Progressive Taxation and Redistribution

Pablo Beramendi[*]
Matthew Dimick[†]
Daniel Stegmueller[‡]

Abstract

In this paper we revisit the relationship between progressivity and redistribution. We develop a model that reconciles the intuition that more progressive taxes ought to be more redistributive with the widely shared understanding that there is a trade-off between the redistributive scope of the fiscal system and the progressivity of its tax structure. Building on a novel measurement approach, which uses tax-benefit simulation models to measure the "pure" policy effect of taxes and benefits on incomes, we show three things: (1) Contrary to common consensus, the progressivity of taxes matters at least as much, if not more, than that of benefits; (2) There is a strong positive association between the relative importance of proportional taxes in the tax function and the progressivity of its design; and (3) The political influence of the poor increases tax progressivity and therefore the redistributive incidence of the fiscal system.

[*]Duke University, pb45@duke.edu
[†]SUNY Buffalo Law School, mdimick@buffalo.edu
[‡]Duke University, mail@daniel-stegmueller.com

I. Introduction

Tax progressivity is a central yet relatively understudied feature of countries' tax-and-transfer systems. Comparative research is only just beginning to generate explanations for differences in progressivity across countries, both in terms of its political origins and its distributional consequences (Kato 2003; Ganghof 2006; Beramendi and Rueda 2007; Martin, Mehrotra, and Prasad 2009; Prasad and Deng 2009; Martin 2015).[4] Yet, despite its incipient nature, this emerging literature already harbors a fundamental tension between two views on the relationship between the progressivity of taxes and the redistributive impact of the fiscal system.

In his seminal contributions, Kakwani (1977a, b) established that the overall redistributive incidence is a function of three factors: the size (effort) of fiscal spending, the progressivity of its design (of both taxes and benefits), and the reordering that both sets of policies, cause, through the behavioral responses of market actors in the distribution of market income in the first place. Leaving aside the latter term for now, it follows from this simple setup that, holding the level of effort constant, more progressive tax and benefit structures are more redistributive (Lambert 2001). In other words, progressivity is a necessary condition for redistribution and there should be, empirically, a positive relationship between the two.[5]

In apparent contrast with this view, a small literature in political science contends that (1) redistribution really takes place on the benefit side, not the tax side (e.g., Kenworthy 2008); and (2) that countries that redistribute *more* have *less* progressive tax structures is quickly gaining widespread acceptance (see, e.g., Martin 2015). Part of the tension emerges from an ambiguity about the level of analysis on the basis of which different authors define progressivity. Some contributions focus on the relative balance *between* tax tools;[6] Others refer to progressivity *within* specific tax instruments (income, consumption, etc.).[7]

Here we focus on the progressivity of the income tax, not its relative position within the tax structure, and its relationship with redistribution and the gap between pre- and post tax inequality. In this paper, we make three contributions. The first one is theoretical: we present a political economy model to understand variation in progressivity across rich democracies as a function of income biased representation. The model provides two major insights. First, as inequality increases, *both* tax progressivity and redistribution will decrease. The main

[4] An exception to this recent trend is the pioneering work of Alt (1983), Steinmo (1989), and Steinmo (1996).

[5] For a vivid illustration of the effect of taxes on distributive outcomes, see Newman and O'Brian (2011).

[6] It seems widely accepted that countries that redistribute more often use taxes that, all else equal, are less progressive. The main example here is the value-added tax (VAT), a type of consumption tax, which is used widely in (more redistributive) Europe, but not in the United States (see, e.g., Beramendi and Rueda 2007).

[7] For example, exemptions on certain goods (e.g., food) can make consumption taxes more progressive. Moreover, the adoption of a consumption tax need not make the aggregate tax structure less progressive if, for example, the income tax schedule is made more progressive to compensate for the regressivity of the consumption tax.

intuition for this result is the increasing weight of the rich in the political process. Rich voters prefer regressive taxation, and as their influence increases, tax policy becomes more reflective of their preferences. Political institutions, particularly PR, moderate this influence. In making the alliance between the poor and the middle classes more durable[8] or the relatively higher likelihood of stable pro-redistributive coalitions (Iversen and Soskice 2006), PR, as opposed to majoritarian systems, tends to generate income designs that maximize overall redistribution at the expense of progressivity. We uncover two reasons for this: (1) the marginal cost of taxation is smaller for poorer individuals, who would rather have more proportional taxes that maximize revenue and increase transfers; and (2) the middle class would rather tax only the rich and not themselves, reducing the overall progressivity of the system. It is worth clarifying this contribution in relation to existing theoretical work. There is a large literature that tries to determine whether progressive tax structures will emerge under general voting equilibrium conditions (see Cukierman and Meltzer 1988; Snyder and Kramer 1988; Marhuenda and Ortuño-Ortín 1995; Roemer 1999; De Donder and Hindriks 2003, 2004; Carbonell-Nicolau and Klor 2003; Carbonell-Nicolau and Ok 2007; De Freitas 2009; Roemer 2011). Yet, this literature neither attempts to explain variation in progressivity across countries nor develops any systematic predictions about the relationship between progressivity and redistribution. In contrast, we develop an explicit model of inequality, redistribution preferences, and biased political representation that addresses precisely these questions.

Our second contribution is methodological. As Beramendi (2001) notes, most policies have second-order, behavioral effects, tax policy not least of all. Taxation does not change the distribution of income only by taking from some and giving to others. It also influences decisions such as how much to work, whether to work, how hard to bargain, when to pay taxes, and so forth. By doing so, taxation also influences the distribution of income before taxes are collected and redistributed. Measures of progressivity based on household income data conflate these two aspects.[9] We measure tax and benefit progressivity using a policy simulation approach. This approach predicts the amount of tax a person would pay (or benefit received) based on all of the complex *statutory* legal rules specifying who should pay how much in taxes (or receive in transfers). This measures the "pure" effect of tax policy by excluding the ex-post effect of taxation on market—that is *pre*-tax—income.

Finally, our third contribution is empirical. Our approach provides compelling evidence that, contrary to a widely held view, tax progressivity, as compared with social spending and benefit progressivity, is a substantively and statistically significant determinant of redistribu-

[8]The persistence of such an alliance occurs through a variety of mechanisms, including the ideological position of the median legislature (Austen-Smith 2000)

[9]Indeed, these behavioral effects can explain the divergence between our results and those of the received wisdom (see, e.g., Piketty, Saez, and Stantcheva 2014).

tion across rich democracies.[10] We show empirically that more progressive tax structures are indeed associated with more redistribution, not less. On average, a standard deviation increase in tax progressivity increases redistribution by about 0.4 standard deviations. We also show that benefit progressivity has no clearly significant relationship to redistribution. While our income biased representation model is consistent with the results that we produce, the explanation based on political institutions is more in line with the current consensus. An increase in inequality lowers both progressivity and redistribution. In contrast, a political system with proportional representation redistributes more, but has a less progressive tax structure than one with majoritarian representation. Our tentative conclusion is thus that the distribution of income itself is a better explanation of tax policies across developed countries. We stress that these are theoretical results arrived at deductively from our model of inequality and political representation. Our model produces several testable (i.e., observable) implications, which can be tested in future work.

The paper proceeds as follows. Section II develops an analytical framework to disentangle the relationship between progressivity and redistribution. The model results provide the basis for our empirical analyses (III). Subsequently, we detail how we measure tax progressivity and discuss our empirical strategy in section IV. In our results section (V) we demonstrate the relevance of tax progressivity (and the irrelevance of benefit progressivity) in determining the levels of redistribution achieved by countries and we subject our results to a variety of specification tests. We finish this section by providing an empirical illustration of further political implications from our model. Section VII concludes the paper.

II. Model

In this section we develop a model to understand the relationship between progressivity and redistribution. After presenting the model set-up, we pay attention to two steps in the process: preference formation among three groups of voters (the rich, the poor, and middle income earners) and preference aggregation into policy.

A. Setup: Actors, tax function and redistribution

Consider a measure one continuum of individuals characterized by their income w and their opportunity cost of working, θ. There are three kinds of income earners, the rich, middle class, and poor, $i \in I = \{R, M, P\}$ with $w_R > w_M > w_P$. Each group has density

[10] Kenworthy (2008) makes both claims when he writes that among a set of developed countries, "[N]one …achieves much inequality reduction via taxes. Instead, to the extent inequality is reduced, it is mainly transfers that do the work." Note that Kenworthy uses (descriptive) country comparisons, while our evidence is based on within-country changes (i.e., fixed effects specifications).

p_i with $\sum_{i \in I} p_i = 1$. We assume that $p_R, p_P < \frac{1}{2}$ so that the median income earner has w_M. The opportunity cost of working has distribution G with support on the positive real line, $\theta \in \Theta = \mathbf{R}_+$. Thus $G(\theta)$ is the fraction of the population with work opportunity cost less than or equal to θ.

Each group of voters pays a marginal tax rate on income, τ_i, and receives a lump sum transfer b. To keep the model as simple as possible, we assume that both M and P voters pay the same tax rate. Therefore, $\tau_R = \tau_2$ and $\tau_M = \tau_P = \tau_1$. This makes the problem bidimensional: $\tau = (\tau_1, \tau_2) \in T$, with $T : [0,1] \times [0,1]$. Formally, following standard definitions, the tax schedule will be (weakly) *progressive* whenever $T(w_i)/w_i$ is nondecreasing in income. Conversely, the tax schedule will be *regressive* whenever $T(w_i)/w_i$ is decreasing in income. In our case, a progressive tax schedule simplifies to the condition $\tau_2 - \tau_1 \geq 0$ (or $\tau_2 - \tau_1 < 0$ for a regressive tax schedule).

We define redistribution in the standard way, as the condition where the distribution of disposable (post-tax, post-transfer) income Lorenz dominates the distribution of market (pre-tax, pre-transfer) income.[11] We can write this condition as:

$$\frac{\sum_{j=1}^{k} p_j c_j}{\bar{c}} - \frac{\sum_{j=1}^{k} p_j y_j}{\bar{y}} \geq 0 \quad \text{for each} \quad k \in \{1, 2, \ldots, n\} \quad \text{and} \quad y_j > y_{j-1} \tag{1}$$

where y_i represents market income (with $\bar{y}$ being average market income) and c_i is disposable income (with $\bar{c}$ being average disposable income). These variables are further defined below. We require integer indexing to make clear that incomes are ordered in an increasing direction, and because we will have more than three income groups since some individuals choose to remain unemployed. Thus, in our case, $n = 4$.

Because much of our paper is concerned about the relationship between tax progression and redistribution, it is important to clarify that one can have a tax schedule that is more progressive, but less redistributive, than another. Although progressivity in either the tax or transfer schedule (or both) is necessary for redistribution to occur, it is not sufficient. For instance, in the case where each tax rate taxes all income fully, $\tau_1 = \tau_2 = 1$, and transfers are lump sum, the tax schedule is proportional (neither progressive nor regressive). Nevertheless, since income inequality is completely eliminated, a lot of redistribution takes place. Conversely, when the tax on M and P is zero, $\tau_1 = 0$, and the tax on R slightly positive, $\tau_2 > 0$, the resulting tax schedule will be more progressive than the previous, yet substantially less redistributive. Of course, it is also possible for a tax schedule to be both more progressive and more redistributive.

[11] Our empirical measure of redistribution is the difference in Gini coefficients of pre- and post-fisc income distributions. Since the Gini coefficient is simply measured by the area between the Lorenz curve and the line of equality, this definition of redistribution translates seamlessly to our empirical measure.

B. Labor supply

Facing a given tax structure, individuals in each group make the discrete choice to work or not. We thus model labor supply on the extensive margin. Not only is this choice more tractable, but there is growing evidence that the extensive margin (i.e., whether to work or not) is more important than the intensive margin (i.e., how much, or how many hours, to work) (see, e.g., Meghir and Phillips 2008). [12]

Individuals choose to work when doing so exceeds the opportunity cost, θ:

$$(1 - \tau_i)w_i + b \geq b + \theta \tag{2}$$

Setting this condition to equality allows us to identify the value of the work opportunity cost, $\bar{\theta}$, for the worker who is indifferent between working and not,

$$\bar{\theta} = (1 - \tau_i)w_i \tag{3}$$

and $G(\bar{\theta}, w_i)$ is then the proportion of type i individuals who choose to work. We also denote

$$\eta(w_i) \equiv \frac{\partial G}{\partial \tau}\frac{\tau}{G} \quad \text{and} \quad \eta_w \equiv \frac{\partial \eta}{\partial w} \leq 0 \tag{4}$$

as, respectively, the elasticity of participation with respect to the tax rate, and the rate of change of this elasticity with respect to income. Consistent with empirical evidence, we assume that $\eta(w_i)$ is *decreasing* in income: richer workers are less responsive to taxes on the extensive margin than are poorer workers.

C. Preferences over tax schedules

Given the above, we define $y_i = w_i G(\theta, w_i)$ and write average income as:

$$\bar{y} = \sum_{i \in I} p_i w_i G(\bar{\theta}, w_i) \tag{5}$$

On part of the government, we impose a balanced budget requirement, so that tax revenues equal tax expenditures:

$$\sum_{i \in I} p_i(\tau_i w_i G(\bar{\theta}, w_i) - b) = 0. \tag{6}$$

[12] Although traditionally the public economics literature has focused on labor supply on the intensive margin, there has been increasing attention given to the extensive margin for these reasons (see, e.g., Saez 2002; Choné and Laroque 2005, 2011; Lehmann, Parmentier, and Van Der Linden 2011).

Solving for b, we obtain:

$$b = \sum_{i \in I} p_i(\tau_i w_i G(\bar{\theta}, w_i) \tag{7}$$

We can now write each working individual's utility as follows:

$$V_i = (1 - \tau_i)w_i + b \tag{8}$$

From this we can characterize the preferences over tax schedules of individuals from each group, which we do in the following lemma:

Lemma 1. *Let $\hat{\tau}_i \in T$ be each individual's group-specific ideal tax schedule. Then $\hat{\tau}_R = (\hat{\tau}_1, 0)$, $\hat{\tau}_1 \geq 0$, $\hat{\tau}_M = (0, \hat{\tau}_2)$, $\hat{\tau}_2 \geq 0$, and $\hat{\tau}_P = (\hat{\tau}_1, \hat{\tau}_2)$, $\hat{\tau}_1, \hat{\tau}_2 \geq 0$. Specifically, the progressivity of an individual's ideal tax schedule is nonmonotonic (increasing and then decreasing) in income. In addition, the level of redistribution implied by an individual's ideal tax schedule is decreasing in income.*

Proof. See Appendix A.1.

Critically, this lemma shows that the progressivity of each group's preferred tax schedule is nonmonotonic in income. In particular, the rich prefer a regressive tax schedule, while the middle class and the poor prefer progressive tax schedules. However, the tax schedule preferred by the middle class is more progressive but redistributes less than the tax schedule preferred by the poor. It therefore seems *a priori* plausible that tax schedules could emerge that redistribute more, but be less progressive, as the current consensus suggests. We would expect this to be case in systems where political representation enhances the poor's voice.

The reasoning for these preferences is as follows. Because the poor are poorer than the middle class, they can gain by increasing the tax rate that applies only to themselves and the middle class (τ_1). In addition, precisely because they are poor, the marginal cost of this tax is smaller to them than to the middle class. In contrast, the middle class cannot gain from a tax that redistributes between them and the poor, and so prefer to set this rate to zero. Hence, the poor prefer less progressive, but more redistributive taxes than the middle-class. Finally, the rich do not bear any of the cost of the tax rate that applies only to middle-class and poor voters, see they prefer to set this rate at a positive level. In addition, they can only lose by increasing any tax on themselves, and so prefer to set this rate to zero. The rich therefore prefer regressive tax schedules.

D. Preference Aggregation: Political environment

To understand how these preferences translate into redistribution policy we need to describe a model of the political process. We adopt a standard political environment along the lines of Coughlin and Nitzan (1981). We later change elements to the political environment to introduce differences in political representation.

The probabilistic framework proposed by Coughlin and Nitzan (1981) considers the case with two office-motivated parties, A and B. In this setting, the probability, π_i^A that a voter from group i will vote for a party A is:

$$\pi_i^A = \frac{V_i(\tau_A)}{V_i(\tau_A) + V_i(\tau_B)} \tag{9}$$

Thus, a voter from group i will be more likely to vote for party A if party A's platform gives her greater economic utility than party B. All voters vote, so $\pi_i^A + \pi_i^B = 1$. Since parties care only about winning, they choose their platforms, τ_A and τ_B, to maximize:

$$\pi^A = \sum_{i \in I} \pi_i^A. \tag{10}$$

As Coughlin and Nitzan (1981) demonstrate, this objective is equivalent to maximizing a Nash social welfare function:

$$N(\tau_A) = \sum_{i \in I} p_i \ln V_i(\tau_A) \tag{11}$$

Note that we have implicitly assumed for simplicity that the unemployed have no weight in the political process. This may well be the case because the unemployed have little influence in politics and turnout may be lower than in the general population.[13]

The final piece in the development of our analytical framework concerns the equilibrium relationship between progressivity and inequality given a probabilistic voting framework. In the Coughlin and Nitzan (1981) environment, it is straightforward to characterize the equilibrium. First, as is well known, even though the model is "probabilistic," there is convergence in the platforms proposed by parties because they have no substantive policy preferences and are strictly office seekers. More importantly for our purposes, under any possible equilibrium *only* progressive tax schedules exist. Furthermore, a necessary and sufficient condition for progressive schedules is the existence of inequality. The following result captures the core intuition:

Proposition 1. ***(Progressive Taxation under Democracy)*** *A symmetric equilibrium tax schedule, $\tau^* = \tau_A^* = \tau_B^*$, exists and is unique. Progressive tax schedules emerge in equilibrium if and only if there is income inequality.*

Proof. See Appendix A.2.

[13]See e.g., Schlozman and Verba (1979); Rosenstone and Hansen (1993) for some classic illustrations.

Moreover, in this setting, progressivity is also enhanced if the labor supply on the extensive margin is increasing in income and the elasticity of the participation rate is strictly decreasing in income. The intuition is simple: because richer individuals participate at higher rates and are less responsive to taxation, they are more attractive targets for taxation.[14]

III. Analysis and Hypotheses

Making use of this analytical framework, we turn now to analyze the relationship between progressivity and redistribution under democracy. We proceed in two different context, one economic, one political. We fist study the response in terms of progressivity and redistribution to a change in inequality. Second, we analyze how political representation, and in particular the institutional facilitation of stronger voice by the poor, conditions the relationship.

A. Context I: Progressivity and Redistribution given a change in inequality

Proposition 2. ***(Progressivity and Redistribution)*** *Consider a mean-preserving spread in the income distribution from X to Y such that $p_{X,P} < p_{Y,P}$, $p_{X,M} > p_{Y,M}$, and $p_{X,R} < p_{Y,R}$. Then progressivity and redistribution are higher under (more equal) distribution X than under distribution Y.*

Proof. See Appendix A.3.

The reasoning for this proposition is as follows: First, the tax rate on rich individuals, τ_2 falls because the political weight of middle-class voters, who favor a high tax rate on rich voters, diminishes with an increased dispersion of the income distribution. In contrast, the effect of the tax rate on lower income groups, τ_1, increases. Both rich and poor voters prefer a higher τ_1 than middle-class voters, which tends to increase τ_1 relative to the lower inequality state. Thus, as inequality increases, the tax schedule becomes clearly less progressive.

With an increase in τ_1 redistribution decreases for several reasons. First, an increase in τ_1 increases the tax burden on poor and middle class voters, which reduces redistribution. Second, as the proportion of middle-class incomes decline and the proportion of poor incomes increases, revenue raised by this tax falls. Finally, the fall in revenue from τ_2 is unlikely to be replaced by the increased revenue from τ_1, both because the elasticity of participation from poor and middle class workers is higher than for rich workers, and simply because the rich have higher incomes than either.

The central implication from this simple exercise is clear: *ceteris paribus*, as progressivity decreases, so does redistribution, a change that is in turn associated with an increase in the

[14] This result demonstrates most clearly the value to modeling labor supply on the extensive margin.

level of post-tax post-transfer inequality. This result calls into question the notion, widely shared and often discussed, that taxes have no redistributive effect. Leaving aside the fact that in fiscal incidence terms such a statement is a bit of an oxymoron, this result provides important correction to a dominant view in the political economy of redistribution. Our model suggests that for any given level of effort and the progressivity of benefits, a change in the progressivity of taxes does have a significant and positive effect on redistribution.

This result also raises a question for another stream of research, namely the relationship between progressivity and the overall scope of redistribution in society. Recall that, as discussed above, several previous contributions point to the existence of a trade-off between the two, especially in those polities where the system of representation has facilitated the creation and sustainability of pro-redistributive cross-class coalitions (Kato 2003; Beramendi and Rueda 2007). Does the fact that progressivity exerts a positive marginal effect on redistribution imply that the relationship between redistribution and the overall tax structure is, contrary to what was previously held, also linear and positive? To address this question we need to evaluate the relationship between progressivity and redistribution under institutional conditions that facilitate the political voice of low income citizens, and therefore an increase in the overall *level* of redistribution.

B. Context II: Political Representation

To account for the presence of political institutions that curb income biased representation, we depart from the Coughlin and Nitzan (1981) framework in two ways: (1) voting is no longer probabilistic and (2) parties have substantive policy preferences that are filtered by different electoral systems. In this sense, we analyze the impact of representative institutions from a setting similar to Austen-Smith (2000) and Iversen and Soskice (2006).

Both these models analyze politics in majoritarian institutions as two-party competition (as in Duverger's Law) with each party vying for the decision votes of middle-class voters. In this case, policy will always reflect middle class interests.[15] In contrast, under proportional representation, there are three parties, each representing a different economic constituency: P, M, R. Furthermore, policy making under proportional representation features coalition government, inasmuch as the winners do not take all. Therefore, following an election stage, we adopt a legislative bargaining stage wherein parties bargain with other parties in coalitions. As is standard in this literature, we assume without loss of generality that party

[15] We simplify their approach somewhat by assuming that there is never a danger of either party deviating to a rich or poor person's preferred platform. Dropping this assumption has uncertain effects. It certainly does not change the result that PR systems redistribute more than majoritarian. The effects of dropping the assumption on progressivity are less clear. Crucial in this regard is their non-regressivity assumption (p. 167), an assumption we do not make in this paper.

M acts as the formateur. Using this political environment, our analysis yields the following proposition.

Proposition 3. ***(Political Representation)***

Suppose that under majoritarian representation, tax policy is coincident with middle-class preferences: $\tau^* = \hat{\tau}_M$. Then redistribution is higher and progressivity is lower under proportional representation than under majoritarian representation.

Proof. See Appendix A.4.

As we have previously demonstrated in Lemma 1, under simple assumptions, individuals' ideal tax schedules are nonmonotonic (increasing and then decreasing) in income. In contrast, the amount of redistribution that occurs under these ideal tax schedules is decreasing in income. From this it is a short step to the conclusion that, all else equal, majoritarianism increases progressivity but reduces redistribution relative to proportional representation. The translation of these preferences into policy differs by system of representation.

Parties in majoritarian systems give greater weight to middle-class voters, while proportional systems give greater weight to poor voters. This effect may reflect that the key decision maker is the median legislator left off median (Austen-Smith 2000) or the formation of coalitions where parties representing the poor and the middle classes must compromise with each other (Iversen and Soskice 2006). In either case, given that middle-class voters prefer more progressivity and less redistribution than poor voters, tax schedules will be more progressive but will redistribute less under majoritarian representation. In contrast, since proportional representation enhances the political representation of the poor's preferences, such systems will exhibit less progressive tax schedules but more redistribution.

This results suggests that as the political influence of the poor increases, that is as income-biased representation declines as a result of political representation, increases in the overall size of redistribution (as supported by low income voters) will require partial sacrifices in terms of the overall progressivity of the income tax design. The higher levels of redistribution are viable because the 'sacrifice' in terms of design allows for a much larger revenue base, along the lines of the non monotonic preference scheme above. In addition, such a design allows the tax burden on the middle classes to be reduced while at the same time facilitating the provision of encompassing systems from which they also benefit.

As we elaborate below, the sacrifice is best captured through the relationship between the two components of tax function: the progressivity of the tax design and the level of flat rate tax income. A central implication of our analysis is that as the former increases the latter must decrease for the mapping of pre-tax to post-tax income, that is for the overall level of redistribution, to remain the same. In other words, there is a ceiling to the amount of redistribution that can be achieved via tax progressivity. Paradoxically, if political actors want to maximize redistribution, as it plausibly happens in systems where the poor exercise significant political influence, the optimal combination is not one that keeps progressivity at its maximum. Rather, it is one in which the flat rate is marginally higher and is kept at intermediate levels. It follows from this logic that in systems where low income voters are influential (such as PR) the ratio between the two components of the tax function (progressive to flat tax rate) will be smaller.

C. Summary of Empirical Implications and Strategy

The theoretical analysis above yields empirical implications about the relationship between progressivity and redistribution:

1. *On the marginal effect of tax progressivity on redistribution (Proposition 2):* For any given level of effort and the progressivity of benefits, a change in the progressivity of taxes does have a significant and positive effect on redistribution (H1)
2. *On the relationship between redistribution and tax designs (Proposition 3):*
    a) There is a negative association between the progressivity of the tax system and the level of (flat rate) taxation[16]
    b) *Corollary*: as the political influence of the poor increases (PR vs SMD), the ratio of progressivity to flat-rate level (proportional) taxation decreases.

IV. Data and Empirical Strategy

We turn now to the empirical assessment of these implications. We begin by describing a novel measurement approach specifically designed to exclude changes in market distributions induced by the very fiscal tools under analysis. On the basis of these analyses we recover our own estimates of λ and τ for 21 OECD countries. The limiting factor in the selection of our sample is the availability of high-quality measures of redistribution (described below). Thus our analysis sample consists of 203 country-years. The countries included in our analysis are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, United Kingdom, and the United States. Table B.1 in the appendix shows which years we observe for each country. The median number of years available by country is 11 (17 countries provide at least 8 years of observations). Thus, we have enough information to employ within-country designs in our statistical analyses.

[16]To avoid confusion, note that we define the tax function below as the mapping from pre (w) to post tax (x) earnings: $x = \lambda w^{(1-\tau)}$. τ captures progressivity, whereas $1 - \lambda$ would capture the level of flat rate taxation

A. Measuring tax progressivity

We measure tax progressivity using a policy simulation approach. This choice is driven by our desire to capture the intended effect of tax policy set by governments given an existing distribution of market incomes. In other words, we want to measure the 'pure' structure of tax policy *net* of possible changes in the distribution of market incomes. Excluding changes in market distributions from the calculation of policy measures is important. As Beramendi (2001) notes most, if not all, policies create second order effects. Policies, when implemented (or announced), do not simply shift the equilibrium distribution of post-government economic outcomes. They also alter the distribution of market outcomes, since individuals react to policy changes. Thus, policies "generate behavioral responses from market actors, who adjust their economic decisions to the nature and changes of policy interventions" (Beramendi 2001: 5).[17] Respondents' market and disposable income derived from household surveys are a mixture of "first order" effects of welfare policies *and* "second order" effects of individuals' behavioral responses (Boadway and Keen 2000: 760). Using policy simulation allows us to separate these two effects. We use the OECD tax-benefit simulation model, which contains detailed rules of both tax and benefit policies covering OECD countries between 2001 and 2015 (OECD 2007). It allows us to calculate the income consequences of each country's tax and benefit structure for households at different point in the income distribution. We then summarize these income effects using progressivity measures, described below.

Simulation details In order to capture a wide range of realistic individual circumstances, our simulation includes four family types: single individuals, single parents with children, and married couples with and without children. In families there is one active working-age individual. Our policy effect simulation computes the amount of taxes (and benefits) across the earnings spectrum for each family type. To normalize difference in average earnings between countries, all earnings are expressed in percentiles relative to those of the average production worker (APW) in a country. They range from 50 to 200 percent of the APW wage.[18] Thus, our results capture the effects of taxes on the incomes of working-age individuals and their families. The taxes included in the simulation are Personal Income taxes, as well as Employee and Employer paid Social Security Contributions. This calculation

[17] This issue is widely acknowledged in macro-economics, where simulating the effect of policy interventions requires not only a knowledge of existing and future policy rules, but also the knowledge of individual preference parameters to predict individuals' behavior responses at both extensive and intensive margins (for extended discussions of issues surrounding policy effect analysis (including shortcomings of our current static approach) see, e.g., Atkinson, King, and Sutherland 1983; Feldstein 1983; Figari, Paulus, and Sutherland 2015).

[18] This is a synthetic (not empirical) income distribution anchored at the wage of the average production worker. It thus holds constant factors such as income-specific incidence of unemployment spells during business cycles, or country-specific gender disparities in occupations.

yields a data set of over 250,000 records of pre and post tax incomes (each country × family type × APW income level).

Tax progressivity With these data on the income consequences of taxes at hand, we can calculate our measure of tax progressivity. We use a tax function approach to approximate each country's tax system, following a long tradition in public finance, where this class of tax-transfer models has been introduced by Feldstein (1969), and later extended to dynamic models by Persson (1983) and Benabou (2002).

In this approach, tax revenue is a function of individual income w_i, since income is the only factor on which taxes can be conditioned on. The retention function of an individual's income is given by:[19]

$$T(w_i) = w_i - \lambda w_i^{1-\tau}. \tag{12}$$

It implies a non-linear mapping between after tax earnings x and pre-tax earnings w:

$$x = \lambda w^{1-\tau}, \tag{13}$$

where $1-\tau$ is the elasticity of before-tax to after-tax income. Thus, τ is a straightforward measure of the progressivity of the tax system. Generally, a tax system can be defined as progressive if the ratio of the marginal tax rate to the average tax rate is greater than one for a given level of income (e.g., Slitor 1948). In terms of our tax function this entails:

$$\frac{T'(w_i)}{T(w_i)/w_i} = \frac{1-\lambda(1-\tau)w_i^{-\tau}}{1-\lambda w_i^{-\tau}} \tag{14}$$

This expression is larger than 1 when $\tau > 0$ yielding a progressive tax structure. If instead $\tau < 0$ the ratio of marginal to average tax is less than one at a given level of income yielding a regressive tax structure. The parameter λ determines the average level of taxation in a country, since it 'shifts' the tax function for a for a given level of τ. If $\tau = 0$ the ratio in (14) is exactly 1 and a country's level of (flat-rate) taxation is then captured by $1-\lambda$.

We fit the tax function in (13) to the pre-post income data from our tax simulations, and obtain estimates for λ and τ for each country and year between 2001 and 2015.[20] Table B.2 in the appendix provides an overview of λ and τ estimates.

B. Definition of other variables

Redistribution. We use data from the OECD income distribution database, which compiles income information from administrative sources and household panel surveys in advanced

[19]The marginal tax rate on an individual's income is thus given by $T'(y) = 1 - \lambda(1-\tau)y^{-\tau}$

[20]The model approximates observed income patters quite well with R^2 values usually above 0.99.

industrialized countries (e.g., Förster and Pearson 2002).[21] The data sources used are of high quality, and estimates derived from it are comparable to those from the Luxembourg Income Study (OECD 2012). The advantage of using this data source is that it provides us with wider data coverage in terms of available country-years. For each country and year we have information on the Gini index of inequality of (square-root scale) equivalized household incomes, (i) at market incomes (before taxes and transfers), and (ii) at disposable incomes (after taxes and transfers). In line with existing studies (e.g., Kenworthy and Pontusson 2005: 455), we operationalize the extent of redistribution as the difference between (i) and (ii).[22]

Benefit progressivity Our measure of benefit progressivity also relies on data from our simulation model. As in our policy simulation for tax structures, we calculate benefits for four different types of households at varying points in the income distribution (ranging from 50 to 200 percent of the average production worker's income). When calculating the income effect of benefits, we include Unemployment, Social Assistance, Housing, Family, and in-work benefits. Based on this detailed data set we calculate the Kakwani progressivity measure of the distribution of benefits (Kakwani 1977b; Kakwani and Podder 1976). See Beramendi and Rehm (2016) for more discussion of this measure.

Social spending and proportion of elderly Two key variables we adjust for in our analyses are existing levels of spending on social programs (e.g., International Labor Organization 2008: 130) and a country's (changing) share of the elderly population. We measure social spending as total public expenditure in percent of gross-domestic product, available in the OECD's SOCX database and the elderly population as the share of individuals aged 65 and older (also obtained from the OECD).

C. Statistical specifications

We estimate several empirical specifications to study the relationship between tax progressivity and redistribution after accounting for levels of social spending and progressivity of benefits. We rely on two-way fixed effects models, which rely on within-country changes in all variables while adjusting for common time shocks. More extended specifications explicitly account for auto-regressive structures, pre-determined regressors, non-unit-constant shocks and slope heterogeneity. Before describing our main model specifications, note that our key measures are sufficiently distinct to jointly include them in a model analyzing redistribution.

[21] The data source for many countries is the EU Survey of Income and Living Conditions, which is one of the primary sources of European Union policy making. Other countries rely on high quality national household panels, their census, or register data.

[22] An argument against this "absolute" definition of redistribution is that it does not take into account initial levels of inequality (as "relative" measures, which express the difference in percentage of initial levels, do). But note that in our empirical specifications we estimate models that rely only on *changes* in redistribution, as well as a model which explicitly includes previous levels of redistribution.

The average correlation between benefit progressivity and tax progressivity is 0.36. This correlation decreased from 0.43 in the 2000s to 0.27 in the 2010s. The correlation of benefit and tax progressivity with spending is -0.05 and 0.03, respectively. All variables also show significant within-country changes over time.

Our first empirical specification for redistribution in country i in year t, y_{it}, is given by

$$y_{it} = \alpha\tau_{it} + x_{it}'\beta + \phi_i + \zeta_t + \epsilon_{it}, \quad i = 1, \ldots, N, \; t = 1, \ldots, T_i. \tag{15}$$

Here, redistribution is a function of tax progressivity τ_{it} with associated coefficient α and vector of controls (including benefit progressivity, social spending, and share of elderly population). We include country fixed effects ϕ_i, which capture (time-constant) unobservables on the country level as well as time fixed effects ζ_t (Baltagi 2013: 39f.). While this setup is the 'default' in many empirical analysis of panel data, it imposes the restriction that all units are affected by time-specific shocks in the same way, a point to which we return to below. Our second empirical specification accounts explicitly for the fact that residuals for the same country are likely correlated over time. We estimate a fixed effects specification where the residuals follow a stationary AR(1) process (taking into account the unbalanced nature of our dataset; cf. Baltagi and Wu 1999):

$$y_{it} = \alpha\tau_{it} + x_{it}'\beta + \phi_i + \epsilon_{it}, \quad \epsilon_{it} = \rho\epsilon_{it-1} + v_{it} \text{ with } |\rho| < 1 \tag{16}$$

In our third, more involved, specification we estimate a dynamic panel specification by including lagged values of redistribution, $y_{i,t-1}$, in our fixed effects model:

$$y_{it} = \rho y_{i,t-1} + \alpha\tau_{it} + x_{it}'\beta + \phi_i + \zeta_t + \epsilon_{it} \tag{17}$$

Simply introducing a lagged dependent variable into models with country fixed effects introduces bias (the lagged dependent variable violates strict exogeneity) particularly in a small-T context (Nickell 1981). We estimate the model using a difference GMM (generalized method of moments) estimator suggested by Arellano and Bond (1991). The estimator operates on a first differenced system of equations, where instruments are generated (i) for the lagged dependent variable from its own lags in each time period, (ii) for the first difference equation from differences of exogenous covariates(cf. Arellano and Bond 1991;

Holtz, Newey, and Rosen 1988).[23] This specification removes two periods and reduces our sample size to 141 country-years.

V. Results

A. Progressivity and redistribution

Table I shows estimates from six specifications. We begin with our basic two-way fixed effects model and sequentially enter levels of social spending, benefit and tax progressivity. For ease of interpretation we standardize all inputs to have mean zero and unit standard deviation. Our results in the first column of Table I show that, not surprisingly, higher levels of spending achieve more redistribution. Of more interest is the role of the progressivity of the tax and transfer system. Interestingly, column (2) shows that changes in the progressivity of benefits have no clear relation to increases or decreases in redistribution. In contrast, it is tax progressivity has a clear, statistically significant, impact. A standard deviation increase in τ increases redistribution by 0.4 (± 0.1) standard deviations. This finding is also illustrated in Figure I, where we plot expected values of redistribution at varying (unstandardized) levels of benefit and tax progressivity. A change from $\tau = 0.15$ (just below the median) to the 75% percentile ($\tau = 0.20$) increases redistribution by 1.7 points, or 15 percent. The corresponding change for benefit progressivity is negligible.

In columns (4) of Table I we allow for serial correlation of residuals (while taking into account unevenly spaced observations in each country) as specified in equation (16). The estimated correlation ($\rho = 0.6$) signifies that residuals are indeed strongly serially correlated. Accounting for this correlation also reduces the regressor fixed-effects dependence in the model.[24] Under this specification we find the coefficient for spending significantly reduced, while there is little change for the estimated impact of changes in benefit progressivity. The effect of tax progressivity is reduced by a third, but still of substantive magnitude and clearly statistically different from zero. A standard deviation increase in progressivity is associated with a 0.28 (± 0.07) standard deviation increase in inequality reduction.

[23] We did not make this model our preferred specification since it depends much more on specific modeling choices (such as the depth of lags to include) impacting the validity of the created instruments. We have limited the lag depth to 8 in our analysis, but ensured that unlimited depth does not change our results. To inspect the appropriateness of the model, we conduct two tests. One key issue is the assumption that the time-dependence follows a first-order Markov process (so that taking second-order lags creates instruments). This assumption can be tested by inspecting the auto-correlation of residuals after the model is estimated (ensuring that the first differenced residuals do not exhibit second-order autocorrelation). We also use the overidentifying restrictions to test the joint validity of the moment conditions of the GMM estimator using the Sargan-Hansen test.

[24] The correlation between regressors and country fixed effects in a fixed effects specification is -0.293, which is reduced to -0.059 in specification (4).

Table I
Redistribution as function of spending and progressivity.

	(1)	(2)	(3)	(4)	(5)	(6)
Spending levels	0.844	0.842	0.989	0.501	0.476	0.343
	(0.119)	(0.120)	(0.117)	(0.070)	(0.089)	(0.102)
Benefit progressivity		−0.036	−0.092	−0.071	−0.131	−0.054
		(0.231)	(0.196)	(0.059)	(0.081)	(0.093)
Tax progressivity			0.439	0.284	0.243	0.229
			(0.098)	(0.072)	(0.091)	(0.089)
ρ				0.607	0.534‡	0.583‡
Two-way fixed effects	✓	✓	✓	✓	✓	✓
Δ economic vars.						✓
Estimator	FE	FE	FE	FE AR(1)	GMM	GMM
R-squared†	0.31	0.32	0.44	0.53	0.42	0.54
N	203	203	203	182	141	141

Note: Unbalanced panel of 21 OECD countries, 2001–2015. All inputs normalized to mean zero and unit standard deviation. Cluster-robust standard errors.

Specifications: (1)-(3) Two-way fixed effects models (country and year). Average T=10.7. (4) AR(1) model with fixed effects (Baltagi and Wu 1999). (5) LDV model with fixed effects (Arellano and Bond 1991; Arellano 2003), GMM IV estimates; estimated on differenced system, using lagged LDV and differenced covariates as instruments: AR test of residuals $p = 0.652$. Sargan overidentifying restrictions test: $p = 0.162$ Specification (6) is (5) with added economic variables (first differences in inflation, real GDP growth, and unemployment rate). AR test $p = 0.157$, Sargan test $p = 0.367$. All models include the share of the 65+ population.

† Refers to "within-panel" R-squared (calculated using doubly demeaned data)

‡ Coefficient on lagged dependent variable.

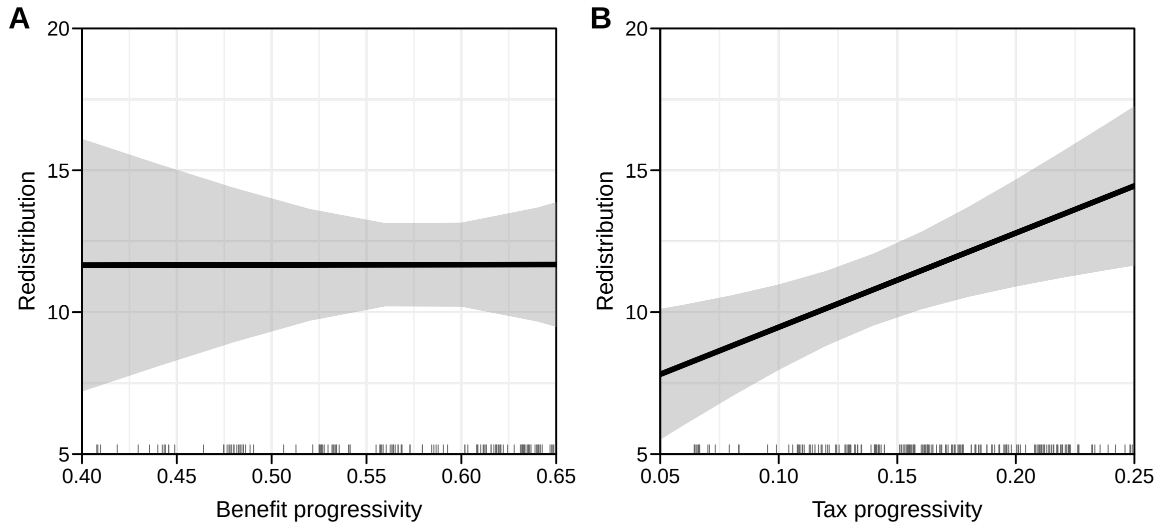

Figure I
The structure of taxes and transfers and redistribution

Expected value (with 95% confidence intervals) of redistribution at levels of tax progressivity and benefit progressivity. Based on two-way fixed effects model fitted to panel of 21 OECD countries, 2001-2015.

We introduce a lagged dependent variable in column (5) following the specification in equation (17). As expected, we find strong persistence of patterns of inequality reduction. The estimate for ρ, the coefficient for y_{it-1}, is sizable and statistically different from zero. Thus, redistribution in year t is in large parts determined by the amount of redistribution carried out in year $t-1$. In this setting, what is the effect of a change in the progressivity of the tax and transfer system? Even in this much more involved specification we find clear evidence for the substantive and statistical relevance of the progressivity of a country's tax structure. The contemporary effect of a unit-change in tax progressivity on redistribution is 0.24 standard deviations, while the long run effect (taking into account both the contemporary change and its feedback via lagged redistribution) is 0.55 (s.e.= 0.22). Finally, this is also confirmed in specification (6) where we add variables representing economic conditions that might effect achieved redistribution in a mechanical way, namely changes in inflation, real GDP growth and unemployment.

B. Specification tests

Before we proceed to a discussion of the political significance of our findings, we subject our results to a number of specification tests. We start with a model that allows for heterogeneous common shocks and flexible cross-sectional dependence. As discussed above, the traditional two-way fixed effects model setup assumes that country and time effects enter the model additively. Relaxing this assumption can be achieved by specifying a model with interactive fixed effects (Bai 2009) implemented via a factor structure:

$$y_{it} = \alpha\tau_{it} + x_{it}'\beta + \xi_i'F_t + \epsilon_{it}. \tag{18}$$

Here, F_t is a vector of r common factors, which represents the structure of unobservables shared by cross-sectional units in a given year, and ξ_i is a $r \times 1$ vector of corresponding factor loadings. ϵ_{it} are idiosyncratic errors. To see how this generates heterogeneity, think of F_t as a vector of common shocks (e.g, the financial crisis). Then ξ_i represents the heterogeneous impact of this shock on country i. We treat both ξ_i and F_t as fixed-effects parameters to be estimated using principal component methods (cf. Bai 2009: 1236f.).[25] The model makes no assumption about the mean of F_t or the structure of its over time dependency (see Pesaran 2006; Bai 2009 for a discussion of its asymptotic properties). Furthermore, we allow spending, benefit and tax progressivity to evolve endogenously:

$$x_{it} = \mu_i + \theta_t + \sum_{k=1}^{r} a_k\xi_{ik} + \sum_{k=1}^{r} b_k F_{kt} + \sum_{k=1}^{r} c_k\xi_{ik}F_{kt} + \pi_i'G_t + \eta_{it} \tag{19}$$

Here, a_k, b_k, and c_k are scalar constants and G_t is a different set of common factors (that do not enter in the outcome equation). Thus, covariates x_{it} can be correlated with ξ_i alone, or with F_t alone, or simultaneously with ξ_i and F_t. Specification (1) of Table II show that a model with interactive fixed effect estimated with one common factor produces estimates for the impact of τ that are very close to our main results. Extending the model to two common factors increases the impact of τ: a standard deviation increase in tax progressivity leads to an increase in redistribution of 0.44 (±0.12) standard deviations.

So far we have assumed that the coefficients on all covariates are constant over countries. If there is heterogeneity in some slopes, e.g., if the impact of spending varies over countries, does it change our understanding of the (average) effect of tax progressivity? In specification (2) we employ an estimation strategy that explicitly allows for heterogeneity in slopes (including country-specific time trends) employing the mean group estimator of (Pesaran and Smith 1995).[26] Our results in specification (2) show that this has little impact on the

[25] Note that we have to choose the number of factors a priori. One factor is enough to introduce cross-sectional dependence (Pesaran 2006: 972) and allow for interactive fixed effects. In our empirical implementation below we also estimate a model with two factors (which is still supported by the data).

[26] The central idea is to estimate N group-specific coefficients for all covariates and a time trend using OLS and then to combine the estimated coefficients across groups. Since this strategy relies on within-country estimates we only include countries with longer series (this amounts to excluding Australia, Japan, New Zealand, and Switzerland). Nonetheless, the pooled mean group estimator assumes large samples in

TABLE II
Specification tests. Estimate of tax progressivity (standard errors in parentheses).

	Tax progressivity	
(1) Interactive fixed effects estimator		
One common factor (r=1)	0.411	(0.084)
Two common factors (r=2)	0.438	(0.114)
(2) Heterogeneous panel estimator (MG)	0.613	(0.293)
(3) Bayesian TSCS model with two-way RE	0.420	(0.068)
(4) Balanced panel (multiple imputation)	0.397	(0.133)
(5) Percentile-t wild bootstrap imposing null	p=0.018	

Specifications: (1) Interactive fixed effects estimator with 1 and 2 common factors (Bai 2009). (2) Allows for heterogeneous regressor slopes and time trends via Pooled Mean Group estimator (Pesaran and Smith 1995). (3) Bayesian hierarchical model with country and year random effects, regressor RE dependence via Mundlak device. Based on 20,000 MCMC samples. (4) Balanced panel, N=313. Regression imputation using country-specific time trend (M=100). MI corrected standard robust errors. (5) Country and time cluster SEs. First entry uses analytic cluster-robust variance estimator. Second entry is test of significance using 1000 percentile-t wild bootstrap samples imposing the null (Cameron, Gelbach, and Miller 2008).

average effect of tax progressivity, but leads to increased standard errors (representing the increased heterogeneity in the model). That notwithstanding, our core result on the role of tax progressivity is confirmed.

The final three specifications are more technical in nature. In (3) we estimate our model in a Bayesian framework providing partial pooling estimates for both country and time random intercepts. See Shor et al. (2007) for the advantages of Bayesian inference with TSCS data. We account for regressor random effect dependence in both dimensions using the Mundlak specification (Rendon 2012).[27] We find little change in our substantive results. In specification (4) we create a balanced panel (with 313 country-years) by filling in values for redistribution under a MAR assumption. Missing years are primarily the results of lack of household panel data in the OECD Income Distribution database, making it more likely that the missingness process is not MNAR. Note that we have complete information on all years for our measures of progressivity, as well as for social spending and the share of the elderly. We create 100 imputed data sets and adjust our standard errors for the increase in

both dimensions (N and T) and thus our specification should be seen as providing suggestive only on heterogeneity only.

[27]We choose non-informative priors with mean zero and variance 100 for all regression type parameters. Variance priors are inverse gamma with shape and scale parameters set to 0.005.

imputation variance. Our results show little substantive change. Finally, specification (5) we test the statistical significance of the estimate for τ using a percentile-t wild bootstrap imposing the null hypothesis and taking into account clustering on both countries and years. See Cameron, Gelbach, and Miller (2008) for an extended discussion. The entry in Table II is the p-value from a two-sided test of the resulting empirical distribution.

VI. Further political implications of our model

Contrary to some previous misconceptions, our results provide solid evidence of a significant and substantively important impact of tax progressivity on redistribution, holding the levels of benefit progressivity and spending constant. This lends support to the first hypothesis (H1) derived from our model. We turn now to assess the second set of implications, those concerning the design of the overall tax structure as the political influence of the poor and the overall level of redistribution increase. The model yielded two implications: the first follows directly from the model; the second emerges as a corollary on the moderating role of political representation.

1. There is a negative association between the progressivity of the tax system and the level of (flat rate) taxation[28]

2. *C*orollary:As the political influence of the poor increases (PR vs SMD), the ratio of progressivity to flat-rate level (proportional) taxation decreases.

Figure II illustrates these implications empirically. Both panels lend empirical support to the theoretical contentions derived from the model. In panel (A) we plot our estimates of the flat rate tax parameter $1 - \lambda$ against τ. The solid line is a nonparametric estimate of their relationship (using a lowess smoother) and the associated 95% confidence interval. Panel A shows that there is a non-zero, negative relationship between progressivity and flat-rate tax levels. In quantitative terms (based on a linear model), as τ increases by one standard deviation $1 - \lambda$ decreases by 0.72 (±0.04) standard deviations. This is an important finding that sheds light on the ambiguity founds in previous contributions on the notion of "redistribution within one class" (Cusack and Beramendi 2006), that is on the idea that for sufficiently high levels of redistribution to be politically feasible workers and consumers must foot a significant share of the bill. Our model sheds light on the mechanism behind this pattern: as progressivity increases, the size of the tax base on which revenues are collec

[28] To avoid confusion, note that we define the tax function below as the mapping from pre (w) to post tax (x) earnings: $x = \lambda w^{(1-\tau)}$. τ captures progressivity, whereas $1 - \lambda$ would capture the level of flat rate taxation

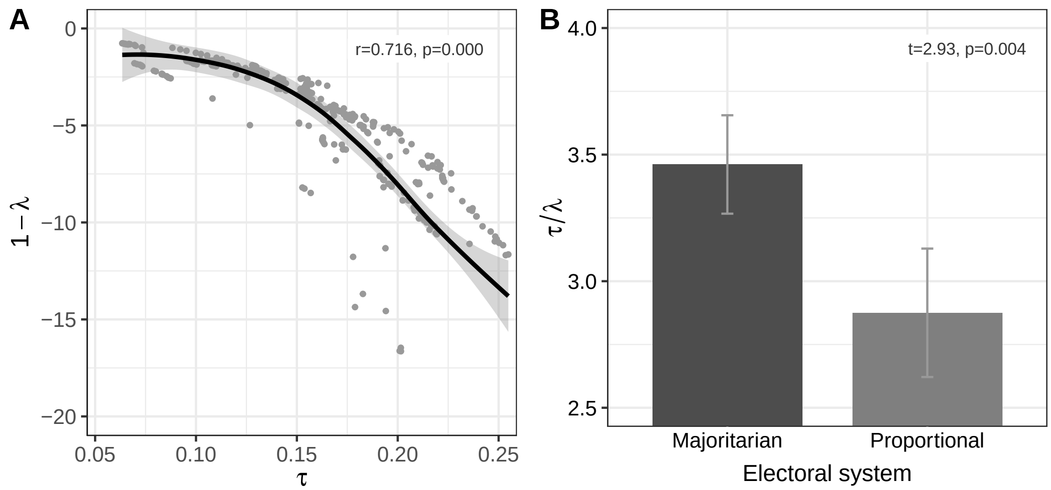

FIGURE II
Empirical illustration of model implications

Panel (A) plots the relationship between progressivity and flat-rate tax parameter. Superimposed is a lowess smoother with 95% confidence bands. Panel (B) shows the average value of τ/λ in majoritarian and proportional electoral systems. Error bars represent 95% confidence intervals.

ted declines, eventually yielding a relatively smaller pool of revenues to be shared. Our findings in panel A display a pattern that is consistent with this logic on the basis of new, and more precise, measurement strategy.

If sufficiently high redistribution comes at the expense of a partial sacrifice of progressivity, it should follow that in democracies where political institutions (PR) facilitate stronger political influence by the poor, the ratio between τ and $1-\lambda$ is lower. To address this corollary directly, panel (B) plots the average (and 95% confidence intervals) of the ratio of progressivity to λ in majoritarian versus PR electoral systems.[29] In line with our theoretical expectations, in countries with majoritarian electoral systems we find the ratio to be 0.80 ($\pm$0.22) points greater than in countries relying on proportional representation.

VII. Conclusion

What governs the relationship between progressive taxation and redistribution? A layman's view would suggest that both are one and the same. And yet, the dominant view

[29] We exclude mixed electoral systems in this calculation. However, note that including them in the reference group does not substantively alter our finding.

so far seems to contend that effective redistribution requires a sacrifice in terms of the progressivity of the design of tax structures. Accompanying these analyses, mostly based on cross-national macro-level data, the case is often made that redistribution is something driven by the "spending side" of the fiscal system, adding yet another layer of ambiguity to the problem.

This paper has developed a formal analysis of the political underpinnings of progressive taxation and a new, incidence-free, measurement strategy to revisit and clarify this relationship. Assuming that actors maximize their preferred level of redistribution, we show formally that: (1) there is a negative relationship between inequality and tax progressivity, and a positive relationship between tax progressivity and redistribution; (2) there is a non-monotonic relationship between income and preferences for progressive taxation: the poor actually may prefer partial sacrifices of progressivity to secure a larger pool of revenue from which to benefit. This larger pool of revenue would come through a larger flat-tax rate. This suggests two empirical expectations: (1) there is a positive marginal impact of tax progressivity on redistribution, contrary to the notion that all the action is on the same time; (2) at the same time, as the political influence of the poor increases, the relative balance between progressive vs flat-rate taxes changes in favor of the latter. Our empirical analyses lend empirical support to both contentions, thus reconciling the layman's view with the need to pay attention to both the marginal effect of progressive taxes and the size of the revenue pool when the role of income taxation in redistributive politics.

Our findings also suggest several paths for future research efforts. The most obvious one concerns the systematic exploration of the non-linear combination of tax progressivity and fiscal effort in a multivariate context. Our multivariate models have uncovered a robust linear effect of progressive taxation. At the same time, our politico-economic analysis of the institutional underpinnings of these relationships suggest that for a subset of countries (PR) the relationship changes. Exploring this contention in a rigorous, systematic way is the focus of ongoing research efforts.

References

Alt, J. E. 1983. "The Evolution of Tax Structures." *Public Choice* 41 (1): 181–222.

Arellano, M. 2003. *Panel Data Econometrics.* Oxford: Oxford University Press.

Arellano, M., and S. Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58 (2): 277–297.

Atkinson, A., M. King, and H. Sutherland. 1983. "The Analysis of Personal Taxation and Social Security." *National Institute Economic Review* 106 (1): 63–74.

Austen-Smith, D. 2000. "Redistributing income under proportional representation." *Journal of Political Economy* 108 (6): 1235–1269.

Bai, J. 2009. "Panel data models with interactive fixed effects." *Econometrica* 77 (4): 1229–1279.

Baltagi, B. H. 2013. *Econometric Analysis of Panel Data.* New York: Wiley.

Baltagi, B. H., and P. X. Wu. 1999. "Unequally spaced panel data regressions with AR (1) disturbances." *Econometric Theory* 15 (6): 814–823.

Benabou, R. 2002. "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70 (2): 481–517.

Beramendi, P. 2001. "The Politics Of Income Inequality In The Oecd: The Role Of Second Order Effects." Luxembourg Income Study Working Paper No. 284.

Beramendi, P., and D. Rueda. 2007. "Social democracy constrained: indirect taxation in industrialized democracies." *British Journal of Political Science* 37 (04): 619–641.

Beramendi, P., and P. Rehm. 2016. "Who Gives, Who Gains? Progressivity and Preferences." *Comparative Political Studies* 49: 529–563.

Binmore, K., A. Rubinstein, and A. Wolinsky. 1986. "The Nash bargaining solution in economic modelling." *The RAND Journal of Economics*: 176–188.

Boadway, R., and M. Keen. 2000. "Redistribution." In *Handbook of Income Distribution*, ed. A. B. Atkinson and F. Bourguignon. Vol. 1. Elsevier.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics* 90 (3): 414–427.

Carbonell-Nicolau, O., and E. A. Ok. 2007. "Voting over income taxation." *Journal of Economic Theory* 134 (1): 249–286.

Carbonell-Nicolau, O., and E. F. Klor. 2003. "Representative democracy and marginal rate progressive income taxation." *Journal of Public Economics* 87 (9): 2339–2366.

Choné, P., and G. Laroque. 2005. "Optimal incentives for labor force participation." *Journal of Public Economics* 89 (2): 395–425.

Choné, P., and G. Laroque. 2011. "Optimal Taxation in the extensive model." *Journal of Economic Theory* 146 (2): 425–453.

Coughlin, P., and S. Nitzan. 1981. "Electoral outcomes with probabilistic voting and Nash social welfare maxima." *Journal of Public Economics* 15 (1): 113–121.

Cukierman, A., and A. H. Meltzer. 1988. "A political theory of progressive income taxation.".

Cusack, T., and P. Beramendi. 2006. "'Taxing work'." *European Journal of Political Research* 45 (1): 43-73.

De Donder, P., and J. Hindriks. 2003. "The politics of progressive income taxation with incentive effects." *Journal of Public Economics* 87 (11): 2491–2505.

De Donder, P., and J. Hindriks. 2004. "Majority support for progressive income taxation with corner preferences." *Public Choice* 118 (3-4): 437–449.

De Freitas, J. 2009. "A probabilistic voting model of progressive taxation with incentive effects." *Hacienda pública española* (190): 9–26.

Feldstein, M. S. 1969. "The effects of taxation on risk taking." *Journal of Political Economy*: 755–764.

Feldstein, Martin, ed. 1983. *Behavioral Simulation Methods in Tax Policy Analysis*. University of Chicago Press.

Figari, F., A. Paulus, and H. Sutherland. 2015. "Microsimulation and Policy Analysis." In *Handbook of Income Distribution*, ed. A. B. Atkinson and F. Bourguignon. Vol. 2. Elsevier.

Förster, M., and M. Pearson. 2002. "Income distribution and poverty in the OECD area: trends and driving forces." OECD Economic Studies 34.

Ganghof, S. 2006. "Tax mixes and the size of the welfare state: causal mechanisms and policy implications." *Journal of European Social Policy* 16 (4): 360–373.

Holtz, E. D., W. Newey, and H. Rosen. 1988. "Estimating Vector Autoregression with Panel Data." *Econometrica* 56 (6): 1371–1395.

International Labor Organization. 2008. *Income Inequalities in the Age of Financial Globalization*. Geneva: International Labor Office.

Iversen, T., and D. Soskice. 2006. "Electoral institutions and the politics of coalitions: Why some democracies redistribute more than others." *American Political Science Review* 100 (2): 165–181.

Kakwani, N. C. 1977a. "Applications of Lorenz curves in economic analysis." *Econometrica* 45 (3): 719–727.

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

Kakwani, N. C., and N. Podder. 1976. "Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations." *Econometrica* 44 (1): 137–148.

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

Kenworthy, L. 2008. “Taxes and Inequality: Lessons from Abroad.” Last accessed June 12, 2018. [URL].

Kenworthy, L., and J. Pontusson. 2005. “Rising inequality and the politics of redistribution in affluent countries.” *Perspectives on Politics* 3 (3): 449–471.

Lambert, P. 2001. *The distribution and redistribution of income.* Manchester University Press.

Lehmann, E., A. Parmentier, and B. Van Der Linden. 2011. “Optimal income taxation with endogenous participation and search unemployment.” *Journal of Public Economics* 95 (11): 1523–1537.

Marhuenda, F., and I. Ortuño-Ortín. 1995. “Popular support for progressive taxation.” *Economics Letters* 48 (3): 319–324.

Martin, C. J. 2015. “Labour market coordination and the evolution of tax regimes.” *Socio-Economic Review* 13 (1): 33–54.

Martin, I. W., A. K. Mehrotra, and M. Prasad. 2009. *The new fiscal sociology: Taxation in comparative and historical perspective.* Cambridge University Press.

Meghir, C., and D. Phillips. 2008. “Labour supply and taxes.” *Available at SSRN 1136210.*

Newman, K. S., and R. O’Brian. 2011. *Taxing the Poor: Doing Damage to the Truly Disadvantaged.* Berkeley: University of California Press.

Nickell, S. 1981. “Biases in dynamic models with fixed effects.” *Econometrica* 49 (6): 1417–1426.

OECD. 2007. *Benefits and Wages.* Paris: OECD.

OECD. 2012. “Quality review of the OECD database on household incomes and poverty and the OECD earnings database.” OECD, Social Policy Division.

Persson, M. 1983. “The distribution of abilities and the progressive income tax.” *Journal of Public Economics* 22 (1): 73–88.

Pesaran, M. H. 2006. “Estimation and inference in large heterogeneous panels with a multifactor error structure.” *Econometrica* 74 (4): 967–1012.

Pesaran, M. H., and R. Smith. 1995. “Estimating long-run relationships from dynamic heterogeneous panels.” *Journal of Econometrics* 68 (1): 79–113.

Piketty, T., E. Saez, and S. Stantcheva. 2014. “Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities.” *American Economic Journal: Economic Policy* 6 (1): 230–271.

Prasad, M., and Y. Deng. 2009. “Taxation and the worlds of welfare.” *Socio-Economic Review* 7 (3): 431–457.

Rendon, S. R. 2012. “Fixed and Random Effects in Classical and Bayesian Regression.” *Oxford Bulletin of Economics and Statistics* 75 (3): 460–476.

Roemer, J. E. 1999. “The democratic political economy of progressive income taxation.” *Econometrica*: 1–19.

Roemer, J. E. 2011. “A theory of income taxation where politicians focus upon core and swing voters.” *Social Choice and Welfare* 36 (3-4): 383–421.

Rosenstone, S. J., and M. J. Hansen. 1993. *Mobilization, Participation, and Democracy in America*. New York: Macmillan.

Rubinstein, A. 1982. “Perfect equilibrium in a bargaining model.” *Econometrica*: 97–109.

Saez, E. 2002. “Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses.” *Quarterly journal of Economics* 117 (3): 1039–1073.

Schlozman, K. L., and S. Verba. 1979. *Injury to insult: Unemployment, class, and political response*. Cambridge: Harvard University Press.

Shor, B., J. Bafumi, L. Keele, and D. Park. 2007. “A Bayesian multilevel modeling approach to time-series cross-sectional data.” *Political Analysis* 15 (2): 165–181.

Slitor, R. E. 1948. “The measurement of progressivity and built-in flexibility.” *Quarterly Journal of Economics* 62 (2): 309–313.

Snyder, J. M., and G. H. Kramer. 1988. “Fairness, self-interest, and the politics of the progressive income tax.” *Journal of Public Economics* 36 (2): 197–230.

Steinmo, S. 1989. “Political institutions and tax policy in the United States, Sweden, and Britain.” *World Politics* 41 (04): 500–535.

Steinmo, S. 1996. *Taxation and democracy: Swedish, British, and American approaches to financing the modern state*. New Haven: Yale University Press.

A. Proofs of Formal Statements

A.1. Proof of Lemma 1

Individuals choose τ_1 and τ_2 to maximize utility. Each individual in group R therefore solves:

$$\frac{\partial V_R}{\partial \tau_1} = \sum_{i \in I \setminus R} p_i w_i G(\bar{\theta}, w_i)(1 - \eta_i) = 0 \tag{20}$$

$$\frac{\partial V_R}{\partial \tau_2} = -w_R + p_R w_R G(\bar{\theta}, w_R)(1 - \eta_R) = 0 \tag{21}$$

and each individual in groups M and P solve:

$$\frac{\partial V_M}{\partial \tau_1} = -w_M + \sum_{i \in I \setminus R} p_i w_i G(\bar{\theta}, w_i)(1 - \eta_i) = 0 \tag{22}$$

$$\frac{\partial V_M}{\partial \tau_2} = p_R w_R G(\bar{\theta}, w_R)(1 - \eta_R) = 0 \tag{23}$$

and

$$\frac{\partial V_P}{\partial \tau_1} = -w_P + \sum_{i \in I \setminus R} p_i w_i G(\bar{\theta}, w_i)(1 - \eta_i) = 0 \tag{24}$$

$$\frac{\partial V_P}{\partial \tau_2} = p_R w_R G(\bar{\theta}, w_R)(1 - \eta_R) = 0. \tag{25}$$

R gains by taxing M and P, and so via equation (20) chooses $\hat{\tau}_1 > 0$. Conversely, R cannot gain by taxing itself, so via equation (21), $\hat{\tau}_2 = 0$. M cannot gain by taxing itself or group P, and so chooses $\hat{\tau}_1 = 0$. In addition, via equation (23), M chooses $\hat{\tau}_2$ such that $\eta_R = 1$. Finally, provided that $\sum_{i \in I \setminus R} p_i w_i G(\bar{\theta}, w_i) > w_P$, then P chooses $\hat{\tau}_1 > 0$ and likewise chooses $\hat{\tau}_2 > 0$ such that $\eta_R = 1$. This establishes the proof.

□

A.2. Proof of Proposition 1

From equation (11), the first-order conditions for party A's proposed platform is:

$$\frac{\partial \pi_A}{\partial \tau_1} = \sum_{i \in I} p_i \left(\frac{1}{V_i} \frac{\partial V_i}{\partial \tau_1} \right) = 0 \tag{26}$$

$$\frac{\partial \pi_A}{\partial \tau_2} = \sum_{i \in I} p_i \left(\frac{1}{V_i} \frac{\partial V_i}{\partial \tau_2} \right) = 0 \tag{27}$$

It is easily checked that the first-order condition for party B is identical which ensures the symmetry of the problem. The Coughlin and Nitzan (1981) set up is sufficient to ensure the existence of a unique equilibrium as long as V_i is quasiconcave on τ.

First, assume that $w_P = 0$. Since $\hat{\tau}_P^1$ and $\hat{\tau}_R^1$ are strictly decreasing in w_P, τ_1^* is at its maximum—i.e., least progressive—at this income level. Therefore, the following proof will hold *a fortiori* for all $w_P \in [0, w_M]$. Therefore, from the first-order conditions, a necessary condition for the tax schedule to be progressive is:

$$1 < \frac{V_R G(\bar{\theta}, w_R)(1 - \eta_R)}{V_M G(\bar{\theta}, w_M)(1 - \eta_M)}. \tag{28}$$

Since $V_R > V_M$, $G(\bar{\theta}, w_R) > G(\bar{\theta}, w_M)$, and $\eta_R < \eta_M$ are always true, this condition will always be satisfied. The only necessary condition is the existence of inequality. Thus, as long as $V_R > V_M$, the condition will be satisfied even if participation rates, G, and participation rate elasticities, η, were equal. However, these participation rates and elasticities also clearly contribute to the existence of a progressive tax schedule. The intuition is that more revenue can be extracted from rich individuals and that without larger economic cost.

□

A.3. Proof of Proposition 2

We show that with a mean-preserving spread in the income distribution, from X to Y, both redistribution and progressivity are lower under Y than X. From equation (22) we know that $\partial V_M / \partial \tau_1 < 0$ for all $\tau_1 \in [0, 1]$. This implies from the first-order condition (26) that

$$\sum_{i \in I \setminus M} p_i \left(\frac{1}{V_i} \frac{\partial V_i}{\partial \tau_1} \right) > 0 \tag{29}$$

Therefore, for any $p_{X,P} < p_{Y,P}$, $p_{X,M} > p_{Y,M}$, and $p_{X,R} < p_{Y,R}$, we have $\tau_{1,Y}^* > \tau_{1,X}^*$.

Next, for any $\tau_{2,X} \in [0, \hat{\tau}_2^M)$ (which is always true in equilibrium), we know that $\partial V_M / \partial \tau_2 > 0$. Therefore, from equation (27), this implies

$$\sum_{i \in I \setminus M} p_i \left(\frac{1}{V_i} \frac{\partial V_i}{\partial \tau_1} \right) < 0 \tag{30}$$

Therefore, for any $p_{X,P} < p_{Y,P}$, $p_{X,M} > p_{Y,M}$, and $p_{X,R} < p_{Y,R}$ we have $\tau^*_{2,Y} < \tau^*_{2,X}$.

Because the change shifts income upwards, and the drop in τ_2 does not compensate for the increase in τ_1, (Lorenz curve) inequality in disposable income will increase, and therefore redistribution will decrease. □

A.4. Proof of Proposition 3

By assumption, but without loss of generality (see below), M is the formateur. Following Binmore, Rubinstein, and Wolinsky (1986), we can use the Nash bargaining solution to model a more involved sequential strategic bargaining model (Rubinstein 1982) as the time period between offers goes to zero. In an MP coalition, the solution to the bargaining problem is written as:

$$\tau^*_{MP} = \arg\max[(V_M(\tau_{MP}) - V_M(\tau_0))(V_P(\tau_{MP}) - V_P(\tau_0))] \tag{31}$$

The first order conditions for this problem are:

$$\frac{1}{V_M(\tau_{MP}) - V_M(\tau_0)} \frac{\partial V_M}{\partial \tau_1} + \frac{1}{V_P(\tau_{MP}) - V_P(\tau_0)} \frac{\partial V_P}{\partial \tau_1} = 0 \tag{32}$$

$$\frac{1}{V_M(\tau_{MP}) - V_M(\tau_0)} \frac{\partial V_M}{\partial \tau_2} + \frac{1}{V_P(\tau_{MP}) - V_P(\tau_0)} \frac{\partial V_P}{\partial \tau_2} = 0 \tag{33}$$

Since $\hat{\tau}_1^P > \hat{\tau}_1^M = 0$, equation (32) implies that $\tau^*_{MP,1} \in [0, \hat{\tau}_1^P]$. However, since from equations (23) and (25), $\partial V_P / \partial \tau_2 = \partial V_M / \partial \tau_2$, we also have $\tau^*_{MP,2} = \hat{\tau}_2^P = \hat{\tau}_2^M$.

By similar reasoning, we can establish equilibrium values for coalitions MR and PR. The critical facts are as follows. Because R voters prefer $\hat{\tau}_2^R = 0$, we have $\tau^*_{MR,2} < \tau^*_{MP,2}$. Moreover, because R voters prefer $\hat{\tau}_1^R > \hat{\tau}_1^P$, $\tau^*_{MR,1} > \tau^*_{MP,1}$. Given group M preferences, group M strictly prefers the MP coalition over the MR coalition, and will choose party P when it is the formateur.[30]

[30] It can also be shown that party R strictly prefers a coalition with party P to M and that party P strictly prefers a coalition with party M to R. This implies a higher likelihood of center-left coalitions (i.e., greater than one half) to center-right or neutral (RP) coalitions under proportional representation when parties are chosen to be the formateur at random. This reasoning is slightly different but the result is equivalent to Proposition II(ii) in Iversen and Soskice (2006). The difference consists precisely of the fact that regressivity is allowed in our framework and not in Iversen and Soskice's (see in particular, their assumption (5)).

It is immediate from the above that if policy in a majoritarian regime gives $\tau^* = \hat{\tau}^M$, then policy in a proportional representation regime reduces progressivity but increases redistribution.

$\square$

B. Empirical details

Table B.1 shows countries and years included in our analysis. In many cases, the limiting factor is information on inequality indices needed to calculate our measure of redistribution. Note that we conduct a robustness using multiple imputation (assuming that the process leading to missing inequality information in a given year is MAR) and found no substantive difference in results (see Table II).

Table B.1
Countries and years included in our analysis

Country	Years included in analysis
Australia	2004, 2008, 2010, 2012, 2014
Austria	2007–2015
Belgium	2004–2015
Canada	2001–2015
Denmark	2005–2014
Finland	2001–2015
France	2005, 2008, 2009–2015
Germany	2004, 2008, 2009-2014
Greece	2004–2015
Iceland	2004-2014
Ireland	2005–2014
Japan	2003, 2006, 2009, 2012
Netherlands	2005–2014
New Zealand	2003, 2008, 2009, 2011, 2012, 2014
Norway	2004, 2008, 2009-2014
Portugal	2004–2015
Spain	2007–2015
Sweden	2004, 2008, 2009-2015
Switzerland	2009, 2011, 2013, 2014
United Kingdom	2001–2015
United States	2005, 2008, 2009–2015

Table B.2 shows average tax function parameter estimates (across all years). Besides the clear differences in tax structures between countries, it also shows substantial over-time variation *within* countries: while the pooled standard deviation for τ is 4.8, the within country standard deviation is 1.6; for λ the overall standard deviation is 4.6 with an within-country SD of 2. We employ this within-country variation in our empirical analysis.

Table B.2
Summary of estimated tax function parameters

Country	τ [×100]	λ
Australia	17.73	5.51
Austria	17.38	5.53
Belgium	21.96	8.79
Canada	19.60	7.52
Denmark	21.23	10.78
Finland	14.82	3.61
France	6.71	1.85
Germany	15.12	4.53
Greece	19.83	7.15
Iceland	19.88	19.97
Ireland	17.99	6.00
Japan	7.98	3.22
Netherlands	24.45	11.49
New Zealand	10.51	2.48
Norway	16.40	6.72
Portugal	12.49	3.25
Spain	13.96	3.84
Sweden	19.76	9.07
Switzerland	13.03	4.85
United Kingdom	13.76	3.49
United States	10.88	2.89
Pooled mean	15.97	6.31
Pooled std.dev.	4.84	4.58
Within-country std.dev.	1.55	2.14

Note: Parameter estimates of equation 13, 2001–2015 averages. Within-country std.dev. calculated on $\tau_{it} - \bar{\tau}_i + \bar{\bar{\tau}}$ (*mutatis mutandis* for λ).

Figure B.1 shows the distribution of tax and benefit progressivity in our sample of 21 OECD countries (for 2001 to 2015).

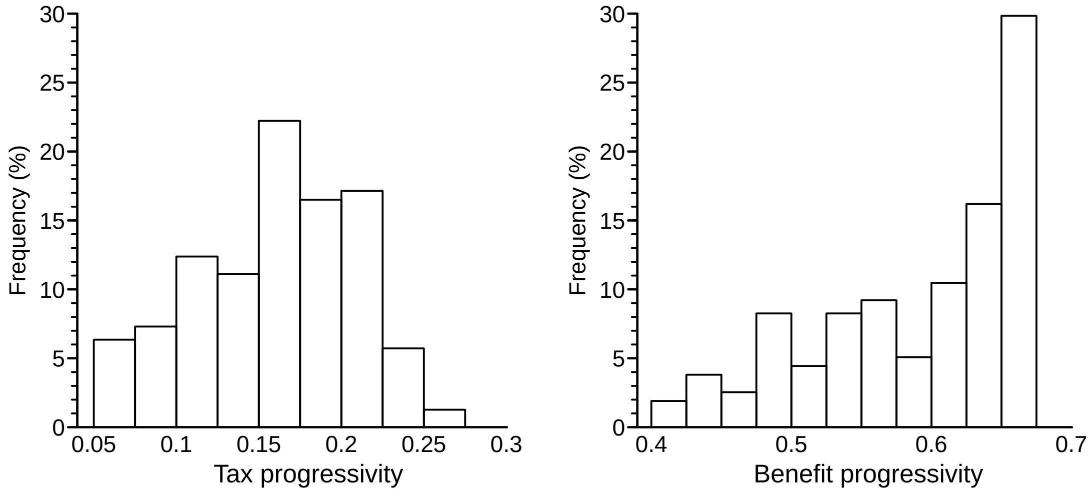

Figure B.1
Distribution of tax and benefit progressivity in 21 OECD countries, 2001-2015