Sonderdruck aus

FinanzArchiv
Public Finance Analysis

2
Volume 77
June
2021

Mohr Siebeck

Income Inequality, Government's Redistributive Preferences, and the Extent of Redistribution

Hannu Tanninen, Matti Tuomala, and Elina Tuominen*

Received 18 December 2019; in revised form 14 December 2020; accepted 18 February 2021

We examine empirically the relation of factor-income inequality and government's redistributive preferences to the extent of redistribution. In the challenging task of measuring taste for redistribution, we utilize the inverse-optimum approach. Our income-inequality and redistribution variables are constructed from the Luxembourg Income Study database, and for our redistributive-preference measure we have collected data from various sources. In addition to traditional linear specifications we use flexible methods. We study 14 advanced countries for approximately four decades and find that factor-income inequality and government's redistributive preferences are associated with the extent of redistribution as suggested by the numerical results of the Mirrlees model.

Keywords: income inequality, nonlinearity, preferences, redistribution

JEL classification: C 14, D 31, H 3

1. Introduction

The postwar history of income inequality in advanced countries can be divided, at least roughly, into two phases. From 1945 to about the mid 1980s, pretax inequality, or the inequality of factor incomes (incomes from earnings and capital, also called market income), 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. Specifically, we have witnessed a significant increase in top income shares in many advanced economies over the past three to four decades (e.g., Atkinson and Piketty, 2010). The Luxembourg Income

* Tanninen: University of Eastern Finland, Finland (hannu.tanninen@uef.fi); Tuomala: Tampere University, Finland (matti.tuomala@tuni.fi); Tuominen: Tampere University, Finland (elina.tuominen@tuni.fi). The authors thank Ronnie Schöb and two anonymous reviewers for their comments and suggestions. The authors also 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. Financial support by the Strategic Research Council (SRC) at the Academy of Finland (project no. 293120, "Work, Inequality and Public Policy") is gratefully acknowledged.

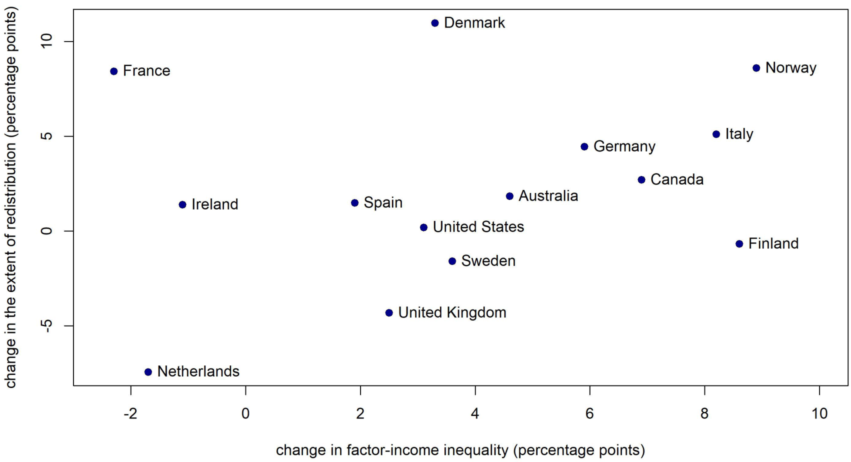
FinanzArchiv 77 (2021), 172–196 doi: 10.1628/fa-2021-0007
ISSN 0015-2218

Study (LIS) 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. Using LIS data, Immervoll and Richardson (2011) reported that in OECD countries governmental redistribution has become less effective in compensating increasing inequalities since the 1990s.

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.

Figure 1

Changes in Factor-income Inequality and the Extent of Redistribution



Note: 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 $RD_{\text{relative}} = 100(Gini_{\text{factor}} - Gini_{\text{disposable}})/Gini_{\text{factor}}$. 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 section 5.1.

There is now a considerable body of empirical literature seeking to explain the observed patterns of redistribution. The starting 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. Persson and

Tabellini (2002) provide a survey. A key element in this literature is the political mechanism – the median-voter theory – through which greater factor-income (market-income) 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 favored by the median voter.

Empirical studies have provided mixed evidence on the association between inequality and demand for redistribution (e.g., Perotti, 1996; Moene and Wallerstein, 2001; Finseraas, 2009). 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. In a more recent study, Milanovic (2010) emphasizes the median-voter hypothesis as only one possible mechanism linking initial inequality and redistribution. Earlier, Alesina et al. (2001) had 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. Luebker (2014) emphasized behavioral aspects in understanding the extent of redistribution.

Our approach in this paper differs from those above in that our inspiration stems from the optimal-income-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 of special interest when discussing the extent of redistribution. First is the concept of inherent (factor-income) inequality, reflecting, among other things, skilled–unskilled wage differentials, asset inequality, and social norms. If there is no intervention by the government, the factor-income inequality will be fully reflected in the disposable income. However, if the government wants to intervene – as seems to be the case in developed countries – it will find the second component of the Mirrlees model, the egalitarian objectives of the government, which are value judgments. Third, 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, and the extent of redistribution is reduced to the following three components: differences in pretax income, distributional objectives, and the responsiveness of income to taxation.[1]

1 There 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.

Numerical results in Kanbur and Tuomala (1994) and Tuomala (2016) show that the optimal income-tax–transfer system becomes more redistributive when inherent (market-income) 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–transfer system. Correspondingly, if inherent inequality decreases, governmental redistribution decreases. The results also suggest that in utilitarian, prioritarian (giving priority to the worse off), and maximin cases, an appropriate response to rising inequality is a shift towards a more redistributive income tax system.

Many of the earlier analytical and numerical results have focused on marginal tax rates, but the computational techniques can also be used to say something about average rates, which are arguably more important indicators of income tax progressivity. Based on numerical simulations, we know how average tax rates link to the different components in the Mirrlees model. However, we lack comparable data on average tax rates. As a result, we utilize the difference between factor-income inequality and disposable-income inequality to measure the extent of redistribution. This approach, despite its shortcomings, benefits from being consistent between countries and is more comparable with the broader empirical literature on the extent of redistribution. It also is worth emphasizing that we are not claiming that redistribution policy (pursued in the studied countries) would be the outcome of optimal policy choice. We utilize the inverse-optimum approach only to measure government's redistributive preferences. While we refer here to the optimal-tax framework, it is quite possible that our empirical model is consistent with some other theoretical frameworks, such as some political economy models.

To the best of our knowledge, the current empirical study is the first to explore the relationship between the extent of redistribution and the components of the Mirrlees framework, namely factor-income inequality, redistributive preferences, and disincentive costs. The LIS database provides data on both factor and disposable incomes for a number of advanced countries over four decades, and the current study focuses on 14 advanced economies. In addition to using Gini coefficients, we also utilize 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. In the case of this measure, we adopt the inverse-optimum approach and, more specifically, utilize the top tax rates to reveal the shape of implicit welfare weights. Our reason to apply the inverse-optimum approach to top earners is the availability of data. Our empirical results lend support to the numerical results of the Mirrlees model: factor-income inequality and redistributive preferences are associated with the extent of redistribution as expected.

The structure of the paper is as follows. Section 2 depicts the theoretical background, empirical specification, data, 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, and making wrong assumptions beforehand may bias the results. Section 3 provides our empirical results, including sensitivity checks. Finally, section 4 concludes.

2. Theoretical Background and Empirical Approach

2.1. Empirical Specification

As discussed in the introduction section, in the Mirrlees framework the extent of redistribution is reduced to the following three components: factor-income inequality, distributional objectives, and the behavioral responses to taxation. We will first provide some background before describing our data and methods.

Unlike the original Mirrlees model on labor income, we adopt the view that the taxation is based on the comprehensive income (i.e., taxing the sum of labor and capital income). In fact, in many countries most ordinary capital income, such as interest from a standard savings account, is taxed jointly with labor income. Moreover, in the case of the Nordic dual income tax model it can be difficult to distinguish between labor and capital income in practice. As Saez and Stantcheva (2016) show with a steady-state assumption, we can extend the Mirrlees model to a comprehensive income taxation model as well. It also turns out, as noted by Saez and Stantcheva, that in the case of the comprehensive income taxation the optimal tax formula takes the same form as in Mirrlees (1971) and Saez (2001).

Let $H(y)$ 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, the optimal nonlinear income tax (t_y) on total or comprehensive income satisfies

$$\frac{t_y}{1-t_y} = \frac{1-\gamma(y)}{\epsilon_y}\frac{1-H(y)}{yh(y)}, \tag{1}$$

where $\gamma(y)$ is the average welfare weight on individuals with total income higher than y, and ϵ_y is the elasticity of total income with respect to $1-t_y$ at income y. The last expression $\frac{1-H(y)}{yh(y)}$ describes the shape of the factor (pre-tax) income distribution. The welfare weights $\gamma(y)$ measure the social value of giving a unit of income to an individual with income y, relative to the social value of dividing it equally among all individuals. For example, in the classical utilitarian case $\gamma(y)$ is constant for all y (given the quasilinearity of preferences), and then the marginal tax rates are uniformly zero. The purpose of the formula (1) is mainly to present the components of the Mirrlees model. In addition, by applying and at the same time simplifying this tax formula for top earners we aim to trace the redistributive preferences in the studied 14 countries.

Next, we turn to examine empirically the relationship between the extent of redistribution and the components of the Mirrlees framework, with a focus on inherent (factor-income) inequality and government's redistributive preferences. The relationship that we explore can be expressed as follows:

$$RD = s(I_f, \gamma_{\{\epsilon\}}; x), \tag{2}$$

where RD is the extent of redistribution. As discussed in the introduction section, this is 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, $RD_{I;\text{relative}} = 100(I_f - I_d)/I_f$. In the sensitivity analysis, we also discuss the alternative case where the extent of redistribution is measured in absolute terms. The function s includes three components that reflect the ingredients of the Mirrlees model: I_f is the factor-income inequality measure, γ is the social marginal welfare weight for top earners (redistributive-preference measure), and the preference measure depends on ϵ, which is the weighted total income (labor and savings) elasticity. In addition, x denotes control variables.

We recognize that measuring government's redistributive preferences is a challenging task. The so-called inverse-optimum research provides one possible approach. This branch of research starts from the existing tax and transfers system and reverse-engineers it to obtain the underlying social preferences. Earlier contributions using this method are by Christiansen and Jansen (1978) and Ahmad and Stern (1984). More recently, detailed micro data on incomes and corresponding marginal tax rates have been used to study the social preferences implicit in tax–benefit systems (e.g., Kleven and Kreiner, 2006; Bourguignon and Spadaro, 2012; Bargain et al., 2014a,b; Spadaro et al., 2015; Lockwood and Weinzierl, 2016; Bastani and Lundberg, 2017; Hendren, 2017). Moreover, Jacobs et al. (2017) used this method to find the redistributive preferences of political parties implicit in the reform proposals. However, these studies typically focus on selected years, and most papers study one or selected countries only.

Following Atkinson (1995) and Diamond (1998), assuming away income effects[2] and constant elasticities,[3] we apply the optimal comprehensive tax formula (1) for top earners. Now γ is the average social marginal welfare weight of top-bracket individuals. It takes the following form:

$$\gamma = 1 - \frac{\tau \alpha \epsilon}{1 - \tau}, \tag{3}$$

where we use data on top income tax rates (τ), estimates for the Pareto coefficients ($\alpha = \frac{yh(y)}{1-H(y)}$), and values for total income elasticities (ϵ). The excellent Pareto fit of the top tail of the distribution has been verified in many countries and many periods, as summarized in Atkinson et al. (2011). Next we discuss our data in more detail.

2.2. Data

We have constructed our main income-distribution variables from the LIS database. LIS has harmonized micro data from (mostly) high- and middle-income countries, and the data are organized into different waves according to the date of the data. The database provides information on both factor and disposable incomes.[4] 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.[5] We use all available LIS data for which data on our other variables are available, and the resulting data set is not balanced.

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 $P90/P50$. Both measures are much used in inequality studies and provide somewhat complementary information. The Gini index is an overall measure of inequality, reflecting the behavior of the whole income

2 Diamond (1998, p. 85) writes: "This assumption seems appropriate at very high income levels, since people at the top of the income distribution are likely to leave large estates – with a linear utility of bequests neither consumption nor earnings vary with the exact level of estate." There also is some empirical (although indecisive) evidence suggesting that income effects are small (e.g., Gruber and Saez, 2002).

3 Diamond (1998) shows that if preferences are represented by a utility function $u = x - l^{1+\frac{1}{\varepsilon}}/\left(1+\frac{1}{\varepsilon}\right)$, where x is a composite consumption good and l hours worked, this presentation implies a labor supply function with a constant elasticity ε.

4 Most studies on inequality and redistribution have utilized 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).

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

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)
UK	50.8	2.57	(1986)	54.3	2.78	(1995)	53.3	2.76	(2004)
USA	46.2	2.47	(1986)	49.1	2.66	(1994)	49.3	2.74	(2004)

Note: Data source is the LIS database.
* This year's observation cannot be used in the empirical models of table 2, due to missing information in other variables.

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 the inequality of factor incomes has risen in most countries over the sample period.

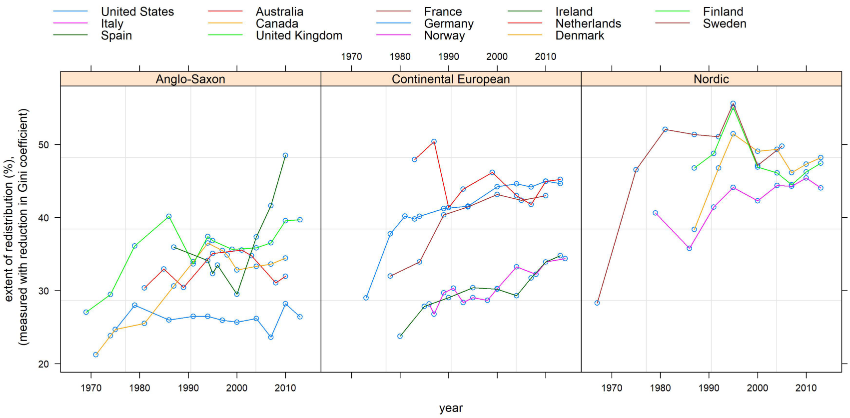
Figure 2 shows the development of the extent of redistribution ($RD_{Gini;\text{relative}}$) in the countries under investigation. The countries are categorized into three groups to provide a concise but readable illustration.[6] The figure shows that the extent of redistribution has increased modestly 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 $RD_{P90/P50;\text{relative}}$ is in appendix section 5.2 (figure 6).

As described by equation (3) in section 2.1, we have used the optimal top-tax formula in calculating government's taste for redistribution (γ), which depends on the top income tax rate (τ), Pareto coefficient (α), and total income elasticity (ϵ).[7] Our data on top income tax rates and Pareto coefficients are col-

6 The categorization 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 $N = 41$).

7 After calculating our estimates for γ, we only include nonnegative values in our data set.

Figure 2
Evolution of the Extent of Redistribution when Redistribution is Measured in Relative Terms: $RD_{Gini;relative}$



Note: 14 advanced countries, unbalanced data. Calculations based on LIS database. More information can be found in appendix section 5.1.

lected from various sources, such as Piketty et al. (2011, 2014) and the World Inequality Database (WID; 2017); see appendix section 5.1 for more detailed information. In estimating α, we have chosen to use the top income shares from the WID because LIS relies heavily on survey data. In comparison, WID combines multiple data sources such as national accounts and fiscal data and should therefore capture the evolution at the top of the distribution better than survey data.

It has been long recognized that behavioral responses to taxation are not confined to participation and hours worked. Feldstein (1995) proposed that we should examine the response of taxable income to changes in tax rates. Some behavioral responses of top incomes to top tax rates seem to be due not to a real change in economic activity and output, but simply to a relabeling of income outlays over various tax bases. We should also bear in mind that the taxable-income elasticity is not derived from immutable preferences, but is affected by the structure of the tax system. However, there is a great deal of uncertainty around these numbers.[8] Feldstein (1995) found very high elasticities, exceeding one, but subsequent research has generated considerably smaller estimates. In a survey on taxable-income elasticities, Saez et al.

8 Using simulation methods, Aronsson et al. (2017) assess the bias and precision of the prevalent methods used in the taxable-income elasticity studies. We are grateful to the reviewer for drawing our attention to this paper.

(2012) conclude that the best available longer-run elasticity estimates range from 0.12 to 0.40. Kleven and Schultz (2014) use Danish data and find that labor income elasticity is low (around 0.1) and that capital income elasticity is somewhat higher (around 0.2–0.3).

Of course, we would like to get estimates of the elasticity of taxable income (preferably of total income) that differ between countries and over time. However, the available data constrain us. As a result, we have investigated our results with some "reasonable" values for ϵ from the prior literature. In our preferred specifications we assume $\epsilon = 0.20$. Given the findings by Kleven and Schultz (2014) and that typically capital income dominates labor income among top income earners, our preferred estimate looks quite plausible. Figure 3 describes the evolution of government's taste for redistribution in the 14 countries of this study. Higher γ reflects lower willingness to redistribute 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-check section.

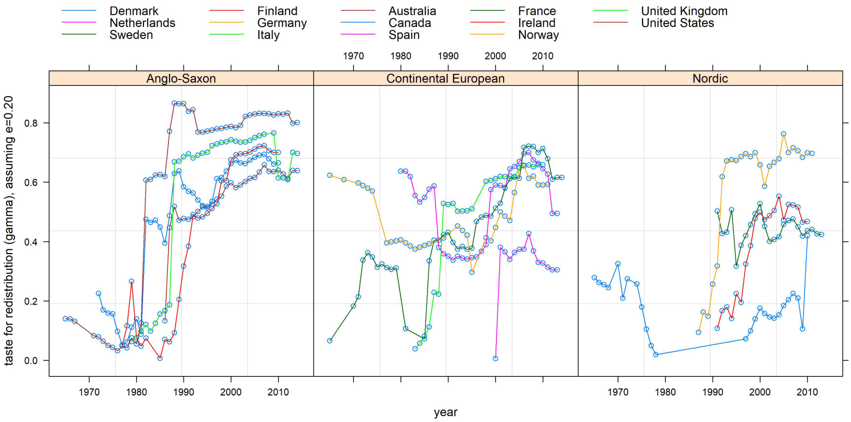
We report most of our results with a fairly broad set of controls. This approach stems from our acknowledging that the extent of redistribution in each country is a result of numerous factors. As discussed earlier, the approach in our paper is more an exploration than a strict test of the optimal tax framework. Government employment, dependency rate, and unemployment rate account for different economies' public expenditure requirements. High trade union density is often associated with a large welfare state, and it may proxy for other political variables as well (see, e.g., Baccaro, 2008, for more discussion). Moreover, it 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.[9] Summary statistics and a complete list of data sources and definitions are provided in appendix section 5.1.

2.3. Estimation Method

We do not impose linearity on all our empirical models. In our preferred specifications, we allow all continuous covariates to enter flexibly so that possible wrong functional forms would not bias our results. Our estimation approach

9 In addition, the authors acknowledge that full assessment of the extent of redistribution should also take account of various services that are publicly provided at less than market value. These are considerable in Nordic countries. Many of these items – health care, education, and social services – are very extensive.

Figure 3
Evolution of Redistributive Preferences



Note: 14 advanced countries, unbalanced data. In calculating the γ-values, we have assumed constant elasticity $\epsilon = 0.20$. Data are constructed from multiple sources, and more information is provided in appendix section 5.1.

is based on penalized cubic regression splines, although we acknowledge that there are numerous alternative approaches to flexible modeling, such as kernel estimation.[10] Moreover, due to the small sample size, we assume an additive structure instead of a fully nonparametric one.[11] The chosen method is accessible in that 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 utilized in the analysis. To estimate our additive models we use the established R software package mgcv, which has previously been utilized in economics studies on various topics.[12]

Additive models provide a flexible framework for investigating the relationship of inequality and redistributive preferences to redistribution. This study follows the approach presented in Wood (2006, 2017). The basic idea is that

10 Li and Racine (2007) describe nonparametric methods extensively, with the focus on kernels. Ahamada and Flachaire (2010) provide a concise overview of nonparametric methods.
11 Additive models are a special case of generalized additive models (GAMs). GAMs were introduced by Hastie and Tibshirani (1986, 1990). They presented a GAM as a generalized linear model where some of the covariates can enter 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.
12 For example, Greiner and Kauermann (2008), Ordás Criado et al. (2011), Bose et al. (2012), and Berlemann et al. (2015) apply (generalized) additive models.

the model consists of a sum of linear and smooth functions of covariates:

$$y_i = X_i^* \theta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}) + \cdots. \tag{4}$$

In the above presentation, y_i is the response variable (extent of redistribution), X_i^* is a row of the model matrix for any strictly parametric model components, θ 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, which is the approach adopted in this study, is to use cubic regression splines. 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 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 the 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 the functions f can be estimated from the data by, for example, maximum likelihood, which is the chosen approach in this study on account of its robustness.

The package mgcv has an automatic choice in the amount of smoothing and wide functionality.[13] 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 section 5.4 and Wood (2006, 2017).

3. Results

3.1. Main Results

In this subsection, we provide some traditional linear models' results (OLS with dummy variables) and compare them with more sophisticated additive

13 The results in this study are obtained using the package's 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, 2017) and the R project's web pages (http://cran.r-project.org/).

models' results. The additive models can be stated as

$$\begin{aligned} RD_{I;\text{relative}} = \theta_0 &+ f_1(I_{f;it}) + f_2(\gamma_{\{\epsilon\};it}) \\ &+ f_3(\textit{government employment}_{it}) + f_4(\textit{dependency}_{it}) \\ &+ f_5(\textit{openness}_{it}) + f_6(\textit{unemployment}_{it}) \\ &+ f_7(\textit{union}_{it}) + f_8(t) + u_i + v_{it}, \end{aligned} \tag{5}$$

where i refers to a country and t to a 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. The f's are smooth functions that are described using penalized cubic regression splines. Country fixed effects are denoted by u_i (traditional dummy variables), and the v_{it} are traditional error terms. The country fixed effects should take into account factors that stay constant over time within each country.

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 functions f are linear, but the additive models allow these functions to be nonlinear with no prespecified 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 ($RD_{Gini;\text{relative}}$), but models (2) and (4) show that the relationship may be more complex than a linear association; note 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 ($RD_{P90/P50;\text{relative}}$). Second, the taste for redistribution (γ_ϵ) is in linear, negative association with the extent of redistribution; linear association is found in all the models in table 2. These findings accord qualitatively with the implications of the Mirrlees model: higher factor-income inequality is linked to more redistribution, and higher γ is associated with less redistribution. We can also see that these 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.

Results for our control variables are not as robust as those for the main variables. We find that government employment is statistically significantly and

Table 2

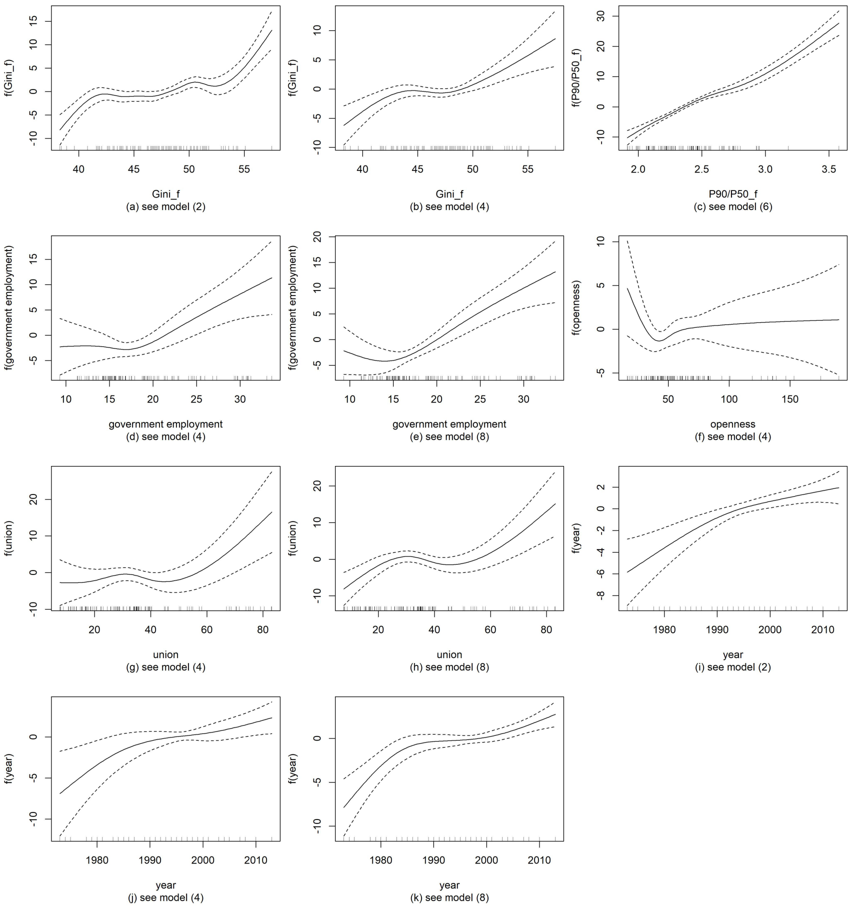
Results of Models where the Dependent Variable is $RD_{I;relative}$, in which I Refers to an Inequality Measure

	Dependent variable: $RD_{Gini;\text{relative}}$				Dependent variable: $RD_{P90/P50;\text{relative}}$			
	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)
government employment	–	–	0.228 (0.232)	f(*gov’t empl.*)*** See fig. 4(d)	–	–	0.320* (0.191)	f(*gov’t empl.*)*** See fig. 4(e)
dependency	–	–	−0.464** (0.203)	−0.195 (0.230)	–	–	−0.353** (0.167)	−0.082 (0.181)
openness	–	–	0.039 (0.027)	f(*open*)** See fig. 4(f)	-	–	0.039* (0.022)	0.034* (0.019)
unemployment	–	–	0.114 (0.124)	0.052 (0.107)	-	–	−0.109 (0.096)	−0.134 (0.093)
union	–	–	0.006 (0.100)	f(*union*)*** See fig. 4(g)	-	–	0.047 (0.083)	f(*union*)*** See fig. 4(h)
year	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

Note: ***, **, *, ’ denote significance at the 1, 5, 10, and 15% levels, respectively. All models include country dummies and have $N = 105$ (years 1973–2013). 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. The smooth terms’ significance levels are based on approximate F-tests. In addition, according to approximate F-tests, the additive models are preferred to their traditional counterparts at the 5% significance level. Tests also guide towards choosing the broadest models with controls.

Figure 4

Illustrations of the Smooth Functions f in the Additive Models of Table 2



Note: The dependent variable in models (2) and (4) is $RD_{Gini;\text{relative}}$, whereas the dependent variable in models (6) and (8) is $RD_{P90/P50;\text{relative}}$. 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.

positively linked with the extent of redistribution; see also plots (d) and (e) in figure 4. There also is some indication of a positive association between trade union density and the extent of redistribution, but the relationship appears nonlinear; see plots (g) and (h) in figure 4. In comparison, our results do not confirm the proposed positive link between openness and the extent of redistribution in the case of models (3) and (4). The unemployment variable is not statistically significant in any of the models of table 2, and the dependency rate fails to be a statistically significant variable after allowing flexible functional forms. We also find that our empirical models are not able to capture all changes in $RD_{I;\text{relative}}$ 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 main 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 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) $\epsilon = 0.25$ ($N = 94$).[14] In case (c), we were left with a very small sample size, and results were not statistically significant for our redistributive-preference measure. In cases (a) and (b) with lower elasticities, our main results were qualitatively similar to those discussed earlier in this paper. Appendix section 5.3 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 categorized into three groups, as in figures 2–3 and footnote 6. Only after dropping the Anglo-Saxon countries from the sample did we find that $Gini_f$ was very nonlinearly linked to $RD_{Gini;\text{relative}}$.[15] Otherwise, we found that our main findings on factor-income inequality and redistributive preferences are fairly robust, although not always statistically significant.

14 As the elasticity increases, the number of observations in our data set decreases. This happens because we use values $\gamma \geq 0$.

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

Table 3
Sensitivity Checks: Alternative Additive Model Specifications

	Alternative I_f as explanatory variable		Alternative definition of *RD*	
	Dependent variable:		Dependent variable:	
	$RD_{Gini;\text{relative}}$ (9)	$RD_{P90/P50;\text{relative}}$ (10)	$RD_{Gini;\text{absolute}}$ (11)	$RD_{P90/P50;\text{absolute}}$ (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)

Note: ***, **, *, ’ indicate significance at the 1, 5, 10, and 15% levels, respectively.
All models have $N = 105$ (years 1973–2013). The coefficients (and standard errors) are provided for the linear terms. Figure 5 shows graphs of the reported smooth functions that are not linear.
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 non-linear form): share of government employment, dependency rate, openness, unemployment rate, trade union density, and a flexible term for year.

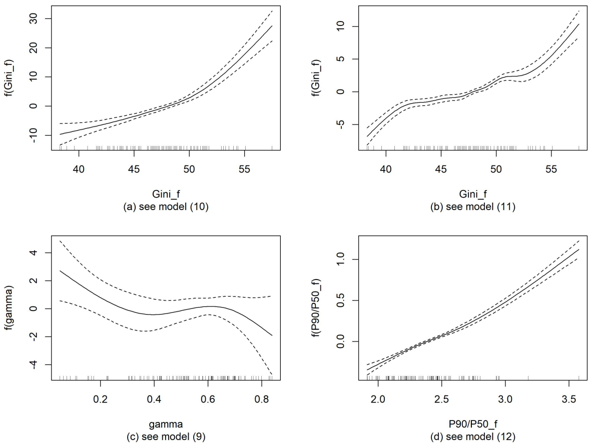
As a third sensitivity check, we estimated models where we did not use the same inequality indicator on both sides of the estimation equation. That is, we tested models with the alternative factor-income inequality measure. Table 3 reports two examples of these specifications; see models (9) and (10). 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 linked 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 $RD_{I;\text{absolute}} = I_f - I_d$ is the dependent variable. Again, the results are qualitatively similar to our main findings.

4. Conclusions

As discussed, we acknowledge that measuring government’s redistributive preferences is a challenging task. The Mirrlees framework – which has dominated the literature on optimal income taxation for decades – provides an interesting possibility to discuss the extent of redistribution in relation to these preferences. The main motivation behind this paper has not been to claim that redistribution policy of the studied countries is the outcome of optimal pol-

Figure 5
Illustration of the Smooth Functions f in the Additive Models of Table 3



Note: The plots also show the 95% confidence bands (dashed) and the covariate values as a rug plot along the horizontal axis.

icy choice, but to utilize this theoretical framework in assessing the taste for redistribution. Thus, in this paper, we have taken the first steps to examine empirically 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. In constructing a measure of redistributive preferences, we collected data from various other sources and utilized the optimal top-tax formula. Instead of relying solely on linear specifications in our empirical models, we also utilized penalized spline methods to allow nonlinearities in a flexible manner. We found – as did numerous empirical studies on inequality and redistribution before us – a positive link between factor-income inequality and the extent of redistribution. This result was clear in all our specifications. Moreover, we found a significant association between the extent of redistribution and our redistributive-preference measure; high γ was linked to less redistribution. These empirical results are qualitatively in accordance with the numerical results of the Mirrlees model.

5. Appendix

5.1. Descriptive Statistics and Data Sources

Table 4
Summary Statistics of Data Used in Models of Tables 2–3

Variable	N	min	mean	max
redistribution: $RD_{Gini;\text{relative}}$	105	23.61	36.71	55.62
redistribution: $RD_{P90/P50;\text{relative}}$	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: $RD_{Gini;\text{absolute}}$	105	9.60	17.35	27.90
redistribution: $RD_{P90/P50;\text{absolute}}$	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 $RD_{I;\text{absolute}} = I_f - I_d$, and relative measures calculated as $RD_{I;\text{relative}} = 100(I_f - I_d)/I_f$.
- Redistributive preference (γ), using the optimal top-tax formula $\gamma = 1 - \tau\alpha\epsilon/(1-\tau)$. Top income tax rates ($\tau$) are from Piketty et al. (2011, 2014) and OECD Tax Database (accessed: 2017-07-24). Piketty et al. data are used for years up to 2010; the OECD data are used to extend series further (series up to 2013 utilized in estimations). As an exception, the whole Finnish series is constructed using data from the OECD Tax Database (including Historical table I.1, accessed: 2017-09-20) and the Association of Finnish Local and Regional Authorities (www.kuntaliitto.fi, accessed: 2017-09-20). Pareto coefficients (α) are calculated using the relative shares of top 10% and top 1% income 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) than our data from the LIS database, we repeat the closest value for that year; (2) when there

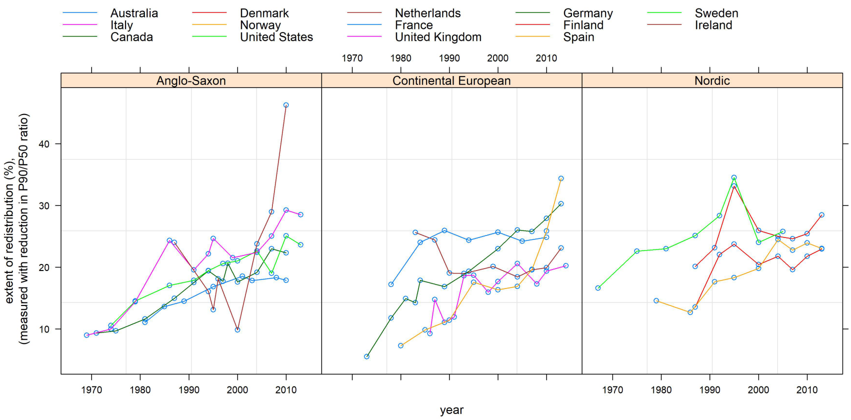
are at most three consecutive observations missing in the series, but we have data from the LIS database, we use linear interpolation. We assume constant elasticity and study cases $\epsilon = \{0.10, 0.15, 0.20, 0.25\}$; in our preferred specifications we assume $\epsilon = 0.20$. We include only $\gamma \geq 0$ in our data set.

- Government employment as percentage of total employment. Source: OECD Economic Outlook No 100 – November 2016 (accessed: 2017-06-02), with supplementary data from OECD Economic Outlook No 75 – June 2004 (data for several countries; accessed: 2016-02-12) and Eurostat (data for Germany; accessed: 2017-01-11) .
- Dependency rate: share of population who are 14 years or under or 65 years or over, as percentage of total population. Source: OECD Population and Labour Force Dataset (accessed: 2017-07-06) .
- Openness: the sum of exports and imports as percentage of GDP. Source: OECD National Accounts database (accessed: 2017-07-06) .
- Unemployment rate as percentage of civilian labor force. Source: OECD ALFS Summary tables (accessed: 2017-08-24) .
- Trade union density, as percentage. Source: OECD Labour Database (accessed: 2017-08-24) .

5.2. Additional Descriptive Figure for $RD_{P90/P50}$

Figure 6

Evolution of the Extent of Redistribution when Redistribution is Measured in Relative Terms: $RD_{P90/P50;relative}$



Note: 14 advanced countries, unbalanced data. Calculations based on LIS database. More information can be found in appendix section 5.1.

5.3. Sensitivity of Results with Respect to Chosen ϵ

Table 5

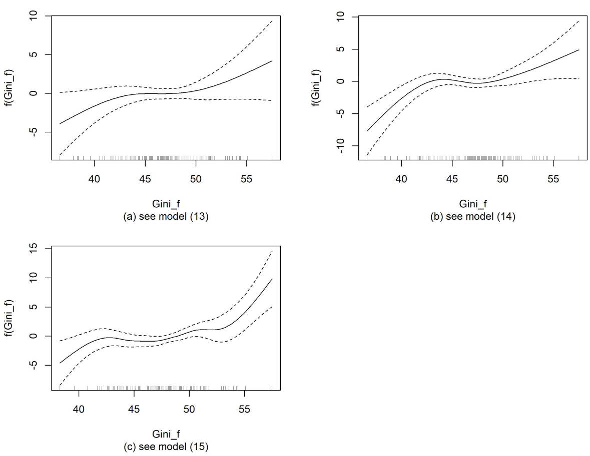
Sensitivity Checks: Alternative Values for the Elasticity Parameter ϵ

	Dependent variable: $RD_{Gini;\text{relative}}$			Dependent variable: $RD_{P90/P50;\text{relative}}$		
	$\epsilon = 0.10$ $N = 120$ (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. 7(a)	$f(Gini_f)$*** See fig 7(b)	$f(Gini_f)$*** See fig. 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)

Note: ***, **, *, ’ indicate significance at the 1, 5, 10, and 15% levels, respectively.
This table provides selected results of additive model specifications. The coefficients (and standard errors) are provided for the linear terms. Figure 7 shows graphs of the smooth functions that are not linear.
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 7

Illustration of the Smooth Functions f in the Additive Models of Table 5



Note: The plots also show the 95% confidence bands (dashed) and the covariate values as a rug plot along the horizontal axis.

5.4. 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 the 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 that 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 $y = X\beta+\epsilon$, where the model matrix X can be represented using basis functions such as those in the cubic regression spline basis. The departure from smoothness can be penalized with $\int f''(x)^2dx$. The penalty $\int f''(x)^2dx$ can be expressed as $\beta^T S\beta$, where 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 minimize $\|y - X\beta\|^2 + \lambda\beta^T S\beta$, with respect to β. 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. $\lambda \longrightarrow \infty$ results in a straight-line estimate for f, and $\lambda = 0$ leads to an unpenalized regression spline estimate. The penalized least-squares estimator of β, given λ, is $\hat{\beta} = (X^T X + \lambda S)^{-1} X^T y$. Thus, the expected-value vector is estimated as $\widehat{\mathbf{E}(y)} = \hat{\mu} = Ay$, where $A = X(X^T X + \lambda S)^{-1} X^T$ 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 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, 2017) for more discussion.

References

Ahamada, I., and Flachaire, E. (2010), Non-Parametric Econometrics, Oxford University Press, Oxford.

Ahmad, E., and Stern, N. (1984), The Theory of Reform and Indian Indirect Taxes, Journal of Public Economics 25, 259–298.

Alesina, A., Glaeser, E., and Sacerdote, B. (2001), Why Doesn't the United States Have a European-Style Welfare State?, Brookings Papers on Economic Activity 2001:2, 187–254.

Aronsson, T., Jenderny, K., and Lanot, G. (2017), The Quality of the Estimators of the ETI, Umeå School of Business and Economics, Umeå Economic Studies 955.

Atkinson, A. B. (1995), Public Economics in Action, Oxford University Press, Oxford.

Atkinson, A. B., and Brandolini, A. (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, 771–799.

Atkinson, A. B., and Piketty, T., Eds. (2010), Top Incomes: A Global Perspective, Oxford University Press, Oxford.

Atkinson, A. B., Piketty, T., and Saez, E. (2011), Top Incomes in the Long Run of History, Journal of Economic Literature 49, 3–71.

Baccaro, L. (2008), Labour Institutions and Inequality, in: Torres, R. (Ed.), World of Work Report 2008: Income Inequalities in the Age of Financial Globalization, International Institute for Labour Studies, 71–114.

Bargain, O., Dolls, M., Neumann, D., Peichl, A., and Siegloch, S. (2014a), Comparing Inequality Aversion across Countries when Labor Supply Responses Differ, International Tax and Public Finance 21, 845–873.

Bargain, O., Dolls, M., Neumann, D., Peichl, A., and Siegloch, S. (2014b), Tax-Benefit Revealed Social Preferences in Europe and the US, Annals of Economics and Statistics 113/114, 257–289.

Bastani, S., and Lundberg, J. (2017), Political Preferences for Redistribution in Sweden, Journal of Economic Inequality 15, 345–367.

Berlemann, M., Enkelmann, S., and Kuhlenkasper, T. (2015), Unraveling the Relationship between Presidential Approval and the Economy: A Multidimensional Semiparametric Approach, Journal of Applied Econometrics 30, 468–486.

Bose, N., Murshid, A. P., and Wurm, M. A. (2012), The Growth Effects of Property Rights: The Role of Finance, World Development 40, 1784–1797.

Bourguignon, F., and Spadaro, A. (2012), Tax-Benefit Revealed Social Preferences, Journal of Economic Inequality 10, 75–108.

Christiansen, V., and Jansen, E. S. (1978), Implicit Social Preferences in the Norwegian System of Indirect Taxation, Journal of Public Economics 10, 217–245.

Deininger, K., and Squire, L. (1996), A New Data Set Measuring Income Inequality, World Bank Economic Review 10, 565–591.

Diamond, P. A. (1998), Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates, American Economic Review 88, 83–95.

Feldstein, M. (1995), The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act, Journal of Political Economy 103, 551–572.

Finseraas, H. (2009), Income Inequality and Demand for Redistribution: A Multilevel Analysis of European Public Opinion, Scandinavian Political Studies 32, 94–119.

Georgiadis, A., and Manning, A. (2012), Spend it Like Beckham? Inequality and Redistribution in the UK, 1983–2004, Public Choice 151, 537–563.

Greiner, A., and Kauermann, G. (2008), Debt Policy in Euro Area Countries: Evidence for Germany and Italy Using Penalized Spline Smoothing, Economic Modelling 25, 1144–1154.

Gruber, J., and Saez, E. (2002), The Elasticity of Taxable Income: Evidence and Implications, Journal of Public Economics 84, 1–32.

Hastie, T., and Tibshirani, R. (1986), Generalized Additive Models (with Discussion), Statistical Science 1, 297–318.

Hastie, T. J., and Tibshirani, R. J. (1990), Generalized Additive Models, Chapman & Hall/CRC, New York.

Hendren, N. (2017), Efficient Welfare Weights, NBER Working Paper No. 20351 (revised version).

Immervoll, H., and Richardson, L. (2011), Redistribution Policy and Inequality Reduction in OECD Countries: What Has Changed in Two Decades?, OECD Social, Employment and Migration Working Papers No. 122.

Jacobs, B., Jongen, E. L. W., and Zoutman, F. T. (2017), Revealed Social Preferences of Dutch Political Parties, Journal of Public Economics 156, 81–100.

Kanbur, R., and Tuomala, M. (1994), Inherent Inequality and the Optimal Graduation of Marginal Tax Rates, Scandinavian Journal of Economics 96, 275–282.

Kleven, H. J., and Kreiner, C. T. (2006), The Marginal Cost of Public Funds: Hours of Work versus Labor Force Participation, Journal of Public Economics 90, 1955–1973.

Kleven, H. J., and Schultz, E. A. (2014), Estimating Taxable Income Responses Using Danish Tax Reforms, American Economic Journal: Economic Policy 6, 271–301.

Li, Q., and Racine, J. S. (2007), Nonparametric Econometrics: Theory and Practice, Princeton University Press, Princeton, NJ.

Lockwood, B. B., and Weinzierl, M. (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, 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., and Richard, S. F. (1981), A Rational Theory of the Size of Government, Journal of Political Economy 89, 914–927.

Milanovic, B. (2000), The Median-Voter Hypothesis, Income Inequality, and Income Redistribution: An Empirical Test with the Required Data, European Journal of Political Economy 16, 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: Balch, M., McFadden, D., and Wu, S. (Eds.), Essays on Economic Behavior under Uncertainty, North-Holland, Amsterdam, 243–258.

Moene, K. O., and Wallerstein, M. (2001), Inequality, Social Insurance, and Redistribution, American Political Science Review 95, 859–874.

Ordás Criado, C., Valente, S., and Stengos, T. (2011), Growth and Pollution Convergence: Theory and Evidence, Journal of Environmental Economics and Management 62, 199–214.

Perotti, R. (1996), Growth, Income Distribution, and Democracy: What the Data Say, Journal of Economic Growth 1, 149–187.

Persson, T., and Tabellini, G. (2002), Political Economics and Public Finance, in: Auerbach, A. J., and Feldstein, M. (Eds.), Handbook of Public Economics Vol. 3, North-Holland, Amsterdam, 1549–1659.

Piketty, T., Saez, E., and Stantcheva, S. (2011), Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities, NBER Working Paper No. 17616.

Piketty, T., Saez, E., and Stantcheva, S. (2014), Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities, American Economic Journal: Economic Policy 6, 230–271.

Rodrik, D. (1998), Why Do More Open Economies Have Bigger Governments?, Journal of Political Economy 106, 997–1032.

Saez, E. (2001), Using Elasticities to Derive Optimal Income Tax Rates, Review of Economic Studies 68, 205–229.

Saez, E., Slemrod, J., and Giertz, S. H. (2012), The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review, Journal of Economic Literature 50, 3–50.

Saez, E., and Stantcheva, S. (2016), Generalized Social Marginal Welfare Weights for Optimal Tax Theory, American Economic Review 106, 24–45.

Scervini, F. (2012), Empirics of the Median Voter: Democracy, Redistribution and the Role of the Middle Class, Journal of Economic Inequality 10, 529–550.

Spadaro, A., Piccoli, L., and Mangiavacchi, L. (2015), Optimal Taxation, Social Preferences and the Four Worlds of Welfare Capitalism in Europe, Economica 82, 448–485.

Tuomala, M. (1984), Optimal Degree of Progressivity under Income Uncertainty, Scandinavian Journal of Economics 86, 184–193.

Tuomala, M. (1990), Optimal Income Tax and Redistribution, Clarendon Press, Oxford.

Tuomala, M. (2016), Optimal Redistributive Taxation, Oxford University Press, Oxford.

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, Chapman & Hall/CRC, Boca Raton, FL.

Wood, S. N. (2017), Generalized Additive Models: An Introduction with R, Second Edition, CRC Press, Boca Raton, FL.

World Inequality Database (2017), http://wid.world/ (data accessed: 2017-11-30).