

LIS

Working Paper Series

No. 743

Income Inequality, Redistributive Preferences and the Extent of Redistribution: An Empirical Application of Optimal Tax Approach

Hannu Tanninen, Matti Tuomala, Elina Tuominen

August 2018
(revised in October 2019)



CROSS-NATIONAL
DATA CENTER
in Luxembourg

Luxembourg Income Study (LIS), asbl

Income inequality, redistributive preferences
and the extent of redistribution:
An empirical application of optimal tax approach[☆]

Hannu Tanninen^a, Matti Tuomala^b, Elina Tuominen^b

^a*University of Eastern Finland, Finland*

^b*Tampere University, Finland*

Revised version: October 2019

(Previous version: July 2018)

Abstract

We examine empirically the relationship between the extent of redistribution and the components of the Mirrlees framework, with a focus on inherent inequality and government's redistributive preferences. We have constructed our income distribution variables from the Luxembourg Income Study (LIS) database, which provides information on both factor and disposable incomes. Our redistributive preference measure is constructed using the optimal tax formula for which we have collected data from various sources. In addition to traditional linear specifications, we use flexible methods to allow nonlinearities because pre-specified functional forms are not easy to justify in empirical investigations of the optimal tax framework. We study 14 advanced countries for approximately four decades and find support for the Mirrlees model: There is a positive relationship between factor-income inequality and the extent of redistribution. We also find a link between our redistributive preference measure and the extent of redistribution.

Keywords: income inequality, nonlinearity, preferences, redistribution

JEL classification: C14, D31, H3

[☆]The authors wish to thank Spencer Bastani, Vidar Christiansen, Thomas Gaube, Emanuel Hansen, Emmanuel Saez, Håkan Selin, Danny Yagan and participants at the UC Berkeley Public Finance Lunch Seminar and the Seventh ECINEQ Meeting for useful comments on an earlier version of this study. Funding from the Strategic Research Council (SRC) at the Academy of Finland (project no. 293120, 'Work, Inequality and Public Policy') is gratefully acknowledged.

Email addresses: hannu.tanninen@uef.fi (Hannu Tanninen), matti.tuomala@tuni.fi (corresponding author) (Matti Tuomala), elina.tuominen@tuni.fi (Elina Tuominen)

1. Introduction

The post-war history of income inequality in advanced countries can be divided, at least roughly, into two phases. From 1945 to about the mid-1980s, pre-tax inequality, or the inequality of factor incomes (incomes from earnings and capital), decreased at least in part because of a reduction in skilled/unskilled wage differentials and asset inequality. The second phase occurred from the 1980s onward, when inequality reversed course and increased. For the 1970s and the above-described latter period we have evidence from the Luxembourg Income Study (LIS) database. This database provides data on both factor and disposable incomes for a number of advanced countries over the past four decades, which facilitates the study of the extent of redistribution. The difference between the factor-income Gini and disposable-income Gini is an often-used measure of the overall redistributive effect of taxes and transfers. These types of difference measures are also used in the current study to discuss the evolution of the extent of redistribution.

For example, Immervoll and Richardson (2011) studied OECD countries using the LIS data. They reported that governmental redistribution has become less effective in compensating increasing inequalities since the 1990s. Moreover, top income shares have increased in many advanced economies over the past three decades (Atkinson & Piketty, 2010), and top tax rates on upper-income earners have declined significantly in many OECD countries during this period (Piketty et al., 2014). Economists have formulated several hypotheses about the causes of increasing inequality, but the explanations are not fully compelling. For example, Atkinson et al. (2011) emphasised that it is very difficult to account for these figures using the standard labour supply/demand explanation. Hence, the role of social policies and progressive taxation should not be dismissed in these discussions.

Figure 1 illustrates changes that have taken place in factor-income inequality and redistribution during a period of 20 years, from the mid-1980s to the mid-2000s. The figure implies a positive association, but there are some outliers. For example, in France inequality decreased during this period, but there was more redistribution. Thus, it seems that factor-income inequality does not fully explain the extent of redistribution in the plotted countries.

There is now a considerable body of empirical literature seeking to explain the observed patterns of redistribution. Here we outline some of this literature. The starting

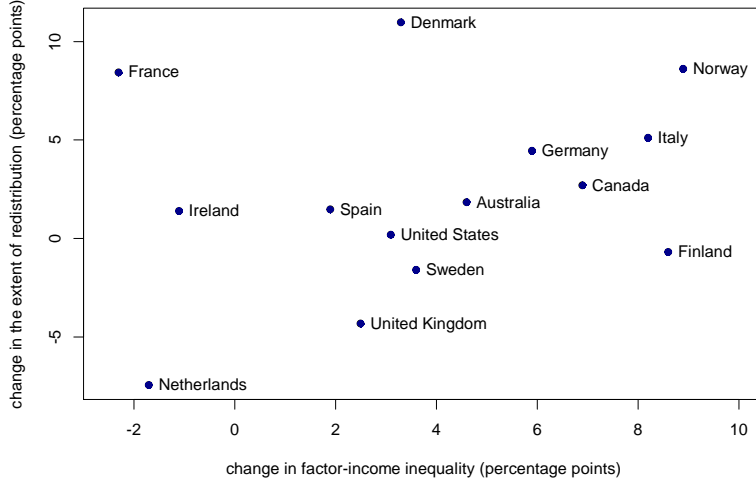


Figure 1: Illustration of evolution from the mid-1980s to the mid-2000s: we plot the change in factor-income Gini against the change in the extent of redistribution over this period; change is defined as the difference between two observations. The Gini coefficients are expressed as percentages, and the extent of redistribution is defined as $RD_{relative} = 100(Gini_{factor} - Gini_{disposable})/Gini_{factor}$. The corresponding factor-income Ginis of the 14 countries are provided in Table 1. Data source is the LIS database, and more information can be found in Appendix A.

point for seeking the determinants of the extent of redistribution, both across countries and over time, is most commonly some model of the political process, originating with Romer (1975), Roberts (1977) and Meltzer and Richard (1981). Persson and Tabellini (2002) provide a survey. A key element in this literature is the political mechanism – the median voter theory – through which greater inherent inequality leads to greater redistribution. The often-cited model of Meltzer and Richard (1981) shows that the larger the gap between mean and median income (that is, inequality), the larger the scale is of income redistribution favoured by the median voter. However, some authors have suggested that redistribution is greater the less inherent inequality there is (e.g. Peltzman, 1980; Persson, 1995; Lindert, 2000). Peltzman (1980) attempts an explanation that greater inequality links to the lower classes’ ability to demand redistribution towards themselves.

Empirical studies have provided mixed evidence on the association between inequality and demand for redistribution (e.g. Perotti, 1996; Moene & Wallerstein, 2001; Finseraas, 2009). For example, Alesina et al. (2001) provided a distinct example of the ‘paradox of redistribution’: a simple comparison of the United States and Europe showed that Europe had lower pre-tax inequality and more redistribution. They pointed out that the extent of altruism – which may be different in different societies – may show in the demand for redistribution. In addition, Georgiadis and Manning (2012) showed that implications resembling altruism may arise when individuals are uncertain about their future incomes –

thus, individuals may be concerned about tax rates at different points in the distribution. Earlier, Benabou and Ok (2001) discussed the ‘prospect of upward mobility’ hypothesis – individuals may consider outcomes in those parts of the income distribution where they expect to end up in the future.

Numerous recent empirical studies on inequality and redistribution have utilised the LIS database. For example, both Milanovic (2000) and Scervini (2012) confirmed the positive association between inequality and redistribution, but Milanovic (2000) found less support for the median voter hypothesis in explaining redistribution decisions. Milanovic (2010) discusses the critique that his earlier study has received and, in particular, emphasises the median voter hypothesis as only one possible mechanism between initial inequality and redistribution. Recently, Luebker (2014) emphasised behavioural aspects in understanding the extent of redistribution. He did not find support for the Meltzer–Richard hypothesis, but he provided some evidence that actual preferences are associated with redistributive outcomes.

The approach in this study is novel and different from those above because our starting point is the optimal tax framework developed by Mirrlees (1971). This model has dominated the economics of redistributive taxation for the past 40 years, and three elements of the model are useful for our purpose. First is the concept of inherent inequality, reflecting, among other things, skilled/unskilled wage differentials, asset inequality and social norms. Unlike the original Mirrlees model on labour income, we adopt the view that the taxation is based on the comprehensive income (i.e. taxing the sum of labour and capital income). In fact, in many countries most ordinary capital income, such as interest from a standard savings account, is taxed jointly with labour income. Moreover, in the case of the Nordic dual income tax model it can be difficult to distinguish between labour and capital income in practice. It also turns out, as noted by Saez and Stantcheva (2016), that in the case of the comprehensive income taxation the optimal tax formula takes the same form as in Mirrlees (1971) and Saez (2001).¹ If there is no intervention by the government, the inherent inequality will be fully reflected in the disposable income. However, if the government wants to intervene – as seems to be the case in developed

¹Let H be the cumulative distribution of the total income y and $h(y)$ the associated density (assuming a linearized tax system at point y). Then using the Saez (2001) procedure optimal nonlinear income tax on total or comprehensive income satisfies $\frac{t_y}{1-t_y} = \frac{(1-\varphi(y))}{\varepsilon_y} \frac{(1-H(y))}{yh(y)}$ where $\varphi(y)$ is the average welfare weight on individuals with total income higher than y' and ε_y is the weighted average elasticity.

countries – it will find the second component of the Mirrlees model, the egalitarian objectives of the government. In addition, if the government tries to redistribute income from high-income people to low-income people, there will be incentive and disincentive effects. In other words, redistribution policy is a product of circumstances and objectives.²

Kanbur and Tuomala (1994) showed that the optimal income tax/transfer system becomes more progressive when inherent inequality increases, taxing the better off at higher rates to support the less well off. Thus, one of the policy responses in view of inherent inequality should be a greater willingness to redistribute through the tax and transfer system. Correspondingly, if inherent inequality decreases, governmental redistribution decreases.

Interestingly, at the same time as the large growth in top income shares over the past few decades, many advanced countries have shifted the tax burden from the top to further down in the distribution. Numerical results in Tuomala (2016) suggest that this shift in tax burden cannot be justified by the standard Mirrlees model, which embodies conventional assumptions about inequality aversion and the trade-off between equity and efficiency. Both in utilitarian and maximin cases, an appropriate response to rising inequality is a shift towards a more progressive income tax system. For more details, see Tuomala (2016).

Previously, Tanninen and Tuomala (2005) discussed how some of the basic features of redistribution can be explained through the Mirrlees model. They examined the relationship between inherent inequality and the extent of redistribution by utilising the LIS data for a number of OECD countries over two to three decades. They found that redistribution is positively associated with inherent inequality. However, their empirical results were based on the assumption that the degree of espoused egalitarianism has remained constant over the period considered. There is now some recent individual-country-level evidence that there could have been a shift in norms, causing governments to become less willing to finance transfers and to levy progressive taxes, leading to reductions in the extent of redistribution. One could argue, in line with Atkinson (1999), that these kinds of changes have been episodic rather than time-trend and are therefore rather difficult to

²There is another strand of optimal redistribution literature (see Mirrlees, 1974; Varian, 1980; Tuomala, 1984, 1990) stressing the social insurance role of redistributive taxation. In this framework, an increase in variability of income would also increase the optimal degree of progressivity because it increases the insurance value of progressive taxation.

justify, for example, in the context of median voter models. Thus, we focus here also on the role of the egalitarian objectives of government, which is an important component of the Mirrlees model.

To the best of the authors' knowledge, the current empirical study is the first to explore inequality, redistributive preferences and the extent of redistribution in the Mirrlees framework. The utilised LIS database provides data on both factor and disposable incomes for a number of advanced countries over the last four decades, and the current study focuses on 14 advanced economies. In addition to using Gini coefficients, we also utilise percentile ratios in measuring inequality and the extent of redistribution. Moreover, we have collected data from various sources to construct a measure of government's taste for redistribution. We adopt the inverse-optimum approach and more specifically we utilise the top tax rates to reveal the shape of implicit welfare weights.

The structure of the paper is as follows. Section 2 depicts the data, empirical specification and methods. The current paper employs, in addition to traditional linear models, flexible methods to address the issue of chosen functional forms. Namely, the shapes of the relationships are not known beforehand. Section 3 provides our empirical results, including sensitivity checks. Finally, Section 4 concludes.

2. Empirical approach

2.1. Specification and data

The relationship that we test can be expressed as follows:

$$RD = h(I_f, \gamma, \epsilon; \mathbf{x}),$$

where RD is the extent of redistribution measured in terms of the difference between factor-income inequality (I_f) and disposable-income inequality (I_d). Our main results are presented for a relative measure; that is, $RD_{I;relative} = 100(I_f - I_d)/I_f$. In the sensitivity analysis, we also discuss the alternative case where the extent of redistribution is measured in absolute terms. Function h includes three components, I_f , γ and ϵ , that reflect the ingredients of the Mirrlees model, and they are inherent inequality, government's redistributive preferences and elasticity of taxable income, respectively. In addition, \mathbf{x} denotes control variables.

We begin by describing our inequality data. Most of the median voter studies have utilised data sets including the largest possible number of countries all around the world (e.g. the panel data set of Deininger and Squire, 1996). However, such data sets have many problematic features that have been discussed in detail by Atkinson and Brandolini (2001). As Milanovic (2000) pointed out, there is no proper distinction between inherent income and redistribution (for different definitions of income see Atkinson et al., 1995). Fortunately, this distinction can be taken into account in the LIS data.

We have calculated inequality measures for two income concepts – factor and disposable household income – from the LIS database. We refer to the difference between these two measures as redistribution. LIS has harmonised microdata from (mostly) high- and middle-income countries, and the data are organised into different waves according to the date of the data. In addition to the LIS historical data (Wave 0), we use the data from Wave I around 1980 to Wave IX around 2013. The lengths of different waves are not uniform. Moreover, some countries may have more than one observation within the same wave.³ We use all available LIS data for which data on our other variables are available, and the resulting data set is not balanced.

Table 1: Levels of factor-income inequality in 14 countries from the mid-1980s to the mid-2000s.

	LIS Wave II around 1985			LIS Wave IV around 1995			LIS Wave VI around 2004		
	$Gini_f$	$P90/P50_f$	(year)	$Gini_f$	$P90/P50_f$	(year)	$Gini_f$	$P90/P50_f$	(year)
Australia	43.7	2.16	(1985)*	47.9	2.34	(1995)	48.3	2.42	(2003)
Canada	40.8	2.16	(1987)	44.9	2.29	(1994)	47.7	2.42	(2004)
Denmark	41.7	1.86	(1987)*	44.7	2.02	(1995)*	45.0	2.00	(2004)
Finland	38.9	1.89	(1987)*	48.1	2.36	(1995)	47.5	2.28	(2004)
France	50.7	2.46	(1984)*	49.2	2.52	(1994)	48.4	2.43	(2005)
Germany	44.3	2.09	(1984)	46.2	2.25	(1994)	50.2	2.46	(2004)
Ireland	51.2	2.75	(1987)	49.5	2.39	(1995)	50.1	2.53	(2004)
Italy	42.6	2.18	(1986)	47.9	2.47	(1995)	50.8	2.56	(2004)
Netherlands	48.0	2.24	(1987)*	46.7	2.14	(1993)*	46.3	2.13	(2004)
Norway	36.6	1.86	(1986)*	42.6	1.93	(1995)	45.5	2.12	(2004)
Spain	43.5	2.25	(1985)	51.0	2.71	(1995)	45.4	2.39	(2004)
Sweden	43.4	2.01	(1987)*	49.8	2.39	(1995)	47.0	2.19	(2005)
United Kingdom	50.8	2.57	(1986)	54.3	2.78	(1995)	53.3	2.76	(2004)
United States	46.2	2.47	(1986)	49.1	2.66	(1994)	49.3	2.74	(2004)

Data source: Luxembourg Income Study (LIS). More information can be found in Appendix A.

* This year's observation cannot be used in empirical models due to missing information in other variables.

We focus on the 14 advanced countries that are listed in Table 1. In addition to studying the traditional Gini coefficients, we investigate the development of the percentile ratio

³For more information about the LIS waves, visit: <http://www.lisdatacenter.org/our-data/lis-database/documentation/list-of-datasets/>.

P90/P50. Both measures are much used in inequality studies and provide somewhat complementary information. The Gini index is an overall measure of inequality, reflecting the behaviour of the whole income distribution, and being particularly sensitive to asymmetries in the central part of the distribution. Percentile ratios focus on two specific sections of the income distribution, providing an idea of how close (or distant) they are from each other. The P90/P50 ratio gives the 90th percentile relative to the median, focusing more on disparities at the top half of the distribution. The table illustrates that over the sample period the inequality of factor incomes has risen in most countries.

Figure 2 shows the development of the extent of redistribution ($RD_{Gini;relative}$) in the countries under investigation. The countries are categorised into three groups to provide a concise but readable illustration. The categorisation is the following: Anglo-Saxon (Australia, Canada, Ireland, the UK and the USA; in main result models $N = 44$), Nordic (Denmark, Finland, Norway and Sweden; in main result models $N = 20$), and Continental European (France, Germany, the Netherlands, Italy and Spain; in main result models $N = 41$). The figure shows that the extent of redistribution has increased modestly in some countries. Thus, it appears that the redistributive role of government has corrected for some of the increase in inherent inequality. A corresponding figure of our alternative measure $RD_{P90/P50;relative}$ is in Appendix B (Figure B.6).

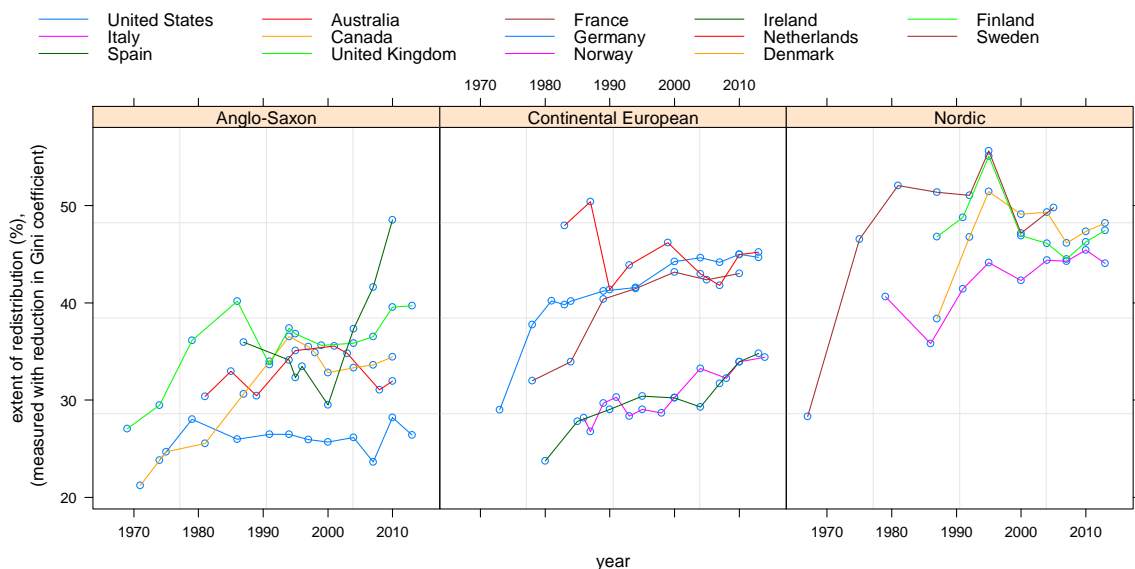


Figure 2: Evolution of the extent of redistribution when redistribution is measured in relative terms: $RD_{Gini;relative}$ (14 advanced countries, unbalanced data, years 1967–2013). Calculations based on LIS database. More information can be found in Appendix A.

How do we measure preferences for redistribution? There is a growing number of stud-

ies trying to reveal social or distributional preferences behind tax/transfer policy. Those studies start from the existing tax and transfers system and reverse-engineer it to obtain the underlying social preferences. Earlier contributions using this method are by Christiansen and Jansen (1978) and Ahmad and Stern (1984). Recently, detailed micro data on incomes and corresponding marginal tax rates have become available to study the social preferences implicit in tax–benefit systems. One of the first studies using micro data was by Bourguignon and Spadaro (2012). They consider the revealed social preferences of the French tax–benefit system. Spadaro et al. (2015) in turn consider the revealed social preferences of the tax–benefit systems for all European countries using European Union Statistics on Income and Living Conditions data. Lockwood and Weinzierl (2016) apply the inverse-optimum approach over time to the United States tax policy and show that the implicit weights decline slowly with income. Jacobs et al. (2017) use this method to find the redistributive preferences of political parties implicit in the reform proposals combining the reform proposals with micro data on the income distribution and the elasticity of the tax base in the Netherlands. Bastani and Lundberg (2017) use Swedish administrative panel data and provide a detailed characterisation of the income distribution, especially at the top. Bargain et al. (2014a) estimate labour supply elasticities from micro data and characterise the redistributive preferences embodied in the welfare systems of 17 EU countries and the United States under the assumption of optimality. They find that social welfare weights are always positive though not monotonically declining for low income groups. They further find that there are significant differences in social welfare weights over income between groups of countries (the United States vs. Continental/Nordic Europe vs. Southern Europe), with rather similar social welfare weights over income within groups of countries.⁴

In our empirical applications, we try to reveal government’s taste for redistribution (γ) using the optimal top tax formula $\tau = (1 - \gamma)/(1 - \gamma + \alpha\epsilon) \Leftrightarrow \gamma = 1 - \tau\alpha\epsilon/(1 - \tau)$, where τ is top income tax rate, α is Pareto-Lorenz coefficient and ϵ is elasticity of taxable income (Saez, 2001). Our data on top income tax rates and Pareto-Lorenz coefficients are collected from various sources, such as Piketty et al. (2011, 2014) and the World Inequality Database (2017); see Appendix A for more detailed information. Moreover, no

⁴For additional related literature see, for example, Kleven and Kreiner (2006), Bargain et al. (2014b) and Hendren (2017).

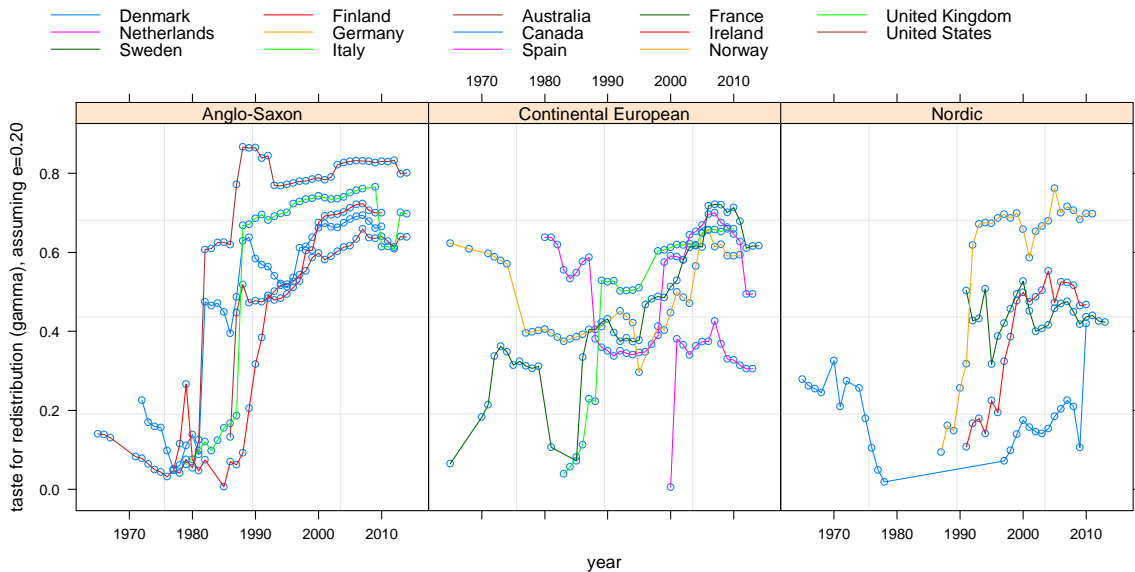


Figure 3: Evolution of redistributive preferences (14 advanced countries, unbalanced data). In calculating the γ (gamma) values, we have assumed constant elasticity of $\epsilon = 0.20$. Data are constructed from multiple sources, and more information is provided in Appendix A.

reliable panel data on behavioural responses ϵ exist, so we have investigated our results with some ‘reasonable’ values from the prior literature. In our preferred specifications we assume $\epsilon = 0.20$. For example, Saez et al. (2012) suggest that aggregate elasticities of taxable income are between 0.1 and 0.4. Moreover, in our empirical applications we consider only values $\gamma \geq 0$. This restriction to nonnegative values limits the number of observations in our empirical models: when $\epsilon = 0.20$, we get $N = 105$. Figure 3 describes the evolution of government’s taste for redistribution in the 14 countries of this study. Higher γ reflects lower redistributive preferences in society. According to Figure 3, government’s preferences to redistribute have decreased in many countries. In addition, we tried three alternative assumptions in calculating values for γ : these cases were $\epsilon = \{0.10, 0.15, 0.25\}$, and they are briefly discussed in the sensitivity checks.

Finally, our control variables are share of government employment, dependency rate, unemployment rate, trade union density and openness.⁵ Summary statistics and a complete list of data sources and definitions are provided in Appendix A.

⁵It has been argued that a larger government is ‘needed’ in more open economies. Rodrik (1998) gives an explanation that open economies are more subject to external shocks and that larger redistribution provides insurance and more stable income for individuals. In addition, the authors of the current study acknowledge that full assessment of the extent of redistribution should also take account of various publicly provided services at less than market value. These are considerable in Nordic countries. Many of these items – health care, education and social services – are very extensive.

2.2. Estimation method

We do not impose linearity into all our empirical models. In our preferred specifications, we allow all continuous covariates to enter flexibly so that potentially wrong functional forms would not bias our results. Our estimation approach is based on penalized cubic regression splines although we acknowledge that there are numerous alternative approaches to flexible modelling, such as kernel estimation.⁶ Moreover, due to small sample size, we assume an additive structure instead of a fully nonparametric one.⁷ The chosen method is accessible as there is a connection to traditional parametric models – traditional linear models are a special case. Moreover, there are ready-made statistical packages that can be utilised in the analysis. To estimate our additive models we use the established R software package ‘mgcv’, which has previously been utilised in economics studies on varying topics.⁸

Additive models provide a flexible framework for investigating the relation between inequality and redistribution. This study follows the approach presented in Wood (2006). The basic idea is that the model’s predictor is a sum of linear and smooth functions of covariates:

$$\mathbb{E}(Y_i) = \mathbf{X}_i^* \boldsymbol{\theta} + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}) + \dots$$

In the above presentation, Y_i is the response variable (extent of redistribution), \mathbf{X}_i^* is a row of the model matrix for any strictly parametric model components, $\boldsymbol{\theta}$ is the corresponding parameter vector and f_\bullet are smooth functions of the covariates, x_\bullet .

The flexibility of these models comes at the cost of two problems. First, one needs to represent the smooth functions f in some manner. One way to represent these functions is to use cubic regression splines, which is the approach adopted in this study. A cubic regression spline is a curve constructed from sections of cubic polynomials that are joined together so that the resulting curve is continuous up to the second derivative. The points

⁶Li and Racine (2006) describe nonparametric methods extensively, with the focus on kernels. Ahamada and Flachaire (2013) provide a concise overview of nonparametric methods.

⁷Additive models are a special case of generalized additive models (GAMs). GAMs were introduced by Hastie and Tibshirani (1986, 1990). They present a GAM as a generalized linear model where some of the covariates can enter the model in linear form, and some terms are smooth functions of covariates. This study is restricted to a special case: it uses an identity link and assumes normality in errors, which leads to additive models.

⁸For example, Greiner and Kauermann (2008), Ordás Criado et al. (2011), Bose et al. (2012) and Berlemann et al. (2015) apply (generalized) additive models.

at which sections are joined (and the end points) are the knots of the spline, and these locations must be chosen. The spline can be represented in terms of its values at the knots. Second, the amount of smoothness that functions f will have needs to be chosen. Overfit is to be avoided and, thus, departure from smoothness is penalized. The appropriate degree of smoothness for functions f can be estimated from the data by, for example, maximum likelihood, which is the chosen approach in this study for its robustness.

The package ‘mgcv’ has an automatic choice in the amount of smoothing and wide functionality.⁹ The relationship between the covariates and the response can be described graphically. Confidence bands for the model terms can be derived using Bayesian methods, and approximate p -values for model terms can be calculated. Models can be compared using information criteria such as the Akaike information criterion (AIC). For further details, see Appendix D and Wood (2006).

3. Results

3.1. Main results

In this subsection we provide some traditional linear models’ results (OLS with dummy variables) and compare them to more sophisticated additive models’ results. The additive models can be stated as:

$$\begin{aligned}
 RD_{I;relative} = & \theta_0 + f_1(I_{f;it}) + f_2(\gamma_{\{\epsilon\};it}) + f_3(government\ employment_{it}) \\
 & + f_4(dependency_{it}) + f_5(openness_{it}) + f_6(unemployment_{it}) \\
 & + f_7(union_{it}) + f_8(t) + u_i + v_{it},
 \end{aligned}$$

where i refers to a country and t to year, and θ_0 is the constant term. In our main analysis, the extent of redistribution (RD) is studied in relative terms as we discussed in Section 2.1. Functions f are smooth functions that are described using penalized cubic regression splines. Fixed country effects are denoted by u_i (traditional dummy variables), and the v_{it} are traditional error terms. The fixed country effects should take into account factors that stay constant over time within each country.

⁹The results in this study are obtained using the package ‘mgcv’, which includes a function ‘gam’. Basis construction for cubic regression splines is used. The maximum likelihood method is used in the selection of the smoothing parameters. The identifiability constraints (due to, for example, the model’s additive constant term) are taken into account by default. More details can be found in Wood (2006) and the R project’s web pages (<http://cran.r-project.org/>).

The additive model above describes the most flexible specification that is studied, and other specifications are special cases of it. In the traditional models all terms (functions f) are linear, but the additive models allow all functions f to be nonlinear with no pre-specified functional form. However, the additive models may also have some linear terms if the data suggest a linear structure. Thus, linear terms are reported for the additive models if linearity was suggested in the initial stage of model fitting. In reporting our results, graphical illustrations are used for nonlinear terms. In comparison, the interpretation of linear terms is straightforward, and these terms are not plotted.

Table 2 reports our main results. The information criteria show that the additive models fit the data better than the corresponding traditional models. However, in many cases the traditional and additive models give qualitatively similar information regarding the variables of interest. First, $Gini_f$ is positively associated with the extent of redistribution ($RD_{Gini;relative}$), but models (2) and (4) show that the relationship may be more complex than a linear association; see the change in the slope of the function in the topmost plots, (a) and (b), in Figure 4. The percentile ratio $P90/P50_f$ also correlates positively with the extent of redistribution in the upper half of the distribution ($RD_{P90/P50;relative}$). Second, the taste for redistribution (γ_ϵ) is in linear, negative association to the extent of redistribution; linear association is found in all the models in Table 2. These findings accord qualitatively with the Mirrlees model. We can also see that the main results for I_f and γ_ϵ are not sensitive to the inclusion of a wide range of control variables: the signs of the slopes do not change after adding controls.

Moreover, we find that government employment is statistically significantly and positively linked with the extent of redistribution; see also plots (d) and (e) in Figure 4. We also find that our empirical models are not able to capture all changes in $RD_{I;relative}$ over time: the shape of $f(year)$ is shown in plots (i)–(k) in Figure 4. This implies that even the broadest models, (4) and (8), do not capture all time-varying factors that relate to the extent of redistribution.

3.2. Sensitivity checks

The remainder of this section provides information about our results' sensitivity. First, we investigate our findings' robustness with respect to the chosen elasticity parameter. Second, we investigate the sensitivity of our findings to leaving some countries out of the sample. Third, we discuss our results when we change the specification so that the same

Table 2: Results of models where the dependent variable is $RD_{I,relative}$, where I refers to inequality measure. All models include country dummies and have $N = 105$. Constant terms and country fixed effects are not reported. The coefficients (and standard errors) are provided for traditional (i.e., OLS with dummies) models. The coefficients for the linear terms in the additive models are also provided to help the reader compare the models. Figure 4 shows graphical illustrations of the smooth functions that are nonlinear.

	Dependent variable: $RD_{Gini,relative}$			Dependent variable: $RD_{P90/P50,relative}$				
	traditional (1)	additive (2)	traditional (3)	additive (4)	traditional (5)	additive (6)	traditional (7)	additive (8)
$Gini_f$	0.574*** (0.132)	$f(Gini_f)$ *** See Fig. 4(a)	0.530*** (0.167)	$f(Gini_f)$ *** See Fig. 4(b)	-	-	-	-
$P90/P50_f$	-	-	-	-	20.540*** (1.406)	$f(P90/P50_f)$ *** See Fig. 4(c)	21.608*** (1.610)	20.023*** (1.369)
$\gamma_{\{\epsilon=0.20\}}$	-6.818*** (1.922)	-7.107*** (1.647)	-4.949** (2.135)	-3.921** (1.845)	-5.277*** (1.648)	-5.521*** (1.572)	-4.204** (1.731)	-2.566* (1.610)
<i>government employment</i>	-	-	0.228 (0.232)	$f(gov't empl.)$ *** See Fig. 4(d)	-	-	0.320* (0.191)	$f(gov't empl.)$ *** See Fig. 4(e)
<i>dependency</i>	-	-	-0.464** (0.203)	-0.195 (0.230)	-	-	-0.353** (0.167)	-0.082 (0.181)
<i>openness</i>	-	-	0.039 (0.027)	$f(open)$ ** See Fig. 4(f)	-	-	0.039* (0.022)	0.034* (0.019)
<i>unemployment</i>	-	-	0.114 (0.124)	0.052 (0.107)	-	-	-0.109 (0.096)	-0.134 (0.093)
<i>union</i>	-	-	0.006 (0.100)	$f(union)$ *** See Fig. 4(g)	-	-	0.047 (0.083)	$f(union)$ *** See Fig. 4(h)
year effects	0.128** (0.051)	$f(year)$ *** See Fig. 4(i)	0.068 (0.064)	$f(year)$ ** See Fig. 4(j)	0.134*** (0.040)	0.146*** (0.041)	0.085* (0.047)	$f(year)$ *** See Fig. 4(k)
AIC	503.3	465.2	499.0	432.4	469.2	461.1	458.5	407.6

***, **, *, ' denote significance at the 1, 5, 10 and 15% levels, respectively.

Note regarding the additive models: The smooth terms' significance levels are based on approximate F-tests.

According to approximate F-tests, the additive models are preferred to their traditional counterparts at the 5% significance level. Moreover, the approximate F-tests guide towards choosing the broadest models, (4) and (8).

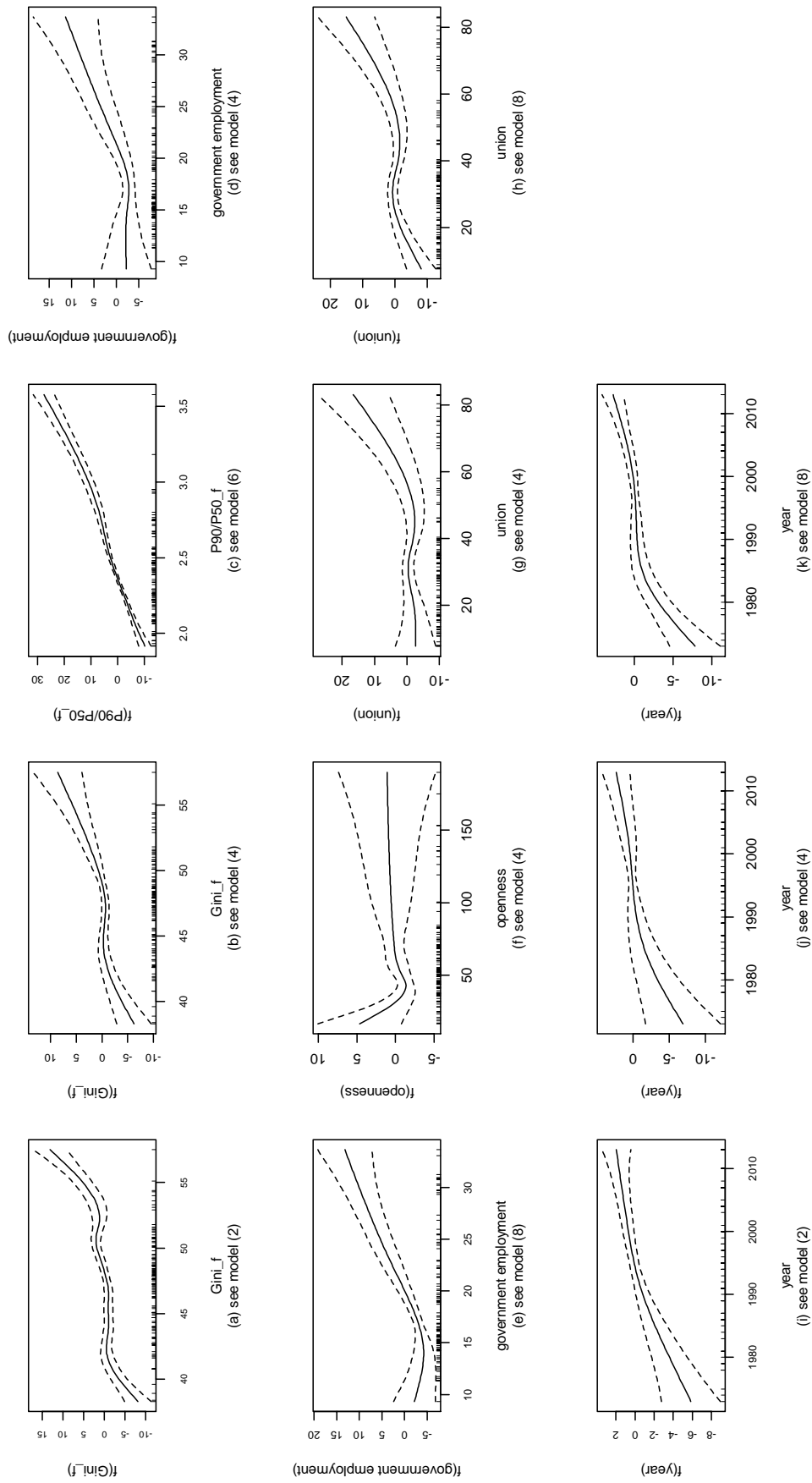


Figure 4: Illustrations of the smooth functions f in the additive models of Table 2. The dependent variable in models (2) and (4) is $RD_{Gini;relative}$, whereas the dependent variable in models (6) and (8) is $RD_{P90/P50;relative}$. The slopes of the functions are of interest. The plots also show the 95% confidence bands (dashed) and the covariate values as a rug plot along the horizontal axis.

explanatory variable (inequality measure) is not used to construct the dependent variable (redistribution). Finally, we use an alternative way to measure the extent of redistribution to check if this affects our main conclusions.

Measuring the government's taste for redistribution (γ) is not an easy task, and for this reason, we tested alternative values for the elasticity (ϵ). The above results were for the case $\gamma_{\{\epsilon=0.20\}}$ ($N = 105$). In our alternative models, we studied cases (a) $\epsilon = 0.10$ ($N = 120$), (b) $\epsilon = 0.15$ ($N = 114$) and (c) $\epsilon = 0.25$ ($N = 94$).¹⁰ In case (c) we were left with a very small sample size, and some of our empirical results were not statistically significant. In cases (a) and (b) with lower elasticities, our main results were qualitatively similar to those discussed earlier in this paper. Appendix C provides details.

Our second sensitivity check is related to the fairly small sample size. Because our main models included only 14 countries, we checked whether some groups of countries drive the main results. We did these investigations by using the specifications of Table 2, leaving each country group out of the sample (one group at a time). The countries were categorised into three groups, as in Figures 2–3. Only after dropping the Anglo-Saxon countries from the sample did we find that $Gini_f$ was very nonlinearly linked to $RD_{Gini;relative}$.¹¹ Otherwise, we found that our main findings on factor-income inequality and redistributive preference variables are fairly robust.

Table 3: Sensitivity checks: alternative additive model specifications. All models have $N = 105$. The coefficients (and standard errors) are provided for the linear terms. Figure 5 shows graphs of the reported smooth functions that are not linear.

	Alternative I_f as explanatory variable		Alternative definition of RD	
	Depending variable:		Depending variable:	
	$RD_{Gini;relative}$ (9)	$RD_{P90/P50;relative}$ (10)	$RD_{Gini;absolute}$ (11)	$RD_{P90/P50;absolute}$ (12)
$Gini_f$	-	$f(Gini_f)^{***}$ See Fig. 5(a)	$f(Gini_f)^{***}$ See Fig. 5(b)	-
$P90/P50_f$	3.545** (1.709)	-	-	$f(P90/P50_f)^{***}$ See Fig. 5(d)
$\gamma_{\{\epsilon=0.20\}}$	$f(\gamma)^{**}$ See Fig. 5(c)	-4.697** (1.988)	-1.342* (0.704)	-0.060' (0.040)

***, **, *, ' indicate significance at the 1, 5, 10 and 15% levels, respectively.

Note: The smooth terms' significance levels are based on approximate F-tests.

All models include country dummies and the following controls (some enter the model in nonlinear form): share of government employment, dependency rate, openness, unemployment rate, trade union density and a flexible term for year.

¹⁰As the elasticity increases, the number of observations in our data set decreases. This happens because we are limited to using values $\gamma \geq 0$.

¹¹To be precise, when only the Nordic and Continental European countries were included in the sample, the association between $RD_{Gini;relative}$ and $Gini_f$ resembled the letter M.

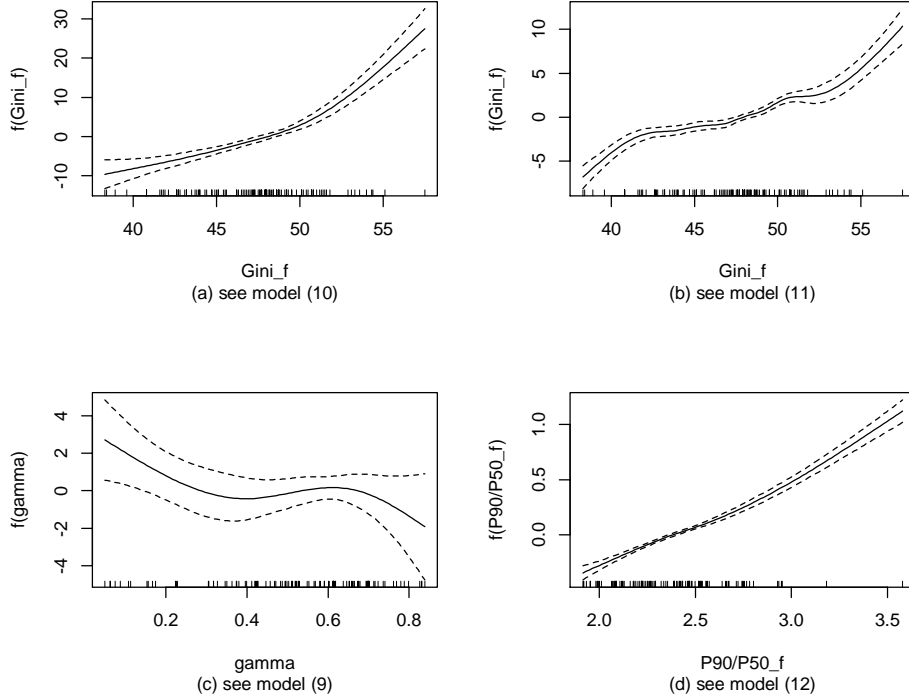


Figure 5: Illustration of the smooth functions f in the additive models of Table 3. The plots also show the 95% confidence bands (dashed) and the covariate values as a rug plot along the horizontal axis.

As a third sensitivity check, we estimated models where we did not use the same inequality indicator on both sides of the estimation equation. That is, we tested models with the alternative factor-income inequality measure. Table 3 reports two examples of these specifications, see models (9) and (10). Both models' results are qualitatively similar to our main findings in Table 2; factor-income inequality is positively linked with redistribution, whereas γ is negatively associated with redistribution.

Finally, we checked how our results change if the extent of redistribution is measured in absolute terms. Table 3 reports models (11) and (12) where $RD_{I;absolute} = I_f - I_d$ is the dependent variable. Again, the results are qualitatively similar to our main findings.

4. Conclusions

This paper examined the relationship between the extent of redistribution and the components of the Mirrlees model. To describe income inequality and redistribution, we used the Gini coefficients and the P90/P50 percentile ratios calculated from the LIS database. We also collected data from various other sources and constructed a measure of redistributive preferences by utilising the optimal top tax formula. Instead of relying solely on linear specifications in our empirical models, we also utilised penalized spline methods to allow nonlinearities in a flexible manner. We found a positive link between

inherent (factor-income) inequality and the extent of redistribution. Moreover, we found a significant association between the extent of redistribution and government's taste for redistribution. These empirical results are qualitatively in line with the Mirrlees model.

Appendix A. Descriptive statistics and data sources

Table A.4: Summary statistics of data used in models of Tables 2–3. Data sources are listed below.

Variable	N	min	mean	max
redistribution: $RD_{Gini;relative}$	105	23.61	36.71	55.62
redistribution: $RD_{P90/P50;relative}$	105	5.52	20.25	46.25
$Gini_f$	105	38.30	47.14	57.50
$P90/P50_f$	105	1.92	2.38	3.58
redistributive preferences $\gamma_{\{\epsilon=0.20\}}$	105	0.05	0.52	0.84
government employment	105	9.33	19.15	33.65
dependency rate	105	30.30	33.53	39.55
openness	105	16.41	62.71	190.11
unemployment rate	105	1.01	8.36	26.19
trade union density	105	7.67	35.06	83.14
redistribution: $RD_{Gini;absolute}$	105	9.60	17.35	27.90
redistribution: $RD_{P90/P50;absolute}$	105	0.11	0.49	1.66

List of data sources and definitions:

- Income inequality (I): $Gini_f$, $Gini_d$, $P90/P50_f$ and $P90/P50_d$ are from the Luxembourg Income Study (LIS) database (2017); subscript ‘ f ’ refers to factor incomes and ‘ d ’ to disposable incomes.
- Redistribution: calculated using the I_f and I_d variables (described above). Absolute measures calculated as $RD_{I;absolute} = I_f - I_d$, and relative measures calculated as $RD_{I;relative} = 100(I_f - I_d)/I_f$.
- Redistributive preference γ , using the optimal top tax formula $\tau = (1 - \gamma)/(1 - \gamma + \alpha\epsilon) \Leftrightarrow \gamma = 1 - \tau\alpha\epsilon/(1 - \tau)$. Top income tax rates (τ) from Piketty et al. (2014), OECD (2017) and the Association of Finnish Local and Regional Authorities (2017): Piketty et al. data are used for years up to 2010; the OECD data are used to extend series up to 2013; as an exception, the whole Finnish series has been updated using data from the OECD and the Association of Finnish Local and Regional Authorities. Pareto coefficients (α) are calculated using the relative shares of top 10% and top 1% income shares from the World Inequality Database (2017). These two series were available for all countries in our sample. To create longer series without breaks, we have imputed data in two cases: (1) when the top income share series begins (ends) one year later (earlier) compared to our data from the LIS database, we repeat the closest value for that year; (2) when there are at most three consecutive observations missing in the series, but we have data from the LIS database, we use linear interpolation. We assume constant elasticity and study cases $\epsilon = \{0.10, 0.15, 0.20, 0.25\}$; in our preferred specifications we assume $\epsilon = 0.20$.
- Government employment as per cent of total employment. Source: OECD (2016, 2017) and Eurostat (2017).
- Dependency rate: share of population who are 14 years or under or 65 years or over, as per cent of total population. Source: OECD (2017).
- Openness: the sum of exports and imports as per cent of GDP. Source: OECD (2017).
- Unemployment rate as per cent of civilian labour force. Source: OECD (2017).
- Trade union density, as percentage. Source: OECD (2017).

Appendix B. Additional descriptive figure of $RD_{P90/P50}$

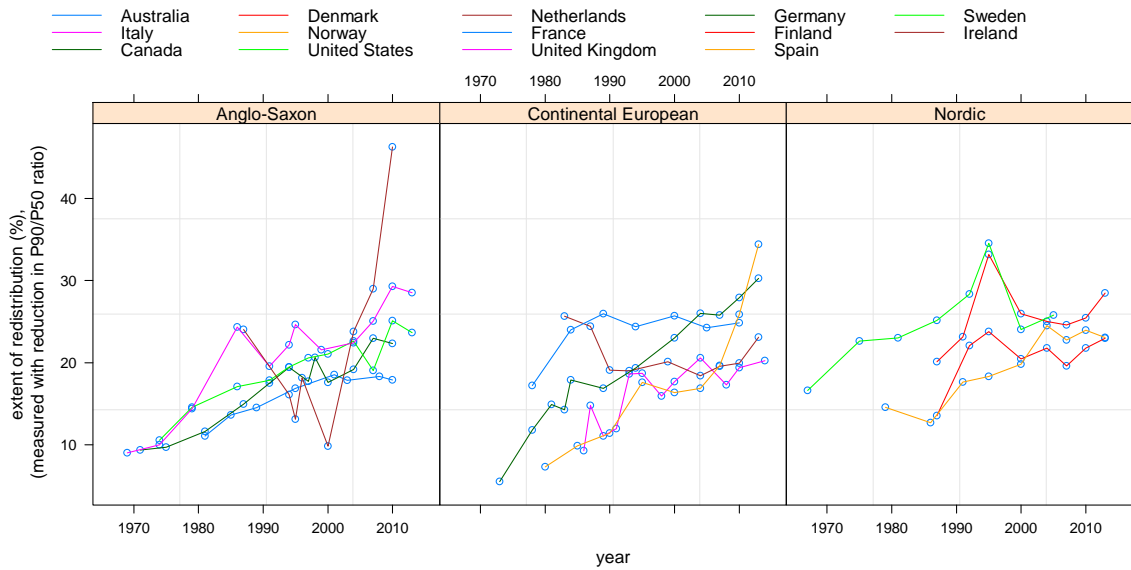


Figure B.6: Evolution of the extent of redistribution when redistribution is measured in relative terms: $RD_{P90/P50;relative}$ (14 advanced countries, unbalanced data, years 1967–2013). Calculations based on LIS database. More information can be found in Appendix A.

Appendix C. Sensitivity of results with respect to chosen ϵ

Table C.5: Sensitivity checks: alternative values for the elasticity parameter ϵ . This table provides selected results of additive model specifications. The coefficients (and standard errors) are provided for the linear terms. Figure C.7 shows graphs of the smooth functions that are not linear.

	Depending variable: $RD_{Gini;relative}$			Depending variable: $RD_{P90/P50;relative}$		
	$\epsilon = 0.10$ $N = 120$ (13)	$\epsilon = 0.15$ $N = 114$ (14)	$\epsilon = 0.25$ $N = 94$ (15)	$\epsilon = 0.10$ $N = 120$ (16)	$\epsilon = 0.15$ $N = 114$ (17)	$\epsilon = 0.25$ $N = 94$ (18)
$Gini_f$	$f(Gini_f)'$ See Fig. C.7(a)	$f(Gini_f)^{***}$ See Fig. C.7(b)	$f(Gini_f)^{***}$ See Fig. C.7(c)	-	-	-
$P90/P50_f$	-	-	-	22.078*** (1.477)	22.163*** (1.472)	21.765*** (1.588)
$\gamma_{\{\epsilon\}}$	-4.999* (2.694)	-5.596** (2.309)	0.050 (2.058)	-6.399*** (2.018)	-3.609** (1.657)	2.041 (2.153)

***, **, *, ' indicate significance at the 1, 5, 10 and 15% levels, respectively.

Note: The smooth terms' significance levels are based on approximate F-tests.

All models include country dummies and the following controls (some enter the model in nonlinear form): share of government employment, dependency rate, openness, unemployment rate, trade union density and a flexible term for year.

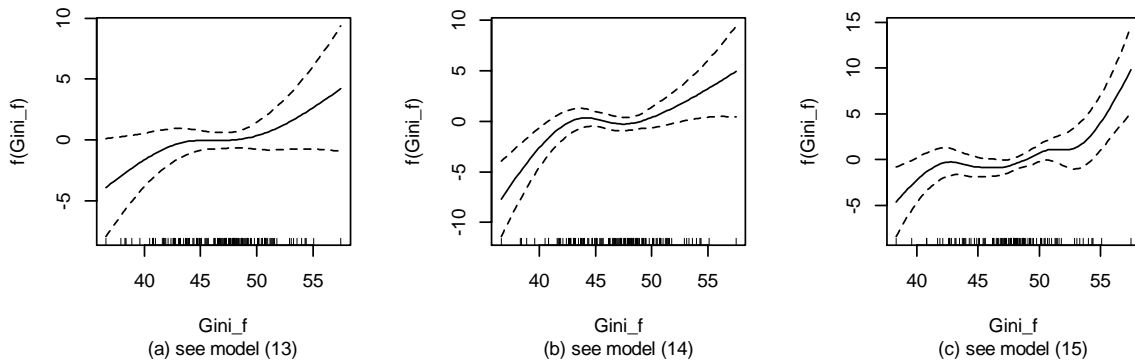


Figure C.7: Illustration of the smooth functions f in the additive models of Table C.5. The plots also show the 95% confidence bands (dashed) and the covariate values as a rug plot along the horizontal axis.

Appendix D. Supplementary information about the estimation method

Consider a model containing only one smooth function of one covariate: $y_i = f(x_i) + \epsilon_i$, where ϵ_i are i.i.d. $N(0, \sigma^2)$ random variables. To estimate function f here, f is represented so that the model becomes a linear model. This is possible by choosing a basis, defining the space of functions of which f (or a close approximation to it) is an element. In practice, one chooses basis functions, which are treated as known.

Assume that the function f has a representation $f(x) = \sum_{j=1}^k b_j(x)\beta_j$, where β_j are unknown parameters and $b_j(x)$ are known basis functions. Using a chosen basis for f implies that we have a linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where the model matrix \mathbf{X} can be represented using basis functions such as those in the cubic regression spline basis. The departure from smoothness can be penalized with $\int f''(x)^2 dx$. The penalty $\int f''(x)^2 dx$ can be expressed as $\boldsymbol{\beta}^T \mathbf{S}\boldsymbol{\beta}$, where \mathbf{S} is a coefficient matrix that can be expressed in terms of the known basis functions.

Accordingly, the penalized regression spline fitting problem is to minimise $\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda\boldsymbol{\beta}^T \mathbf{S}\boldsymbol{\beta}$, with respect to $\boldsymbol{\beta}$. The problem of estimating the degree of smoothness is a problem of estimating the smoothing parameter λ . In the estimation, one faces a bias–variance trade-off: on the one hand, the bias should be small, but on the other hand, the fit should be smooth. One needs to compromise between the two extremes. $\lambda \rightarrow \infty$ results in a straight line estimate for f , and $\lambda = 0$ leads to an unpenalized regression spline estimate. The penalized least squares estimator of $\boldsymbol{\beta}$, given λ , is $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{S})^{-1} \mathbf{X}^T \mathbf{y}$. Thus, the expected value vector is estimated as $\widehat{\mathbf{E}(\mathbf{y})} = \hat{\boldsymbol{\mu}} = \mathbf{A}\mathbf{y}$, where $\mathbf{A} = \mathbf{X}(\mathbf{X}^T \mathbf{X} + \lambda \mathbf{S})^{-1} \mathbf{X}^T$ is called an influence matrix.

This setting can be augmented to include several covariates and smooths. Given a basis, an additive model is simply a linear model with one or more associated penalties. The size of basis dimension for each smooth is usually not critical in estimation, because it only sets an upper limit on the flexibility of a term. Smoothing parameters control the effective degrees of freedom, and the effective degrees of freedom can be used to measure the flexibility of a model. See Wood (2006) for more discussion.

References

- Ahamada, I. & E. Flachaire (2013). *Non-Parametric Econometrics*. Oxford: Oxford University Press.
- Ahmad, E. & N. Stern (1984). The theory of reform and Indian indirect taxes. *Journal of Public Economics* 25(3), 259–298.
- Alesina, A., E. Glaeser, & B. Sacerdote (2001). Why doesn't the US have a European-style welfare state? NBER Working Paper No. 8524.
- Atkinson, A.B. (1999). Is rising income inequality inevitable? A critique of transatlantic consensus. *Wider Annual Lectures 3*. Helsinki: UNU-WIDER.
- Atkinson, A.B. & A. Brandolini (2001). Promise and pitfalls in the use of "secondary" data-sets: Income inequality in OECD countries as a case study. *Journal of Economic Literature* 39(3), 771–799.
- Atkinson, A.B. & T. Piketty (Eds.) (2010). *Top Incomes – A Global Perspective*. Oxford: Oxford University Press.
- Atkinson, A.B., T. Piketty, & E. Saez (2011). Top incomes in the long run of history. *Journal of Economic Literature* 49(1), 3–71.
- Atkinson, A.B., L. Rainwater, & T.M. Smeeding (1995). *Income Distribution in OECD Countries; Evidence from the Luxembourg Income Study*. Paris: OECD.
- Bargain O., Dolls, M., Neumann, D., Peichl, A., & S. Siegloch (2014a). Comparing inequality aversion across countries when labor supply responses differ. *International Tax and Public Finance* 21(5), 845–873.
- Bargain O., Dolls, M., Neumann, D., Peichl, A., & S. Siegloch (2014b). Tax-benefit revealed social preferences in Europe and the US. *Annals of Economics and Statistics*, No. 113/114, 257–289
- Bastani, S. & J. Lundberg (2017). Political preferences for redistribution in Sweden. *Journal of Economic Inequality* 15, 345–367.
- Benabou, R. & E.A. Ok (2001). Social mobility and the demand for redistribution: The POUM hypothesis. *Quarterly Journal of Economics* 116(2), 447–487.
- Berlemann, M., S. Enkelmann, & T. Kuhlenkasper (2015). Unraveling the relationship between presidential approval and the economy: A multidimensional semiparametric approach. *Journal of Applied Econometrics* 30(3), 468–486.

- Bose, N., A.P. Murshid, & M.A. Wurm (2012). The growth effects of property rights: The role of finance. *World Development* 40(9), 1784–1797.
- Bourguignon, F. & A. Spadaro (2012). Tax–benefit revealed social preferences. *Journal of Economic Inequality* 10(1), 75–108.
- Christiansen, V. & E.S. Jansen (1978). Implicit social preferences in the Norwegian system of indirect taxation. *Journal of Public Economics* 10(2), 217–245.
- Deininger, K. & L. Squire (1996). A new data set measuring income inequality. *World Bank Economic Review* 10(3), 565–591.
- Finseraas, H. (2009). Income inequality and demand for redistribution: A multilevel analysis of European public opinion. *Scandinavian Political Studies* 32(1), 94–119.
- Georgiadis, A. & A. Manning (2012). Spend it like Beckham? Inequality and redistribution in the UK, 1983–2004. *Public Choice* 151(3–4), 537–563.
- Greiner, A. & G. Kauermann (2008). Debt policy in euro area countries: Evidence for Germany and Italy using penalized spline smoothing. *Economic Modelling* 25(6), 1144–1154.
- Hastie, T. & R. Tibshirani (1986). Generalized additive models (with discussion). *Statistical Science* 1(3), 297–318.
- Hastie, T.J. & R.J. Tibshirani (1990). *Generalized Additive Models*. New York: Chapman & Hall/CRC.
- Hendren, N. (2017). Efficient welfare weights. NBER working paper no. 20351 (revised version). National Bureau of Economic Research.
- Immervoll, H. & L. Richardson (2011). Redistribution policy and inequality reduction in OECD countries: What has changed in two decades? *OECD Social, Employment and Migration Working Papers*, No. 122, OECD Publishing. <http://dx.doi.org/10.1787/5kg5d1khjq0x-en> (January 24, 2016).
- Jacobs, B., E.L.W. Jongen, & F.T. Zoutman (2017). Revealed social preferences of Dutch political parties. *Journal of Public Economics* 156, 81–100.
- Kanbur, R. & M. Tuomala (1994). Inherent inequality and the optimal graduation of marginal tax rates. *Scandinavian Journal of Economics* 96, 275–282.
- Kleven, H.J. & C.T. Kreiner (2006). The marginal cost of public funds: Hours of work versus labor force participation. *Journal of Public Economics* 90, 1955–1973.

- Li, Q. & J.S. Racine (2006). *Nonparametric Econometrics: Theory and Practice*. Princeton, NJ: Princeton University Press.
- Lindert, P.H. (2000). Three centuries of inequality in Britain and America. In: A.B. Atkinson & F. Bourquignon (Eds.), *Handbook of Income Distribution*, Vol 1 (Chapter 3). Amsterdam: North-Holland.
- Lockwood, B.B. & M. Weinzierl (2016). Positive and normative judgments implicit in U.S. tax policy, and the costs of unequal growth and recessions. *Journal of Monetary Economics* 77, 30–47.
- Luebker, M. (2014). Income inequality, redistribution, and poverty: Contrasting rational choice and behavioral perspectives. *Review of Income and Wealth* 60(1), 133–154.
- Luxembourg Income Study (LIS) Database (2017). <http://www.lisdatacenter.org> (multiple countries; data compiled during July–August 2017). Luxembourg: LIS.
- Meltzer, A.H. & S.F. Richard (1981). A rational theory of the size of government. *Journal of Political Economy* 89(5), 914–927.
- Milanovic, B. (2000). The median-voter hypothesis, income inequality, and income redistribution: an empirical test with required data. *European Journal of Political Economy* 16(3), 367–410.
- Milanovic, B. (2010). Four critiques of the redistribution hypothesis: An assessment. *European Journal of Political Economy* 26, 147–154.
- Mirrlees, J.A. (1971). An exploration in the theory of optimal income taxation. *Review of Economic Studies* 38, 175–208.
- Mirrlees, J.A. (1974). Notes on welfare economics, information and uncertainty. In: M. Balch, D. McFadden, & S. Wu (Eds.), *Essays on Economic Behaviour under Uncertainty*. Amsterdam: North-Holland.
- Moene, K.O. & M. Wallerstein (2001). Inequality, social insurance, and redistribution. *American Political Science Review* 95(4), 859–874.
- Ordás Criado, C., S. Valente, & T. Stengos (2011). Growth and pollution convergence: Theory and evidence. *Journal of Environmental Economics and Management* 62(2), 199–214.
- Peltzman, S. (1980). The growth of government. *Journal of Law & Economics* 23(2), 209–287.

- Perotti, R. (1996). Growth, income distribution, and democracy: What the data say. *Journal of Economic Growth* 1(2), 149–187.
- Persson, M. (1995). Why are taxes so high in egalitarian societies? *Scandinavian Journal of Economics* 97(4), 569–580.
- Persson, T. & G. Tabellini (2002). Political economics and public finance. In: A. J. Auerbach & M. Feldstein (Eds.), *Handbook of Public Economics*, Vol. 3 (Chapter 24). Amsterdam: North-Holland.
- Piketty, T., E. Saez, & S. Stantcheva (2011). Optimal taxation of top labor incomes: A tale of three elasticities, NBER Working Paper No. 17616.
- Piketty, T., E. Saez, & S. Stantcheva (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy* 6(1), 230–271.
- Roberts, K. (1977). Voting over income tax schedules. *Journal of Public Economics* 8(3), 329–340.
- Rodrik, D. (1998). Why do more open economies have bigger governments? *Journal of Political Economy* 106(5), 997–1032.
- Romer, T. (1975). Individual welfare, majority voting, and the properties of a linear income tax. *Journal of Public Economics* 4, 163–185.
- Saez, E. (2001). Using elasticities to derive optimal tax rates. *Review of Economic Studies* 68, 205–229.
- Saez, E., J. Slemrod, & S. Giertz (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature* 50(1), 3–50.
- Saez, E., & S. Stantcheva (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review* 106(1), 24–45.
- Scervini, F. (2012). Empirics of the median voter: democracy, redistribution and the role of the middle class. *Journal of Economic Inequality* 10(4), 529–550.
- Spadaro, A., L. Piccoli, & L. Mangiavacchi (2015). Optimal taxation, social preferences and the four worlds of welfare capitalism in Europe. *Economica* 82(327), 448–485.
- Tanninen, H. & M. Tuomala (2005). Inherent inequality and the extent of redistribution in OECD countries. CESifo DICE Report 1/2005, 48–53.
- Tuomala, M. (1984). Optimal degree of progressivity under income uncertainty. *Scandinavian Journal of Economics* 86(2), 184–193.

- Tuomala, M. (1990). *Optimal Income Tax and Redistribution*. Oxford: Clarendon Press.
- Tuomala, M. (2016). *Optimal Redistributive Taxation*. Oxford: Oxford University Press.
- Varian, H. (1980). Redistributive taxation as social insurance. *Journal of Public Economics* 14, 49–68.
- Wood, S.N. (2006). *Generalized Additive Models: An Introduction with R*. Boca Raton, FL: Chapman & Hall/CRC.
- World Inequality Database (2017). <http://wid.world/> (data downloaded in November 2017).