countries have a similar tax structure. First of all, they have, on average, a high consumption tax. This is the result of the desire for harmonization within the single market, and is therefore more political than economic in origin, but it is a powerful factor for convergence. Second, European countries are characterized by high labour taxes (52.9% of tax revenues in 2019). Figure 8 shows that the heterogeneity of the European fiscal system concerns mainly the taxation of labour.

These European trends do not fully capture the diversity of capitalisms in Europe. France and Europe have relatively close fiscal systems, but they differ in another important economic dimension. Indeed, comparative political economy contributions such as Iversen and Soskice (2012) or Iversen, Soskice, Hope (2016) have convincingly showed that Coordinated Market Economies (typically Germany), had an export-led growth model, whereas other member states of the European Monetary Union (typically France) had a demand driven growth. This difference is generating destabilizing economic and political tensions. From this ascertainment, we conclude that the proximity of fiscal systems is neither the cause nor the consequence of similar growth strategy. This may be an additional piece of evidence of the relative autonomy of social political preferences from economic strategies in the design of fiscal systems.

Finally, European heterogeneity is increased by the inclusion of the United Kingdom and Ireland. The exit of the UK from the European Union in 2021 therefore mechanically increased the proximity of the remaining countries of the European Union.

5. Conclusion

The analysis of the proximity of fiscal systems confirms the existence of three groups of countries in 2019, qualified as high-redistribution, liberal and a third group of countries between these two models, qualified as intermediate, which are consistent with previous typology of welfare states or

capitalisms. The evolution of fiscal systems can be the effect of economic constraint within each type of capitalism, or, more directly, the expression of "culture" or social preferences mediated by heterogeneous political institutions. The suggestive evidence of the current paper favors the second explanation. First, the amount of redistribution within each country is independent of the initial level of inequality. Second, we identify a divergence of these types of fiscal systems. The liberal model pursues a logic of a low tax burden in the economy. The high-redistribution model is more consistent, and most of the countries in this group have stable or increasing size of their welfare states. The trajectory of the UK seems an example of a political agenda driving fiscal change. Third, the proximity of the fiscal systems of France and Germany, whereas the two countries have different growth model, demand-driven and export-led, shows that the same fiscal system can be consistent with different growth models. Finally, in the comparative political economy literature, many papers study political changes considering various economic constraints, whereas economists tend to consider economic forces toward efficiency considering political constraints. The first approach may be the right one to understand the dynamics of fiscal systems, and maybe the diversity of capitalism.

References

Amable, B. (2004) *The Diversity of Modern Capitalism*, Oxford University Press.

Amable, B., R. Barré, and R. Boyer (1997), *Les systèmes d'innovation - À l'ère de la globalisation*, Paris, Economica.

Amoureux, V., E. Guillaud, and M. Zemmour (2018) 'Le financement par cotisations freine-t-il la redistribution ? Une analyse en comparaison internationale', *Revue française des affaires sociales*, 34.

Andre, M. and A.-L. Biotteau (2021) 'Medium-Term Effects of a Rise in VAT on Standard of Living and Inequality: a Microsimulation Approach', *Economie et Statistique/Economics and Statistics,* 522-523, 5-21.

Atkinson, A. B. and J. E. Stiglitz (1976) 'The design of tax structure: Direct versus indirect taxation', *Journal of Public Economics*, 6, 55–75.

Boyer, R. (2007) 'Comment concilier solidarité sociale et efficacité économique à l'ère de la globalisation : une lecture régulationniste', in *Repenser la solidarité. L'apport des sciences sociales*, ed. by S. Paugam, PUF, 887–914.

Blasco, J., Guillaud, E. and Zemmour, M. (2021) "The Impact of Consumption Taxes on Income Inequality: An International Comparison," *LIS Working papers* 785, LIS Cross-National Data Center in Luxembourg

Carey, D. and H. Tchilinguirian (2000) 'Average Effective Tax Rates on Capital, Labour and Consumption', *OECD Economics Department Working Papers*, 258.

Causa, O. and M. Hermansen (2018) 'Income Redistribution Through Taxes and Transfers across OECD Countries', *LIS Working paper Series*.

DG Taxation and Customs Union (2020) 'Taxation Trends in the European Union', Report, European Commission.

Ding, C., and He, X. (2004, July). K-means clustering via principal component analysis. In Proceedings of the twenty-first international conference on Machine learning (p. 29).

Esping-Andersen, G. (1990), *The Three Worlds of Welfare Capitalism*, Princeton University Press.

Fullerton, D. and G. E. Metcalf (2002) 'Tax Incidence', in *Handbook of Public Economics*, vol. 4, 1787–1872.

Genschel, P. and P. Schwarz (2011), "Tax competition: a literature review", *Socio-Economic Review*, Volume 9, Issue 2, April 2011, Pages 339–370,

Guillaud, E., M. Olckers, and M. Zemmour (2020) 'Four Levers of Redistribution: The Impact of Tax and Transfer Systems on Inequality Reduction', *Review of Income and Wealth*, 66, 444–466.

Hakelberg, L. and P. Rixen (2021) Is neoliberalism still spreading? The impact of international cooperation on capital taxation, *Review of International Political Economy,* 28:5, 1142-1168, DOI: 10.1080/09692290.2020.1752769

Hall, P. and D. Soskice (2001), *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*, Oxford University Press.

Hasse, A., B. Palier, and S. Avlijas (2020) 'The pursuit of growth. Growth regimes, growth strategies and welfare reforms in advanced capitalist economies', *Stato e mercato*.

Ihori, T., 2017, *Principal of Public Finance*, Springer, 2017.

Iversen, T. and Soskice D. (2012), "Modern Capitalism and the Advanced Nation State: Understanding the Causes of the Crisis », in « Coping with Crisis: Government Reactions to the Great Recession", Bermeo N. and Pontusson, J. Eds (2012), Russel Sage Foundation, New York.

Iversen, T. Soskice, D. and David Hope (2016), "The Eurozone and Political Economic Institutions », Annu. Rev. Polit. Sci. 2016. 19:163–85.

Jessop, B. (2012) 'Rethinking the diversity of capitalism: varieties of capitalism, variegated capitalism, and the world market', in *Capitalist diversity and diversity within capitalism*, ed. by G. Wood and C. Lane, London:Routledge, 209–237.

Kassambara, A. (2017) 'Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning: Volume 1 (Multivariate Analysis)', CreateSpace Independent Publishing Platform.

Kemmerling, A. (2011), 'From Convergence to Diffusion: The EU's Influence on National Tax Systems', Kemmerling, "EU's Influence on National Tax Systems," ACES Cases 2011.

Korpi, W. and Palme, J. (1998) The Paradox of Redistribution and Strategies of Equality: Welfare State Institutions, Inequality, and Poverty in the Western Countries, *American Sociological Review,* Vol. 63(5), pp. 661-687

Lindert, Peter H. 2004. Growing Public, Volume 1: Social Spending and Economic Growth since the Eighteenth Century. Cambridge: Cambridge University Press.

Mendoza, E., A. Razin, and L. Tesar (1994) 'Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption', *Journal of Monetary Economics*, 34, 297–323.

Mirrlees, J. A. (1971) 'An Exploration in the Theory of Optimum Income Taxation', *Welfare, Incentives, and Taxation*, 38.

Musgrave, R. A. (1981) *Fiscal Systems*, Yale University Press.

OECD, (2022), 'Tax Policy Reforms 2022', OECD,7th Annual report, September 2022.

Piketty, T. (2014) '*Capital in the 21st century*', Harvard University Press, 698.

Plümper, T., V. E. Troeger, and H. Winner (2009): "Why is there no race to the bottom in capital taxation?" *International Studies Quarterly*, 53, 761–786.

Prasad, M. and Y. Deng (2009) 'Taxation and the worlds of welfare', *Socio-Economic Review*, 7, 431–457.

Rousselon, J. and M. Viennot (2020) 'Primary inequality and redistribution: where does France stand in Europe?', *La note d'analyse,* 77, 1-16.

Schneider, M. R. and M. Paunescu (2012) 'Changing varieties of capitalism and revealed comparative advantages from 1990 to 2005: a test of the Hall and Soskice claims', *Socio- Economic Review*, 10, 731–753.

Swank, D. (2006) 'Tax Policy in an Era of Internationalization: Explaining the Spread of Neoliberalism', International Organization, 60, 847 –882.

Tibshirani, R., G. Walther, and T. Hastie (2001) 'Estimating the number of clusters in a data set via the gap statistic', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63, 411–423.

Yonzan, N., Milanović, B., Morelli, S. and Gornick, J. C. (2020) : Drawing a line: Comparing the estimation of top incomes between tax data and household survey data, *LIS Working Paper Series,* No. 805, Luxembourg Income Study (LIS), Luxembourg

Evolution of fiscal systems: Convergence or divergence?

Appendix

This appendix presents the sources and construction of the data. All of the databases constructed are provided in the separate file, available upon request.

1 Breakdown of taxation

1.1 Database

We use mainly OECD data, to allow for international comparison. Tax data are taken from Revenue Statistics 2022, and national accounts data from National Accounts 2022. Some have been supplemented for European countries by Eurostat accounting data, when OECD data were not available. In particular, data on the distribution of taxes on household income between labour, capital, and the self-employed is taken from the European Union's Taxation Trends report (2022).

The OECD tax revenue data are net of tax credits (unlike the Eurostat data following the revision of the ESA 2010 accounts). They are considered as revenue shortfalls and, in case the tax credits exceed the tax amount, the difference is recorded in government expenditure.

The OECD classification of taxes is described below.

1.2 Structure of the State budget

Government resources come from tax revenues (T) and the budget deficit (noted D). Government expenditures are composed of government final consumption expenditures (health, education, civil servants' salaries, etc.), noted G, interest payments on government debt rB (where r is the apparent interest rate on government debt and B is the total amount of government debt) and transfers to households and firms, noted Tr. These transfers are measured net of taxes and other contributions. For each country i and each year t, we can therefore represent the government budget constraint as

$$G_{it} + r_{it}B_{it} + Tr_{it} = T_{it} + D_{it}$$

Then, the tax rates T can be decomposed according to the contribution of different taxes. We decompose the fiscal system by distinguishing between taxes on the factors of production, i.e. capital and labour, and on consumption. Next, we decompose taxes by the payers, i.e. firms or households. Finally, for the taxation of capital, we distinguish between the taxation of capital income (e.g. tax on dividends) and the taxation of capital stock (e.g. tax on real estate). Finally, consumption taxation

concerns only households. We therefore consider six taxes: consumption tax (T_C), tax on labour paid by households (T_{L_hou}) or by corporations and the self-employed ($T_{L\_corp\_SE}$), tax on capital income paid by households ($T_{K\_inc\_hou}$) or by corporations and the self-employed ($T_{K\_inc\_corp\_SE}$) and tax on capital stock (T_{K_sto}), as in the following Equation.

$$T = T_C + T_L + T_K = T_C + T_{L\_hou} + T_{L\_corp\_SE} + T_{K\_inc\_hou} + T_{K\_inc\_corp\_SE} + T_{K\_sto}$$

Tables A1a and A1b present the six taxes by country, in 2006 and 2019. All variables are expressed in percentage of GDP.

1.3 Methodology

The calculation of implicit rates (ITR) consists of associating tax revenues (the numerator) with their tax base (the denominator). It allows a better assessment of the tax burden on capital, labour and consumption. In the literature, they are also called average effective tax rates (AETR). In our study, we are mainly interested in tax yields, i.e. tax revenues (numerator) relative to GDP.

This is a method initiated by Mendoza and al. (1994), then completed by the OECD from Carey and Tchilinguirian (2000). It is used annually by the European Commission in the Taxation Trends report (2020). We have used the methods of Carey and Tchilinguirian (2000) and the Taxation Trends report (2020) in a complementary way, so that each compensates for their respective limitations.

However, this combination raises the problem of the different classification of taxes: Revenue Statistics adopts the OECD classification used by Carey and Tchilinguirian (2000), while Taxation Trends uses the ESA 2010 classification (European system of national and regional accounts). It has therefore sometimes been necessary to adjust the allocation of some taxes between labour, capital and consumption. For example, we have allocated tax 6200 to consumption, as indicated by its "translation" into national accounts in the Taxation Trends report. The 6100 tax is not allocated anywhere in either method but is paid by firms. These amounts were large in some countries, so we chose to allocate it between the labour and capital taxes of firms, without changing the ratios, to ensure aggregate consistency. Similarly, the 2400 tax is a labour tax, which we split between household, corporation, and self-employment labour taxes. One of the main limitations of the Carey and Tchilinguirian (2000) method is the allocation of income taxes between labour and capital. The Taxation Trends report (2020) addresses this limitation by providing micro-data on the distribution of taxes on labour, capital and self-employment income (part_L, part_K and part_SE) by year and by country. The European average is attributed to the missing countries. We follow Carey and Tchilinguirian (2000) to allocate the different taxes on the self-employed between labour and capital,

rather than attributing them only to capital. Finally, the denominators are those described in the Taxation Trends report, which also shows the decomposition of the numerators between firms and households.

The implicit consumption tax rate is as follows:

$$ITR_C = \frac{5000 - (5124 + 5125 + 5127 + 5212 + 5213) + 6200}{P3\_S14dom + P31\_S15 + P3\_S13 - P3\_D1PAY}$$

The implicit tax rate on labour is as follows, where the numerator is decomposed into taxation of households, corporations and self-employed workers:

$$ITR_L = \frac{(\text{part\_L} * 1100 + 2100 + 2300\text{CN}) + (3000 + 2200 + 6100\_\text{L}) + (\text{part\_SE\_L} * (\text{part\_SE} * 1100 + 2300\text{CS})) + 2400}{P3\_S14dom + P31\_S15 + P3\_S13 - P3\_D1PAY}$$

The implicit capital tax rate is the most complex, and we have made changes to it, which we describe below.

1.4 Modification of the construction of the implicit tax rate on capital

The two studies on which we rely point to the important limitations of the implicit tax rate on capital and the problems of its international comparison. Indeed, capital taxation is complex (double taxation, tax credits, wealth tax...) and the construction of a tax base is very imperfect. To try to improve it, we rely once again on the method presented in the OECD, Carey and Tchilinguirian (2000), then on the modifications made by the European Commission in the report Taxation Trends (2020), and on the recommendations of France Stratégie (2020). First, the choice of OECD data rather than Eurostat data allows us to obtain tax revenues "net" of tax credits, i.e., they are taken into account as less tax revenue and not as transfers to firms. The use of Eurostat's micro-data for the distribution of income between labour and capital greatly improves on the OECD's original method, which noted the inaccuracies of this distribution with macroeconomic variables. Instead of attributing all taxes on the self-employed to capital, we attribute, according to the OECD method, the share of self-employed with higher-than-average income to capital, and the rest to labour. In our sample, the share attributed to labour is the largest, which seems consistent because the self-employed include liberal professions (doctor, lawyer), with a higher salary than if they were employees, but above all precarious workers, who are less favored than if they were employees. Concerning the tax base, we add a suggestion from France Stratégie (2020) that it would be more representative to consider as a base the sum of capital income received by households and received by foreigners, rather than the sum of net income received by institutional sectors. However, this measure is still very imperfect, especially for taxes on capital stocks. Indeed, it is very difficult to associate a tax base with these taxes, so that countries with high

stock taxes see their implicit rate overvalued, or at least disconnected from reality. This is why we have preferred to use the implicit tax rate on capital income.

Thus, the implicit rate of capital taxation is as follows, where the numerator is decomposed into taxation of the capital income of households, firms and the self-employed and taxation of capital stocks:

$$ITR_K = \frac{(\text{part\_K} * 1100) + (1200) + (\text{part\_SE\_K} * (\text{part\_SE} * 1100 + 2300CS)) + (4000 + 5124 + 5125 + 5127 + 5212 + 5213)}{B2N\_S11 + B2N\_S12 + B2N\_S14\_15 + D4R\_S14\_15 + D42R\_S13 + D42R\_S2 + part\_SE\_K\} * B3}$$

The implicit rate of taxation on capital income is identical, except that it omits taxes on capital stocks:

$$ITR_{K\_inc} = \frac{(\text{part\_K} * 1100) + (1200) + (\text{part\_SE\_K\}} * (\text{part\_SE} * 1100 + 2300CS))}{B2N\_S11 + B2N\_S12 + B2N\_S14\_15 + D4R\_S14\_15 + D42R\_S13 + D42R\_S2 + part\_SE\_K\} * B3}$$

Then, the implicit rate of taxation on household and self-employed capital income is isolated (here, the tax base cannot distinguish between households and self-employed):

$$ITR_{K\_inc\_hou\_SE} = \frac{(\text{part\_K} * 1100) + (\text{part\_SE\_K\}} * (\text{part\_SE} * 1100 + 2300CS))}{BB2N\_S14\_15 + D4R\_S14\_15 + D42R\_S13 + D42R\_S2 + part\_SE\_K\} * B3}$$

Finally, the implicit rate of taxation on corporate capital income can be further decomposed:

$$ITR_{K\_inc\_corp} = \frac{(1200)}{B2N\_S11 + B2N\_S12}$$

This methodology allows us to disentangle the effects of the decrease in the tax revenues by separating the evolution of the tax base and tax rate. Table A2 displays the evolution of the tax yields (numerator), the tax base (denominator) and the implicit tax rate, which is the ratio of the two. This complements the analysis of the downward convergence of the corporate capital income taxation discussed in Section 2. However, in this Table, only the corporate capital income taxation is considered, excluding the capital income taxation of the self-employed, in order to consider a consistent tax base.

This shows that the decrease in the corporate capital income tax yields is explained by both a decline in the tax base and in the tax rate. Indeed, the tax base decreased by 1.7 percentage points between 2006 and 2019. The ITR allows to control for the pure effect of tax base. It has decreased by 2.0 percentage points, which confirm the downward convergence of the corporate capital income taxation.

	Tax yields (% of GDP)	Tax base (% of GDP)	Implicit tax rate (%)
2006	3.62	16.20	23.52
2007	3.62	16.41	23.30
2008	3.37	15.53	22.38
2009	2.70	13.77	20.78
2010	2.77	15.03	19.46
2011	2.82	15.23	19.84
2012	2.80	14.92	20.12
2013	2.80	14.97	20.35
2014	2.74	15.24	19.46
2015	2.76	15.39	19.48
2016	2.85	15.01	20.44
2017	2.93	15.21	20.59
2018	3.09	14.92	21.61
2019	2.93	14.43	21.46

Table A2: Taxation on corporate capital income – Total sample

1.5 Convergence tests

In Section 2 and 4, we assess the dynamics in tax rates with beta and sigma convergence tests. For more details on the methods, see Kimbala (2020).

Beta-test

Beta-convergence is a concept originally used by Barro and Sala-i-Martin (1990) to evaluate the convergence over time in levels of per capita income and product. Intuitively, it tests whether an observation far from the mean will get closer to it faster than a closer observation. It is measured with a linear regression. The variation of the tax rate between 2006 and 2019 is regressed over the initial level of the tax rate, in 2006. There is potential convergence if the coefficient (beta) is negative and statistically significant.

Table A3 displays the results on the overall sample of countries, whereas Tables A4a to A4d present the results for each cluster and the European Union countries.

Sigma-test

However, beta-convergence test has some limitations (see Monfort (2008)). We use sigma-convergence test to complete our analysis. Sigma-convergence simply refers to a reduction of disparities among regions in time. The most frequently used measures of sigma-convergence are the standard deviation or the coefficient of variation. The coefficient of variation is a normalized measure of dispersion of a probability distribution, defined as the ratio of the standard deviation to the mean.

Table A5 displays the results on the overall sample of countries, whereas Tables A6 present the results for each cluster and the European Union countries.

Variable	beta	p.value
T	-0.016	0.075
hou	-0.017	0.040
corp SE	-0.010	0.264
C	-0.003	0.620
L	-0.014	0.025
K	-0.012	0.118
K inc corp SE	-0.006	0.787
K inc hou*	-0.006	0.386
K sto	-0.001	0.832
L hou	-0.016	0.007
L corp SE	0.005	0.387
ITR K	-0.012	0.161
ITR K inc	-0.017	0.098
ITR K inc corp	-0.015	0.427
ITR K inc hou SE**	0.003	0.748
ITR L	-0.018	0.002

Table A3: Beta-convergence test – Total sampe

	Variable	beta	p.value
EU	T	-0.016	0.221
	hou	-0.017	0.182
	corp SE	-0.010	0.351
	C	-0.011	0.576
	L	-0.011	0.105
	K	-0.006	0.501
	K inc corp SE	0.018	0.599
	K inc hou*	-0.004	0.617
	K sto	-0.002	0.774
	L hou	-0.008	0.167
	L corp SE	-0.002	0.830
	ITR K	-0.012	0.219
	ITR K inc	-0.016	0.183
	ITR K inc corp	-0.016	0.456
	ITR K inc hou SE**	0.007	0.522
	ITR L	-0.019	0.007

Table A4a: Beta-convergence test – European Union

	Variable	beta	p.value
HR	T	-0.046	0.000
	hou	-0.027	0.009
	corp SE	-0.014	0.257
	C	-0.020	0.163
	L	-0.033	0.000
	K	-0.024	0.036
	K inc corp SE	-0.035	0.024
	K inc hou	0.012	0.397
	K sto	-0.003	0.686
	L hou	-0.025	0.045
	L corp SE	0.004	0.626
	ITR K	-0.009	0.549
	ITR K inc	-0.016	0.266
	ITR K inc corp	-0.048	0.001
	ITR K inc hou SE	-0.005	0.578
	ITR L	-0.038	0.034

Table A4b: Beta-convergence test – High-redistribution cluster

	Variable	beta	p.value
INT	T	-0.055	0.006
	hou	-0.012	0.541
	corp SE	-0.040	0.078
	C	-0.012	0.56
	L	-0.043	0.004
	K	-0.003	0.844
	K inc corp SE	0.034	0.598
	K inc hou*	-0.004	0.553
	K sto	0.009	0.494
	L hou	0.003	0.754
	L corp SE	-0.034	0.053
	ITR K	-0.022	0.120
	ITR K inc	-0.023	0.245
	ITR K inc corp	-0.012	0.751
	ITR K inc hou SE**	0.060	0.008
	ITR L	-0.043	0.016

Table A4c: Beta-convergence test – Intermediate cluster

* Latvia and Slovakia are excluded because of zero values.
** Latvia is excluded because of zero values

	Variable	beta	p.value
LIB	T	-0.056	0.100
	hou	-0.042	0.082
	corp SE	0.007	0.888
	C	-0.022	0.203
	L	-0.036	0.185
	K	-0.072	0.126
	K inc corp SE	-0.048	0.283
	K inc hou	-0.039	0.079
	K sto	0.027	0.583
	L hou	-0.036	0.042
	L corp SE	0.029	0.097
	ITR K	0.045	0.293
	ITR K inc	0.005	0.921
	ITR K inc corp	0.014	0.792
	ITR K inc hou SE	-0.036	0.005
	ITR L	-0.046	0.058

Table A4d: Beta-convergence test – Liberal cluster

	2006		2019		
Variable	Sd	CV	Sd	CV	DifCV
T	6.339	0.181	6.100	0.170	-0.011
hou	4.892	0.220	4.696	0.199	-0.021
corp SE	3.625	0.285	3.294	0.270	-0.015
C	2.682	0.254	2.987	0.276	0.021
L	4.831	0.277	4.583	0.245	-0.032
K	2.646	0.381	2.159	0.347	-0.034
K inc corp SE	1.930	0.503	1.209	0.391	-0.113
K inc hou	0.581	0.697	0.502	0.636	-0.061
K sto	1.139	0.499	1.214	0.520	0.021
L hou	3.147	0.310	3.044	0.273	-0.037
L corp SE	3.332	0.459	3.385	0.448	-0.011
ITR K	7.684	0.388	7.607	0.399	0.012
ITR K inc	5.673	0.410	4.557	0.368	-0.041
ITR K inc corp	9.346	0.397	8.678	0.404	0.007
ITR K inc hou SE	7.033	0.789	5.726	0.634	-0.156
ITR L	9.280	0.268	8.438	0.231	-0.037

Table A5: Sigma-convergence test – Total sample

		2006		2019		
	Variable	Sd	CV	Sd	CV	DifCV
EU	T	5.419	0.148	5.401	0.144	-0.005
	hou	4.531	0.194	4.463	0.180	-0.014
	corp SE	3.481	0.265	3.247	0.254	-0.011
	C	1.654	0.143	2.108	0.177	0.034
	L	4.715	0.252	4.399	0.221	-0.031
	K	2.281	0.363	2.226	0.387	0.024
	K inc corp SE	0.935	0.278	1.116	0.395	0.117
	K inc hou	0.641	0.773	0.565	0.709	-0.063
	K sto	1.189	0.568	1.266	0.596	0.028
	L hou	3.355	0.321	3.368	0.296	-0.025
	L corp SE	3.141	0.382	3.144	0.371	-0.011
	ITR K	7.927	0.431	7.652	0.427	-0.003
	ITR K inc	5.490	0.424	4.437	0.378	-0.046
	ITR K inc corp	9.811	0.438	7.865	0.396	-0.042
	ITR K inc hou SE	7.820	0.843	6.332	0.664	-0.179
	ITR L	8.411	0.227	7.385	0.189	-0.038

Table A6a: Sigma-convergence test –European Union

		2006		2019		
	Variable	Sd	CV	Sd	CV	DifCV
HR	T	3.786	0.092	1.987	0.048	-0.045
	hou	4.732	0.182	3.647	0.134	-0.048
	corp SE	4.091	0.270	3.216	0.221	-0.049
	C	1.765	0.154	1.479	0.127	-0.027
	L	3.697	0.168	2.520	0.108	-0.060
	K	3.169	0.414	2.212	0.323	-0.090
	K inc corp SE	2.847	0.655	1.357	0.393	-0.262
	K inc hou	0.599	0.583	0.568	0.556	-0.027
	K sto	1.164	0.509	1.203	0.509	0.000
	L hou	3.020	0.236	2.732	0.197	-0.039
	L corp SE	3.548	0.385	3.499	0.368	-0.016
	ITR K	9.341	0.456	8.929	0.452	-0.005
	ITR K inc	7.179	0.489	5.116	0.386	-0.103
	ITR K inc corp	10.784	0.425	5.357	0.217	-0.207
	ITR K inc hou SE	8.011	0.728	7.731	0.682	-0.046
	ITR L	4.877	0.110	3.929	0.087	-0.024

Table A6b: Sigma-convergence test – High-redistribution cluster

		2006		2019		
	Variable	Sd	CV	Sd	CV	DifCV
INT	T	3.384	0.103	2.177	0.062	-0.041
	hou	2.485	0.120	2.944	0.127	0.007
	corp SE	2.633	0.217	2.113	0.178	-0.039
	C	1.549	0.131	1.829	0.142	0.012
	L	2.680	0.168	1.696	0.096	-0.072
	K	1.956	0.382	1.812	0.393	0.011
	K inc corp SE	1.080	0.341	1.133	0.458	0.117
	K inc hou	0.350	0.809	0.323	0.679	-0.130
	K sto	0.930	0.612	1.158	0.697	0.085
	L hou	1.589	0.195	1.875	0.199	0.004
	L corp SE	2.135	0.276	1.925	0.235	-0.040
	ITR K	7.341	0.435	5.881	0.359	-0.076
	ITR K inc	5.577	0.447	3.804	0.347	-0.100
	ITR K inc corp	10.391	0.487	8.473	0.502	0.015
	ITR K inc hou SE	3.732	0.615	3.980	0.547	-0.069
	ITR L	4.699	0.147	3.386	0.098	-0.049

Table A6c: Sigma-convergence test – Intermediate cluster

		2006		2019		
	Variable	Sd	CV	Sd	CV	DifCV
LIB	T	4.211	0.146	3.898	0.138	-0.008
	hou	4.186	0.221	3.368	0.178	-0.044
	corp SE	1.336	0.133	2.285	0.243	0.110
	C	2.839	0.370	2.507	0.348	-0.022
	L	2.320	0.179	2.670	0.194	0.016
	K	1.099	0.133	1.262	0.171	0.038
	K inc corp SE	0.871	0.220	0.851	0.252	0.032
	K inc hou	0.574	0.539	0.425	0.495	-0.044
	K sto	0.522	0.161	0.836	0.266	0.104
	L hou	2.269	0.251	1.839	0.191	-0.06
	L corp SE	1.330	0.335	1.870	0.456	0.120
	ITR K	4.487	0.198	7.481	0.350	0.152
	ITR K inc	3.314	0.230	4.748	0.367	0.137
	ITR K inc corp	5.662	0.239	10.946	0.479	0.240
	ITR K inc hou SE	8.325	0.867	3.416	0.423	-0.444
	ITR L	4.284	0.172	4.184	0.154	-0.018

Table A6d: Sigma-convergence test – Liberal cluster

2 Measuring redistributive effects

2.1 Database

In order to measure the redistributive effects of socio-fiscal systems, we use the Luxembourg Income Study data, which is a harmonized ex-post micro-data base derived from national surveys. This database contains data by households and individuals on the different incomes (labour, capital, private and public transfers) and levies (income tax, employee contributions, sometimes capital taxes). However, it is sometimes incomplete, as employer and sometimes employee contributions are not collected. Guillaud and al. (2020) manage to complete the social contributions at the individual level by micro-simulation, using OECD data on the rules for tax rates by country. This allows us to cover almost 54% of total tax revenues, compared to 35% with the initial LIS data. We therefore use the tax rate per income percentile from these data, calculated in Amoureux and al. (2018).

We do not include taxes on consumption, even though it has a significative impact on the progressivity of the tax systems, as it is assessed by Prasad and al. (2009). Indeed, on our sample, there is not enough data regarding consumption expenditure of the households. However, Basco and al. (2021) use the same database with a broader set of observations and run simulations. They show that regressivity of consumption taxes are determined by the implicit tax rate. Nonetheless, the negative effect of consumption taxes on redistribution is largely compensated by the positive effects of tax and transfers. Therefore, we only assess for the reduction in income inequality.

2.2 Methodological choices for household income

In LIS, we process data from 17 countries, some over multiple years, between 2006 and 2016, for a total of 48 datasets. For each dataset, we exclude the retired population, selecting only individuals between 25 and 55 years old. The difference in pension systems between countries changes the nature of pensions, which are sometimes considered as market income in funded systems or as transfers in pay-as-you-go systems. Since we do not wish to overestimate inequalities before taxes and transfers or to minimize the measurement of the redistributive effects of pensions, we have chosen to exclude the retired population from our samples. The transfers therefore do not include the various benefits for the elderly. Then, we group the individuals by households, and calculate the different incomes per consumption unit, according to the OECD method. We distinguish primary income (also called market income), before taxes and transfers, then income before taxes after transfers, and disposable income, after taxes and transfers.

2.3 Multivariate regression on Gini and taxation variables

We test the relationship between variation of Gini and tax revenues, and control for other variables with multivariate regressions. In Table A7, we run several OLS regression without intercept over the variation of Gini. *Dgini* is the variation between the Gini market income (before tax and transfer) and the Gini disposable income (after tax and transfer), measured in percentage. *T_GDP* is the total tax revenues over GDP. *gini_mi* is the Gini of market income. *aagr* is the annual average growth rate from 2006 to 2019.

	Dependent variable:			
	Dgini			
	(1)	(2)	(3)	(4)
T_GDP	−0.653***	−0.471***	−0.586***	−0.465***
	(0.027)	(0.105)	(0.046)	(0.105)
gini_mi		−0.175*		−0.132
		(0.098)		(0.103)
aagr			5.472*	4.085
			(3.111)	(3.274)
Observations	48	48	48	48
R^2	0.928	0.933	0.933	0.935
Adjusted R^2	0.926	0.930	0.930	0.931
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A7: Regression results on Gini reduction

3 Clustering

3.1 K-means method

Description of the method

We use the *K-means* of machine learning which is one of the most widely used method to perform automatic clustering. More precisely, the *K-means* algorithm is an unsupervised non-hierarchical clustering algorithm. It allows to group the data (here the countries) in a finite number of sets on the basis of the provided characteristics (here the variables describing the socio-fiscal systems). From a given number of clusters, the algorithm provides the groups of countries that minimize the distance within clusters (see Kassambara (2017) for a presentation of the algorithm and its implementation).

As it is explained in introduction, K-means and PCA are closely linked. Therefore, in Figure A2a below, the clusters are represented over the PCA dimensions.

3.2 Choosing the number of clusters

The difficulty of the K-means method lies in choosing the optimal number of clusters. For each exercise, we determine the number of clusters endogenously by using three different methods. First, the Elbow method indicates the number of clusters that minimizes the total within-cluster sum of square. This measures the compactness of the clustering. The gap-statistics tests several values of k and determines the optimal number in order to maximize the statistical gap with respect to a random dispersion. The average silhouette approach determines how well each object lies within its cluster. The optimal number of clusters k is the one that maximizes the average silhouette over a range of possible values for k. For 2006 and 2019, the methods do not point out the same optimal number of clusters. We selected the number that was selected by two of them. The results are displayed in Figures A1a and A1b.

3.3 Cluster description

Here are presented the variables that characterized the best the clusters. Table A8 describes global characterization of the overall sample by displaying the most discriminating variables. Then, in Tables A9, A10 and A11, each variable is tested by cluster to determine whether the cluster average is significantly different from the overall average. Only variables with a test value (in absolute value) greater than 1.96 are considered important to describe the class (and therefore displayed). The greater the test value in absolute value, the more the variable characterizes the class.

3.4 Cluster construction

To construct the clusters in Figure 6, we consider ten variables for the years 2006 and 2019. These are transfers (Tr) and nine fiscal variables: tax revenue over GDP (T_{it}) , the six rates as a proportion of total taxation that represents the tax structure ($T_{L\_hou}, T_{L\_corp}, T_{K\_inc\_hou}, T_{K\_inc\_corp}, T_{K\_sto}, T_C$ as a percentage of total tax revenues) and the implicit rates on labour and capital income (ITR_L and ITR_{K_inc}). We test their stability here. The year clusters remain unchanged if we consider only the nine tax variables, the ten initial variables and the measure of Gini change, the ten initial variables and the Esping-Andersen decommodification index. The latter two clusterings are shown in Figures A2 and A3. Figure A2a represents the clusters under the PCA dimensions and Figure A3b displays the dimensions composition. In Figure A5, we add the annual average growth rate among the initial variables.

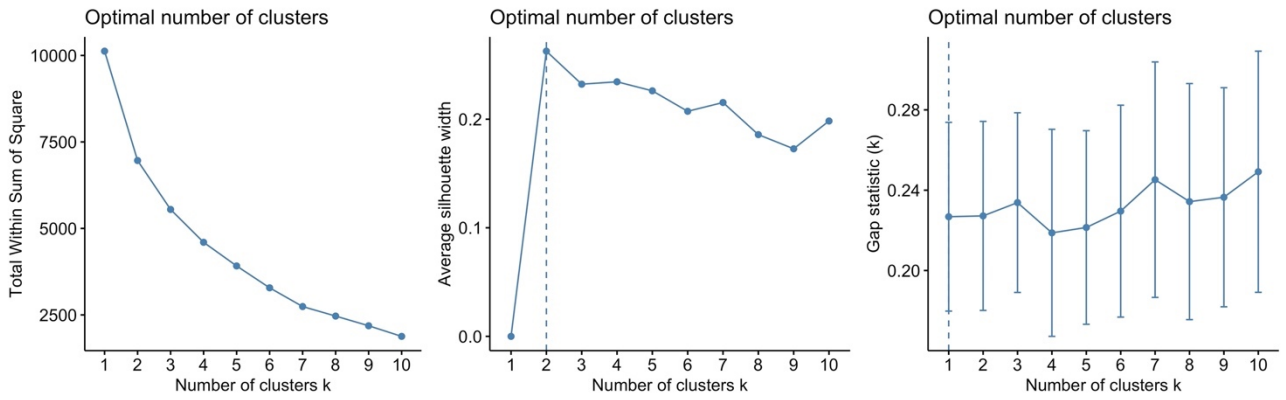

Figure A1a: Choosing the number of clusters, with 2006 data from Figure 6

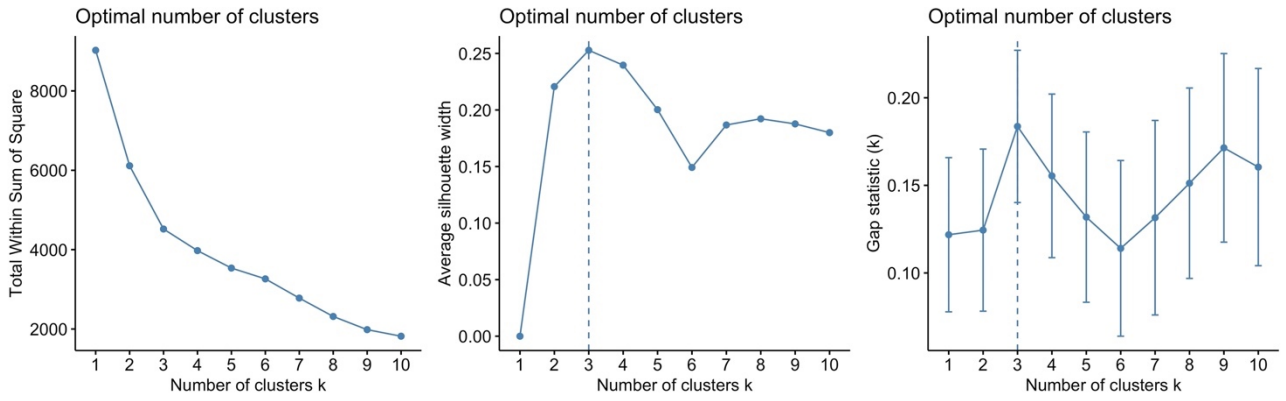

Figure A1b: Choosing the number of clusters, with 2019 data from Figure 6

	Eta2	P-value
T	0.82	0.00
ITR L	0.81	0.00
K sto t	0.51	0.00
C t	0.48	0.00
Tr	0.42	0.00
K inc corp SE t	0.29	0.01
L hou t	0.27	0.02
K inc hou t	0.27	0.02
L corp SE t	0.26	0.02

Table A8: Main decisive variables – overall sample

HR	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
ITR L	4.40	45.42	36.60	3.75	8.29	0.00
T	4.18	41.84	35.78	1.89	5.99	0.00
Tr	2.82	17.67	14.92	4.28	4.02	0.00

Table A9: Main decisive variables – High-redistribution cluster

INT	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
C t	3.60	36.74	30.18	5.03	6.99	0.00
K sto t	-2.20	4.65	6.77	2.82	3.70	0.03
K inc hou t	-2.56	1.36	2.21	0.88	1.28	0.01
L hou t	-2.72	26.76	31.25	4.19	6.32	0.01

Table A10: Main decisive variables – Intermediate cluster

LIB	v.test	Mean in category	Overall mean	sd in category	Overall sd	p.value
K sto t	3.74	11.01	6.77	2.12	3.70	0.00
K inc corp SE t	2.74	12.03	8.90	2.74	3.73	0.01
K inc hou t	1.98	2.99	2.21	1.15	1.28	0.05
C t	-2.28	25.30	30.18	6.99	6.99	0.02
L corp SE t	-2.71	14.49	20.68	5.43	7.46	0.01
Tr	-3.09	11.12	14.92	2.36	4.02	0.00
ITR L	-3.71	27.17	36.60	3.91	8.29	0.00
T	-4.06	28.32	35.78	3.65	5.99	0.00

Table A11: Main decisive variables – Liberal cluster

PCA representation

We perform here the principal component analysis (PCA), to compare with the K-means method. PCA and K-means are closely linked (see Ding and He (2004) for more details) but they have different objectives. Intuitively, PCA seeks to represent the data by reducing its dimensions, while K-means groups the observations around the cluster centroids, based on their similarity.

Figure A2a and A2b show that the PCA and K-means deliver quite similar results. Figure A10b illustrate that the main variables that contribute to dimensions 1 and 2 are the tax revenues over GDP and the implicit tax rate of labour, similar to the results of Table A6 pointing out the main

discriminative variables between clusters. Then, on Figure A2a, the 2019 representation of the clustering show interesting features. First, the PCA representation seems to be consistent with the K-means clustering distinction. The liberal cluster demonstrates a strong coherence. This is also the case for the “intermediate” and high-redistributive groups.

With redistributive effects

In Table A5, we add to the original ten variables the relative change in the Gini index. The methodology is described in Section 4 and the results are presented in Table 4. However, the years studied in LIS vary by country, which biases their comparison. Therefore, we chose to keep 2019 as the initial ten variables, the purpose being to test the stability of the clustering in 2019 and to add the change in the Gini from the last year. Among the 17 countries available in LIS, we find that the optimal number of clusters is two.

With Esping-Andersen decommodification index

In Figure A4, we add to the ten initial variables the Esping-Andersen (1990) index of decommodification. This index allows us to measure demarketing in the states according to several qualitative and quantitative criteria, in order to distinguish between different types of welfare state. As explained above, we keep the year 201 for the initial variables. Among the 17 countries where the index is measured, none is in the intermediate cluster, the clustering however finds the liberal and high-redistribution model.

With growth rate

In Figure A5, the average annual growth rate is added to the ten initial variables.

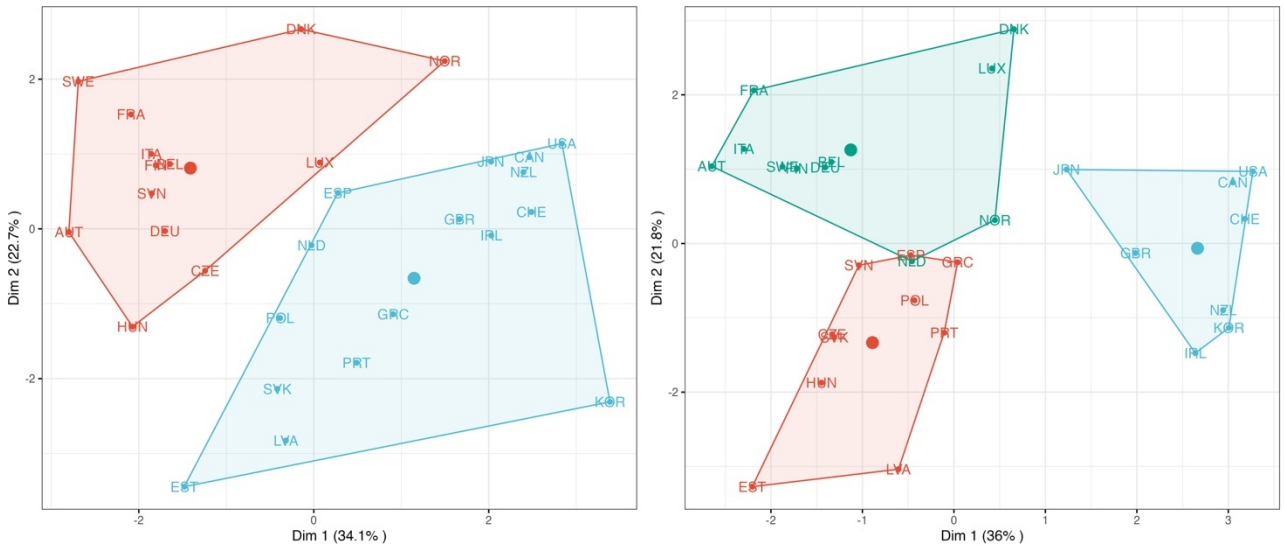

Figure A2a: Clustering PCA visualization

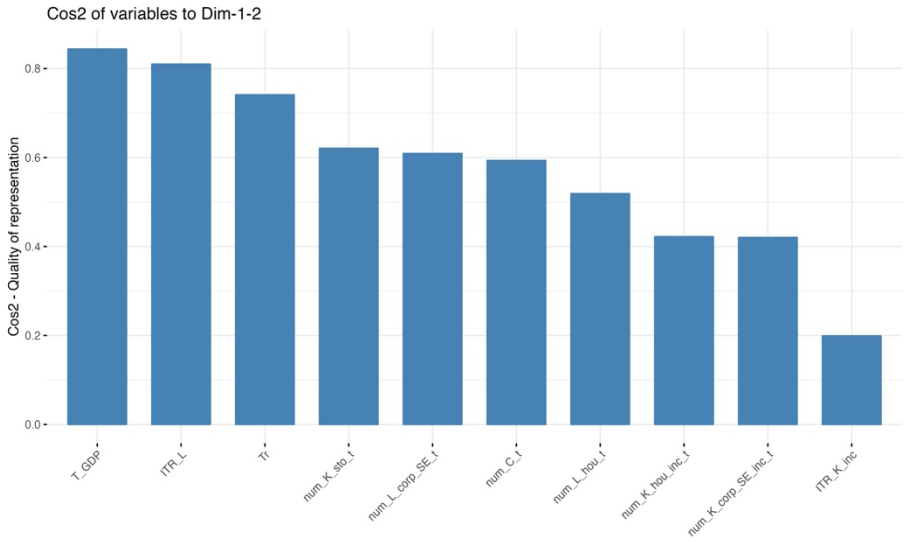

Figure A2b: Variables contributions to the two first dimensions

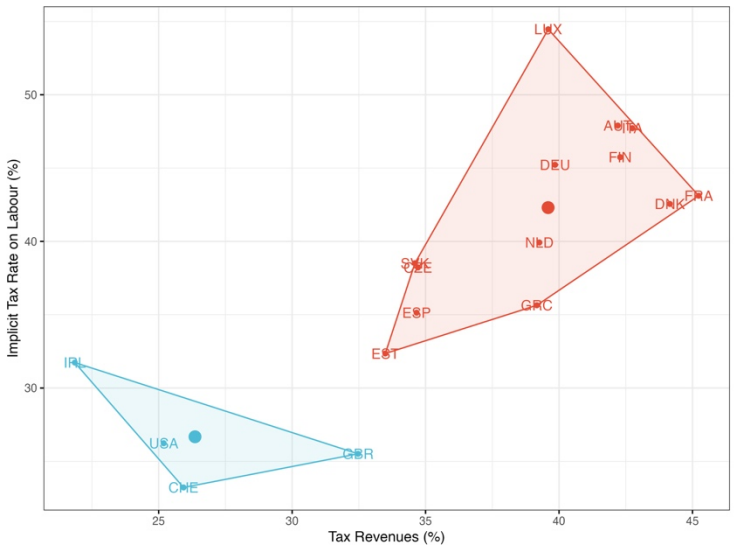

Figure A3: Clustering with redistributive effects

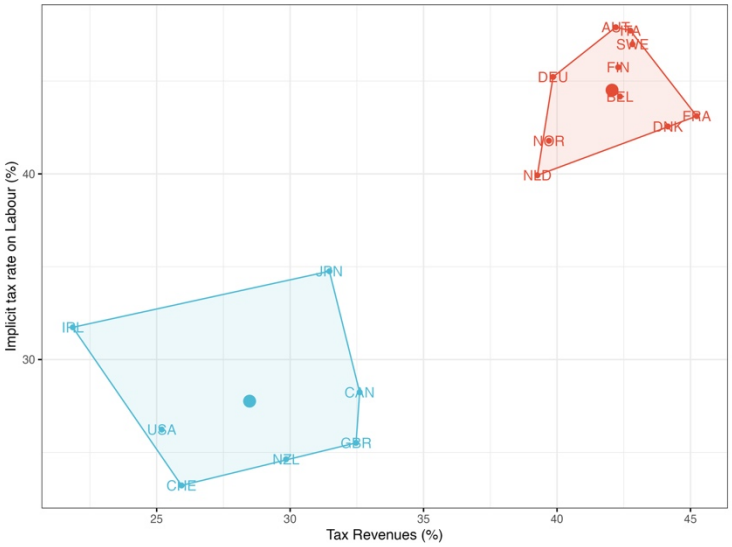

Figure A4: Clustering with Esping-Andersen's (1990) index of decommodification

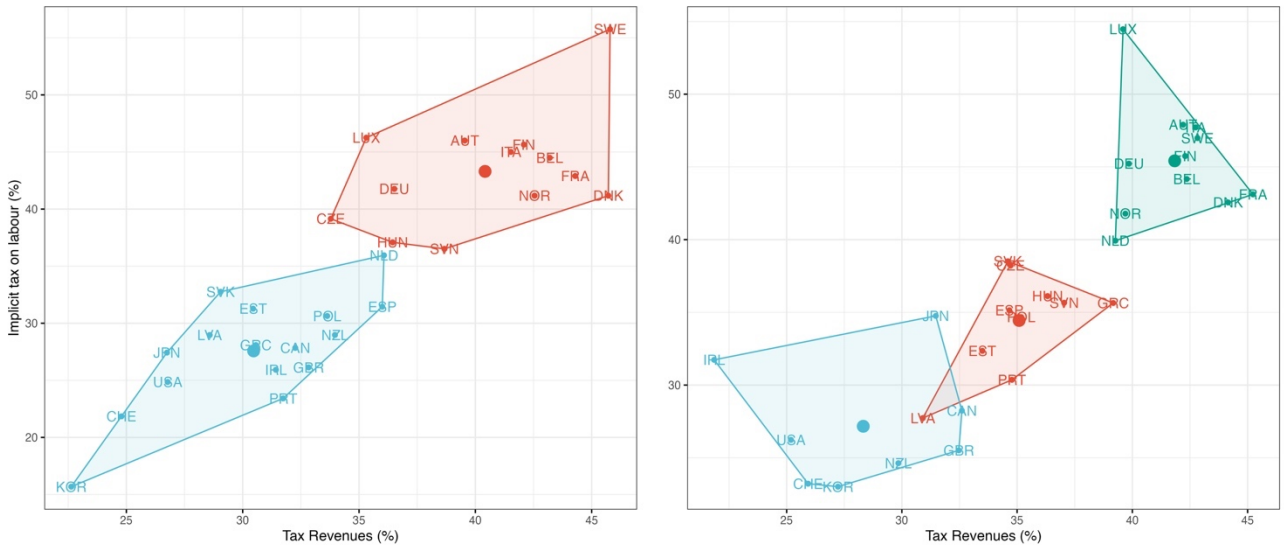

Figure A5: Clustering with annual average growth rate

4 OECD classification of taxes

1. 1000. Taxes on income, profits and capital gains

1100. Taxes on income, profits and capital gains of individuals

1110. On income and profits

1120. On capital gains

1200. Corporate taxes on income, profits and capital gains

1210. On income and profits

1220. On capital gains

1300. Unallocable as between 1100 and 1200

2. 2000. Social security contributions

2100. Employees

2110. On a payroll basis

2120. On an income tax basis

2200. Employers

2210. On a payroll basis

2220. On an income tax basis

2300. Self-employed or non-employed

2310. On a payroll basis

2320. On an income tax basis

2400. Unallocable as between 2100, 2200 and 2300

2410. On a payroll basis

2420. On an income tax basis

3. 3000. Taxes on payroll and workforce

4. 4000. Taxes on property

4100. Recurrent taxes on immovable property

4110. Households

4120. Other

4200. Recurrent taxes on net wealth

4210. Individual

4220. Corporate

4300. Estate, inheritance and gift taxes

4310. Estate and inheritance taxes

4320. Gift taxes

4400. Taxes on financial and capital transactions

5. 5000. Taxes on goods and services

5100. Taxes on production, sale, transfer, leasing and delivery of goods and rendering of services

5110. General taxes

5111. Value added taxes

5112. Sales taxes

5113. Turnover and other general taxes on goods and services

5120. Taxes on specific goods and services

5121. Excises

5122. Profits of fiscal monopolies

5123. Customs and import duties

5124. Taxes on exports

5125. Taxes on investment goods

5126. Taxes on specific services

5127. Other taxes on international trade and transactions

5128. Other taxes on specific goods and services

5130. Unallocable as between 5110 and 5120

5200. Taxes on use of goods, or on permission to use goods or perform activities

5210. Recurrent taxes

5211. Paid by households in respect of motor vehicles

5212. Paid by others in respect of motor vehicles

5213. Other recurrent taxes

5220. Non-recurrent taxes

5300. Unallocable as between 5100 and 5200

6. 6000. Other taxes

6100. Paid solely by business

6200. Paid by other than business or unidentifiable

Countries	Year	T	C	L hou	L corp SE	K inc hou	K corp SE	K sto
AUT	2019	42.20	11.35	14.07	12.63	0.52	2.72	0.91
	2006	39.54	11.21	13.30	11.41	0.57	2.10	0.96
BEL	2019	42.36	10.63	14.13	10.20	0.13	3.70	3.56
	2006	43.20	10.92	14.76	10.29	0.32	3.49	3.42
CAN	2019	32.60	7.53	11.84	3.96	1.00	4.29	3.98
	2006	32.26	7.86	11.14	4.20	1.12	4.34	3.60
CHE	2019	25.93	4.86	10.13	3.94	0.71	3.41	2.88
	2006	24.79	5.24	9.23	3.72	0.78	3.02	2.80
CZE	2019	34.72	10.71	8.70	10.64	0.09	3.79	0.79
	2006	33.79	9.85	8.06	9.94	0.08	4.96	0.90
DEU	2019	39.85	10.00	16.78	9.25	0.41	2.00	1.41
	2006	36.52	9.90	14.55	8.28	0.37	2.18	1.25
DNK	2019	44.15	13.86	20.43	2.32	2.06	3.16	2.34
	2006	45.72	15.85	19.06	2.79	2.30	3.70	2.01
ESP	2019	34.66	9.63	8.46	10.11	1.01	2.60	2.86
	2006	36.00	9.64	7.34	9.54	1.07	4.83	3.58
EST	2019	33.49	13.92	5.78	11.28	0.19	1.83	0.49
	2006	30.44	12.83	5.32	9.82	0.24	1.58	0.66
FIN	2019	42.29	14.16	13.62	9.11	1.32	2.53	1.55
	2006	42.09	13.22	12.68	10.42	1.38	3.25	1.15
FRA	2019	45.23	12.12	10.73	14.03	1.50	2.51	4.36
	2006	44.29	10.69	9.66	14.63	0.89	4.31	4.11
GBR	2019	32.47	10.49	8.87	4.87	1.75	2.29	4.21
	2006	32.83	9.66	9.47	4.73	1.35	3.41	4.21
GRC	2019	39.17	14.42	10.87	6.07	0.85	2.56	4.39
	2006	30.57	10.64	8.41	5.22	0.70	2.84	2.76
HUN	2019	36.32	15.82	9.99	7.08	0.46	1.62	1.36
	2006	36.44	13.79	8.70	9.66	0.61	2.62	1.07
IRL	2019	21.86	6.47	6.91	3.28	0.62	3.05	1.53
	2006	31.43	11.16	7.56	4.07	2.12	3.62	2.90
ITA	2019	42.75	11.85	11.21	12.74	1.12	2.65	3.17
	2006	41.53	10.12	9.47	12.74	1.13	4.31	3.75
JPN	2019	31.47	5.76	10.65	7.70	0.50	3.82	3.04
	2006	26.74	4.43	7.95	6.31	0.48	4.57	3.00
KOR	2019	27.23	7.32	7.12	4.60	0.40	4.28	3.51
	2006	22.63	7.90	4.68	3.28	0.33	3.24	3.20
LUX	2019	39.60	9.11	13.79	5.14	1.56	6.06	3.94
	2006	35.31	9.71	10.13	5.50	1.69	4.93	3.35
LVA	2019	30.91	13.52	8.32	6.93	0.70	0.16	1.29
	2006	28.58	11.75	7.59	6.01	0.00	2.11	1.11

Table A1a: Taxation structure – by country

Countries	Year	T	C	L hou	L corp SE	K inc hou	K corp SE	K sto
NLD	2019	39.26	11.62	12.07	9.14	0.71	3.69	2.02
	2006	36.07	11.00	12.00	6.77	0.62	3.46	2.23
NOR	2019	39.70	11.36	12.21	7.75	0.86	5.95	1.57
	2006	42.54	11.30	9.83	6.62	0.85	12.54	1.42
NZL	2019	29.85	11.11	10.18	1.01	1.03	3.83	2.69
	2006	33.96	11.00	11.66	1.58	1.41	5.68	2.63
POL	2019	35.18	12.71	10.01	6.65	0.45	3.90	1.46
	2006	33.63	12.75	8.58	6.43	0.44	3.79	1.63
PRT	2019	34.77	13.46	9.19	6.41	0.56	3.27	1.88
	2006	31.74	13.72	8.12	4.80	0.64	3.01	1.44
SVK	2019	34.60	11.49	9.99	9.09	0.00	3.08	0.96
	2006	29.05	10.75	7.83	6.40	0.00	3.04	1.03
SVN	2019	37.04	12.99	12.97	7.53	0.44	1.96	1.15
	2006	38.67	12.84	11.76	9.59	0.55	2.92	1.00
SWE	2019	42.83	11.91	13.52	12.16	1.06	3.03	1.15
	2006	45.79	12.02	15.52	12.05	1.20	3.52	1.48
USA	2019	25.19	4.09	11.34	3.48	0.86	2.06	3.35
	2006	26.78	4.10	10.51	3.85	0.94	3.82	3.55

Table A1b: Taxation structure – by country

Endnotes

[1] A traditional example is the difference between the taxation of dividends received by households and the taxation of corporate profits. Both taxes are on the capital factor. If firms reinvest all their profits, then the first tax does not affect investment, while the second reduces it.
[2] In the rest of the paper, corporate and self-employed capital income are directly referred as corporate capital income.
[3] The historical timetables for all countries are available upon request.
[4] Ireland's tax rate has fallen by almost nine points, from 31.4% in 2006 to 24.7% in 2019. However, this drop does notcorrespond to a drastic fall in tax revenues, but to an accounting change in GDP: between 2014 and 2015, GDP grew by25.2% and tax rate decreased by six points.
[5] The drop in the corporate and self-employed capital income yields during the financial crisis can be partly explained by a decrease in the tax base. However, we find that the ITR on corporate capital income and self-employed decreases by 1.3 percentage points between 2006 and 2019, which confirm the downward convergence. The tax rate and tax base effects are further detailed in Appendix.
[6] The spike of the taxes on capital stocks in 2017 for the US comes from a special non recurrent (Tax Cuts and Jobs Act repatriation tax). It is a one-time tax on past profits of US corporations' foreign subsidiaries and should not be considered as a permanent change.
[7] Other measures of inequality can be used. The Gini coefficient on income is the most widely used indicator, even though it is more sensitive to changes in inequality around the median, and much less to inequalities at the top of the income distribution.
[8] We are very thankful for Elvire Guillaud for sharing her data.
[9] Initial work uses principal component analysis (Amable (2005)). The use of machine learning methods proves to be more powerful to identify the coherences and then discuss the variables that determine these coherences, as the number of clusters can be made endogenous.
[10] The similiarity of the results comes from the correlation between the Gini reduction and the tax revenues over GDP discussed in Section 3, so the redistributive effects are largely captured by the macro variables.
[11] In Appendix, we also provide results where we use the Esping-Andersen (1990) index of decommodification.
[12] The measurement of the evolution of the standard deviation is of course biased since the three groups were formed in the year 2019, so it is normal to find rather a decrease in the standard deviation. Therefore, we use it mostly in comparison between the groups, or when it increases. More consistent convergence tests are performed in Appendix.
[13] see sufra note