One Dollar, One Vote *

Igor Barenboim
City of Rio de Janeiro

Loukas Karabarbounis †
Harvard University

October 2009

Abstract

We revisit the relationship between inequality and redistribution in a panel of advanced OECD countries. Using panel data methods that hold constant a variety of determinants of the public redistributive policy, we find a non-monotonic relationship between distribution of income and redistribution. Relatively to mean income, a more affluent rich and middle class are associated with lower and a richer poor class with higher public spending for redistribution. These results are consistent with what we define as a one dollar, one vote politico-economic equilibrium: When the income of a group of citizens rises (relative to mean income), aggregate redistributive policies tilt towards this group's most preferred public policy.

JEL-Codes: C23, D31, H50, P16.

Keywords: Inequality, Distribution, Redistribution.

*We are grateful to Alberto Alesina for many valuable comments and useful discussions. We also thank Elias Albagli, James Alt, Filipe Campante, Benjamin Friedman, Gita Gopinath, Dale Jorgenson, Lawrence Katz, James Robinson, Andrei Shleifer, Giacomo Ponzetto, Matthew Weinzierl and seminar participants at the Harvard Macro Lunch for helpful suggestions.

†Contact author: karabarb@fas.harvard.edu.

1 Introduction

What determines the amount of resources that a society redistributes between its members? The current consensus in the literature, summarized in the books of Persson and Tabellini (2003) and Alesina and Glaeser (2004), is that the pre-tax and transfer distribution of income is not a significant determinant of the size of the welfare state. The striking contrast between the US and Europe illustrates the "paradox of redistribution". The two workhorse models of distribution and redistribution, the normative model of Mirrlees (1971) and the positive theory of Meltzer and Richard (1981), in general predict that more inequality should be associated with more redistribution. However, in reality the more pre-tax and pre-transfer unequal US redistributes less than the more equal Europe.

Contrary to the conventional view in the literature, we argue that there is no paradox if we introduce different dimensions of income inequality in the same empirical framework and interpret our findings as a one dollar, one vote result. The motivation for interpreting cross country differences in redistribution through the lens of this framework is simple. Since heterogeneous groups of citizens have conflicting goals regarding the redistribution of resources and since income is strongly correlated with various measures of political participation, in principle a single summary statistic, such as the Gini coefficient or the distance of the median from the mean of the income distribution, is unlikely to account for all conflicting preferences regarding the size of the modern welfare state.

We measure inequality using three ratios. Inequality at the bottom of the distribution is given by the ratio of the gross earnings of the worker in the $10th$ percentile of the distribution relative to that of the mean. The ratio of the gross earnings of the worker in the $90th$ percentile relative to the mean captures the relative affluence of the rich. Finally, the familiar median over mean ratio measures inequality at the middle of the distribution. Figure 1 explains why we summarize inequality with these three indices. In the vertical axis we measure the unconditional correlation between total public social expenditure (as a percentage of GDP) and the earnings of individuals located in each percentile of the gross earnings distribution relative to the mean. In the horizontal axis we depict the percentile of the gross earnings distribution from a sample of advanced democratic OECD countries. The Figure shows, for instance, that the correlation of social expenditures with the ratio of the gross earnings of the worker located in the $10th$ percentile relative to mean gross earnings is around .50, when the latter is measured with a three year lag.[1]

Clearly, there are three areas of interest. The correlation of redistribution with the relative gross earnings of individuals below the median is positive, it falls to zero for the median and is negative for individuals richer than the median. This pattern arises when we look at our pooled sample, but

[1]Correlations for the 40th and the 60th percentile are not displayed because of the scarcity of this data.

we obtain a similar result at the cross country dimension of our sample for separate time periods. As Figure 1 shows, when we measure our inequality ratios with a 20 year lag relative to redistribution, the correlations for the poor and the rich do not decrease in magnitude. Based on this fact, we hypothesize that the omission of relevant variables is a more serious concern than reverse causation in attempting to estimate a causal link from distribution to redistribution. However, our econometric methodology is meant to address both concerns.

To test our hypothesis that inequality is relevant for redistribution, in Section 2 we construct a panel of advanced democratic OECD countries in 1978-2002. Our methodology differs from previous empirical papers in three key dimensions. First, and most importantly, we introduce three different indices of income inequality. Second, we measure inequality with gross individuals earnings which, as we explain later, addresses a number of econometric issues. Third, we focus on the within country variation of our sample, and therefore our methodology is in line with a prominent strand of literature that emphasizes persistent determinants of redistribution (such as institutions, culture, and geography). In other words, our estimated effect of inequality on redistribution holds constant all "long run" factors that may drive some societies to redistribute more than others, independently of differences in their income distributions.

We present our results in Section 3. Overall, we find a robust positive relationship between the relative gross earnings of the poor and aggregate spending for cash redistribution, a solid negative association for the median over mean ratio, and a broadly consistent negative association between relative earnings of the rich and redistributive spending. Therefore, *inequality at the tails* of the distribution, in the sense of a richer rich class and a poorer poor class, is associated with less redistribution. That the ratio of the median over mean gross earnings, the *equality at the center* of the income distribution, is associated with less public spending shows how different indices of income inequality are important for redistribution.

One contribution of our paper is to show how previous empirical results in the literature, similarly to the unconditional correlation of Figure 1, do not provide appropriate tests of the basic median voter equilibrium mechanism proposed by Meltzer and Richard (1981). The reason is simple. If we omit the earnings of the rich and the poor from the regression, we are implicitly assuming that the median voter is decisive. In reality, however, the political system is more complicated than what the one person, one vote model assumes, and various groups of citizens may have a "say" for the equilibrium outcome. Hence, the basic prediction that the median over the mean ratio of gross earnings is negatively associated with redistribution should be conditioned on the earnings of other groups of voters.

We provide a battery of robustness checks. Our results remain unchanged when we include a variety of controls and interaction terms with time invariant determinants of the welfare state. We also repeat our regressions with pensions and personal income taxes as measures for redistribution to address measurement issues in the dependent variable. Dropping influential observations, slicing the sample in different fashions, redesigning the structure of the panel errors, accounting for country specific trends, and including the lag of the dependent variable into the model does not affect our results. Finally, we verify most of our results by instrumenting for the inequality ratios, and therefore our empirical estimates are in general robust to the strict exogeneity assumption necessary for the fixed effects estimator.

Section 4 discusses our results. A well known theoretical proposition is that income and most preferred tax system are negatively related across groups of voters. As we show, this theoretical result finds considerable empirical support in microeconomic data. Therefore, the rich prefer less redistribution than the median, and the median prefers less redistribution than the poor. Starting from this result, we observe that the effects of the inequality ratios on redistribution are consistent with what we define as a *one dollar, one vote* politico-economic equilibrium. In the one dollar, one vote equilibrium, when a group of citizens becomes richer (holding constant mean income), redistribution tilts towards its bliss point. For the rich and the median this means that public spending decreases when they become more affluent, while for the low income group, public spending increases with their income. In other words, our one dollar, one vote definition captures the positive association between group income and distance of group most preferred outcome to equilibrium redistribution. To conclude we present the implications of our empirical results for previous theories of inequality and redistribution and, in light of our results, we show how the US-Europe difference in the size of the welfare state can be explained by differences in their income distribution.

2 Data and Methodology

In this Section, we describe the quality of the data, define our variables and discuss our estimation strategy.

2.1 Data Sources and Sample

We use a variety of sources to assemble an unbalanced panel of OECD countries covering the 1978-2002 period for 14 OECD countries.[2] The political economy model that we wish to test concerns

[2]These are: Australia, Austria, Canada, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, New Zealand, Sweden, UK and USA. Canada is observed only in the 1980s. For Italy we have data until 1995, i.e. before the change

economies that have an established tradition of democratic institutions. By focusing on advanced developed OECD countries we control for a variety of factors that are constant within and across countries such as the stability and the quality of the democratic institutions and the level of economic development. In Appendix 1 we provide detailed links to the original data sources. Table 1 presents the summary statistics.

2.2 Redistribution and Inequality Variables

Redistribution: The dependent variable is drawn from the OECD SOCX Database. Our measure of redistribution is Total Public Social Expenditure as a percentage of GDP which includes expenditures on old age (pensions), survivors, incapacity related benefits, health, family, active labor market programs, unemployment, housing and other categories but not education, which is not always progressive. We examine the robustness of our results to alternative measures of redistribution slicing total public social expenditures in health, pensions and expenditures net of health and pensions. As an alternative dependent variable we also use OECD's Tax Wedge variable, which measures the sum of personal income taxes and Social Security contributions (including employer's contributions).[3]

Earnings Distribution: An important issue in our empirical investigation is the definition and quality of the distributional variables. The quality of the distributional data has improved in the recent years, but it is still far from perfect (Atkinson and Brandolini, 2001). Our source is the OECD LFS dataset which has a number of advantages. The proper implementation of the basic political economy model requires the use of *gross earnings*, that is, pre-tax and pre-transfer earnings and not expenditures, consumption, wages, net earnings, net or disposable income. Using gross instead of net earnings has the important advantage that our measures of redistribution do not overlap with the inequality ratios. In addition, the OECD LFS database offers the possibility to construct a variety of distributional measures and focus on different aspects of the earnings distribution over a period of two decades. For most countries we construct the inequality ratios using the gross earnings of full-time workers. We have verified that our results are not sensitive to the few observations for which gross earnings data is not available.

Contrary to most empirical studies in redistribution, growth and inequality, we map the theory fairly close to the data and proxy the Meltzer and Richard (1981) ratio with the ratio of the gross

in its political institution.

[3] In addition, we have verified our results using Milanovic's (2000) reduction in inequality variables as a proxy for redistribution. Still, we prefer to use total public social expenditure for two reasons. First, the number of observations drops significantly when we overlap Milanovic's (2000) with our sample. Second, Lind's (2005) argument that there is a mechanical positive correlation between the Gini coefficient and Milanovic's measure of redistribution may also apply in our case.

earnings of the worker who is located in the 50th percentile over the mean, $y^{50}/\bar{y}$.[4] Typically, the literature uses either the Gini coefficient or the decile ratio of earnings, y^{90}/y^{10} (or the ratio y^{90}/y^{50}), to proxy for inequality. Our intuition is that the Gini coefficient cannot capture the non monotonic relationship between inequality and redistribution. An example that contrasts the distribution of factor income in Sweden and the US in the early 1980s illustrates this point. While Sweden has a Gini coefficient of 0.463 and the US one of 0.464, the US is much less equal than Sweden when we examine other dimensions of inequality. According to the OECD LFS dataset, the ratio of the gross earnings of the worker located in the 90th percentile of the gross earnings distribution over these of the worker at the 10th percentile is 3.79 for the US while 2.05 for Sweden. If we use the Gini coefficient to summarize the income distribution, then we overlook the fact that inequality at the tails of the distribution is much more pronounced in the US than in Sweden. In addition, as we discuss in Section 4, the demand of the rich and that of the poor can also affect the size of the welfare state, and in reality the median voter may not be pivotal. We proxy for the rich's and the poor's demand for redistribution with the $y^{90}/\bar{y}$ and the $y^{10}/\bar{y}$ ratios respectively.[5]

2.3 Empirical Specification

We consider the estimating equation:

$$RED_{i,t} = c_i + \tau_t + \beta_1 \left(\frac{y^{90}}{\bar{y}}\right)_{i,t-1} + \beta_2 \left(\frac{y^{50}}{\bar{y}}\right)_{i,t-1} + \beta_3 \left(\frac{y^{10}}{\bar{y}}\right)_{i,t-1} + \sum_{k=1}^{K} \gamma_k X_{k,i,t-1} + \varepsilon_{i,t} \qquad (1)$$

where $RED_{i,t}$ is Total Public Social Expenditure as a share of GDP for country $i = 1, ..., N$ in period $t = 1, ..., T$, and $y^j/\ \bar{y}$ is the ratio of gross earnings of the worker in the jth percentile to mean gross earnings, for $j = 90, 50, 10$. $X_{k,i,t-1}$ denotes other potential explanatory variables, c_i is a time invariant country specific unobservable effect, τ_t is a common unobservable time effect and $\varepsilon_{i,t}$ is the time varying country specific idiosyncratic error. All variables except for the coordination index, the

[4] There are two notable exceptions. Perotti (1996) uses the size of the middle class which is close to our $y^{50}/\bar{y}$ variable. He finds that more inequality is associated with less growth, but he rejects that the link between these two variables occurs through redistribution, as hypothesized by workhorse inequality and growth models. Rodriguez (1999) carefully tests the Meltzer-Richard (1981) model using pre-tax median over mean earnings. He presents time series and cross section evidence from the US states which reject the median voter model.

[5] Other models use the y^{90}/y^{10} ratio to proxy for the median voter's demand and interpret the negative estimated coefficient as a failure of the Meltzer-Richard model. See for instance, Moene and Wallerstein (2001) who offer an insurance interpretation to rationalize the negative effect and Iversen and Soskice (2006) who also find a negative or close to zero correlation coefficient for the 90-50 ratio. In light of the correlations in Figure 1, the negative correlation of the 90-10 with redistribution is not surprising, since both the 90th percentile in the numerator and the 10th percentile in the denominator drive a strong negative overall correlation. The same intuition applies in the case of the 90-50 ratio: If the 90-mean ratio is strongly negatively correlated with redistribution but the 50-mean is uncorrelated or weakly negatively correlated with redistribution, then the 90-50 ratio (90-mean over 50-mean) should be (weakly) negatively correlated with redistribution.

government surplus and growth are in logs, and therefore the regression coefficients are the elasticities of redistributive spending.[6]

Assessing the causal effect of inequality on redistribution is certainly a challenging task; in general we cannot expect inequality to be exogenous. The first source of endogeneity is contemporaneous reverse causation and in fact the often reported negative relation between inequality and redistribution may simply reflect the case that societies redistribute to decrease their economic inequality. Using gross earnings instead of net earnings or consumption to construct the inequality indices relaxes somewhat this constraint because the latter vary mechanically with the fiscal system, while the former only through the endogenous response of labor supply. To mitigate this concern we follow the more formal procedure to average all variables using three year intervals and lag regressors one period. For example, in period $t = 2001$, RED_t is an average of total public spending in years 1999, 2000 and 2001. This is regressed on right hand side variables which are averaged in years 1996, 1997 and 1998, and therefore, on average, inequality leads redistribution by three years. Our averaging procedure has additional advantages, as it removes transitory fluctuations due to economic business cycles, it captures meaningful variations along the political cycle and it reduces serial correlation and measurement error.

Probably the most serious concern for our empirical investigation is the potential inconsistency from the omission of relevant observable or unobservable determinants of redistribution. The panel that we construct readily controls for *time non-varying* observable or unobservable country specific effects. This is important since the literature generally emphasizes long-run, persistent factors as determinants of redistribution. The electoral rules, form of the government, judicial review and federalism (Persson and Tabellini, 2003; Alesina and Glaeser, 2004, and references therein), persistent cultural characteristics such as beliefs in a just world (Benabou and Tirole, 2006) and trust (Tabellini, 2007), ethnic fragmentation (Easterly and Levine, 1997; Alesina and Glaeser, 2004), prospects of upward mobility (Piketty, 1995; Benabou and Ok, 2001), social beliefs about fairness (Alesina and Angeletos, 2005), legal origins, initial technological capabilities, and geography may vary across countries but are fixed within country for the twenty year long period that we consider. While all these factors are important determinants of the cross country variation in redistribution, there is no loss in generality if we delegate them to the country specific intercept, since our interest lies in the distributional variables. Given that it is very likely that our inequality ratios are endogenously induced by these factors, we allow for any arbitrary correlation between the fixed effect and our regressors.

In addition, we include time effects in (1) which control for common time trends, such as the

[6]The results are similar when we specify the equation in levels.

technological slowdown caused by the oil shocks in the beginning of the sample or the world rise of rightist movements in the 1980s. Time dummies also reduce the effects of spurious trends and contemporaneous panel error correlations. Adding $N+T-2$ dummy variables into the model certainly creates a demanding environment to test our hypothesis since it removes most of the cross and within country variation in the data. However, we adopt this technique because it mitigates the endogeneity problem and helps us identify the effects of distribution on redistribution.

The consistency of the fixed effects estimator is subject to the strict exogeneity assumption. Specifically, factors affecting redistribution in period t but omitted in the error term, are assumed orthogonal to past, present and future explanatory variables (Wooldridge, 2002). This assumption may be too restrictive and we relax it in Section 3.2.

3 Income Inequality and Redistribution

3.1 Main Results

Table 2, column 1, presents our baseline specification. There is a clear positive association between the poor to mean gross earnings ratio and total social public expenditures. This coefficient is significant at 1% level. The median to mean ratio is negatively correlated with redistribution and it is significant at 5% level. Finally, the rich to mean ratio is negatively related to redistribution. This relationship is significant at 1% level. In the following columns, we slice the Total Social Public Expenditures variable (RED henceforth) in its major three components, namely Health, Pensions and RED net of Health and Pensions. We find a similar pattern in terms of signs and significance for the distribution coefficients when we use Pensions and RED net of Health and Pensions as our dependent variable. However, the distribution variables do not seem relevant to explain the within country variation of Public Health Expenditures. This is a reasonable result in light of the fact that our one dollar, one vote hypothesis should be stronger for more progressive social programs, and health is understood as one of the least progressive social expenditures.

The fact that our results remain broadly unchanged when we use the Pensions variable instead of RED, shows clearly how the association between our distributional ratios and redistributive expenditures is not artificial. The reason is simple. Our inequality ratios concern the *gross* earnings of full time workers, while by definition pensions are given to retirees.[7] As a result, our inequality measures do not overlap mechanically with redistribution.

[7] Tabellini (2000) argues that since pensions redistribute also within generation, a small degree of altruism of the grandchild towards the grandfathers can sustain pensions in equilibrium even in the absence of commitment. In general, the consensus in the literature appears to be that pensions are mildly redistributive within-cohorts. See Liebman (2001) for a microsimulation model applied at the US Social Security system.

Our baseline specification controls for four other variables. GDP controls for the size of the tax base and also for Wagner's Law. There is no evidence that GDP affects redistribution, which is reasonable since our sample is the most advanced from the developed democracies. We control for economic growth, to take into account the counter-cyclicality of fiscal policy driven by automatic stabilizers (e.g. unemployment compensations).[8] We find strong evidence of counter-cyclical redistribution, in line with previous studies for OECD countries. The share of the old population holds constant demographic characteristics and the bargaining power of the elderly when demanding more directed transfers, i.e. pensions. We find that a larger share of elderly population is associated with a larger redistributive scheme. Finally, the deadweight loss, which measures the opportunity cost of taxation, is associated with less redistribution. When we omit all these controls the estimated parameters are very similar (not shown).

In column 5 of Table 2 we use the sum of personal income taxes and payroll taxes, a measure of the net tax burden for the average earner, as the dependent variable. We repeat our regressions with this measure of redistribution to verify that our results are not sensitive to cyclical fluctuations in output. During economic booms, spending on automatic stabilizers declines. If in booms the income benefit for the rich exceeds that of the poor, then the $y^{90}/\bar{y}$ ratio grows and the $y^{10}/\bar{y}$ falls, which could explain the negative coefficient for the former and the positive coefficient for the latter. The tax wedge variable addresses this concern, since personal income taxes and Social Security contributions are statutory. Note that if poor's relative earnings are counter-cyclical and rich's relative earnings are pro-cyclical, then fluctuations in output tend to bias downward the poor's coefficient and upward the rich's coefficient, as in both cases relative earnings vary but taxes are constant. Column 5 of Table 2 shows how the estimated coefficients for the rich and the poor remain robust, while the coefficient for the median still enters negatively but without significance. In addition, we note that there is not anymore an effect of growth on redistribution, which shows how the tax wedge variable is immune to cyclical variations, as initially hypothesized.

We reject the null hypothesis that all country fixed effects are not significant at the 1% level. Time dummies also enter jointly significantly. The exclusion of the time effects affects only the significance of the median over the mean ratio. Estimation with random effects does not alter our results and therefore our estimation is robust to a potential sensitivity of the FE estimator due to measurement error. Still, we view the omitted variable bias as of first order concern and prefer the use of fixed effects.

The economic significance of the three faces of inequality is also large. According to the first

[8] Growth is the only control variable which we do not lag in our regressions, because theoretically the automatic stabilization should take place contemporaneously.

column in Table 2, the conditional elasticity of redistribution with respect to the demand of the rich, the Meltzer-Richard ratio and the demand of the poor are respectively $-1.06, -1.44$ and 1.01. To understand the magnitude of these coefficients, suppose that some country redistributes the mean value of total social public expenditure in our sample, 21.82%. A 10% increase in each of the three ratios independently would change redistribution respectively by $-2.31, -3.14$ and 2.20 percentage points of GDP. The magnitude of these variations is reasonable both in the within and the across dimension of our sample.[9] In Section 4.3 we show how our estimates can explain the US-Europe difference in redistribution.

In Table 3 we add other time varying controls to our baseline specification. We add controls in isolation in columns 1-5 and include all variables in column 6. We include respectively, the rate of unemployment, openness, share of left vote, voter turnout and the wage coordination variable. None of these controls are significant at the conventional levels, but they do present signs in accordance with theory. Unemployment is positively associated with redistribution, which would be expected since unemployment benefits compose RED. Openness also enters with a positive coefficient which is in line with Rodrik (1998), given that the size of the country is absorbed by the fixed effect. Left vote enters positively independently as left wing parties are associated with more redistributive policies, however its sign changes when more controls are included. Voter turnout enters positively in the regression as supposedly the poor voters are the ones more elastic to voting. The wage coordination variable enters with a negative coefficient which shows a substitutability between labor market rigidities and cash redistribution.

We have controlled for other potentially relevant variables—to save space we do not report these results. Controlling for the percentage of right vote in the last elections, unionization, and the budget surplus of the government does not change the results shown in Tables 2 and 3. Adding the employment ratio into our regressions also does not overturn the magnitude and significance of the three inequality ratios. The employment ratio addresses a potential sensitivity of the poor's coefficient due to the normality of leisure. If a more generous lump sum transfer from the welfare state causes some low ability workers to exit the labor market or participate only part-time, then the earnings of the worker located in the 10th percentile of the distribution of full-time gross earnings would rise artificially with redistribution.[10]

[9] A 10% increase in the mean over the median ratio is an increment of more than two standard deviations. In 1978, France's $y^{50}/\bar{y}$ ratio was 10% lower than the same statistic for the US. A 10% increase in the rich to mean ratio is slightly more than a one standard deviation increment, whereas a similar change in the poor to mean ratio corresponds to slightly less than one standard deviation. For instance, the $y^{90}/\bar{y}$ ratio in the US is 20% above that of Sweden and its $y^{10}/\bar{y}$ ratio 40% below that of Sweden in 1999.

[10] Note for the median and the rich, this effect goes against finding the negative coefficients.

We have interacted our inequality variables with other determinants of the welfare state (results available upon request). In the simplest formulation of a voting model it is assumed that everyone votes, but obviously the level of inequality matters for redistribution to the extent that the poor vote. To test this conjecture, we interact the Meltzer-Richard ratio with the turnout of the voters in the last elections. This coefficient has the predicted sign, but it is not significant and our results regarding the impact of the inequality ratios on redistribution do not change. Similarly, we find the effect of poor's income on redistribution increases in voter turnout, but the effect is quantitatively small. Second, we interact the Meltzer-Richard ratio with the ELF index of Roeder (2001) to test Lind's (2007) conjecture that only in less fragmented societies inequality increases redistribution. In accordance to his intuition, we find that less heterogeneous societies redistribute more after an increase in inequality, while our results for the rich's and the poor's demand remain the same. Finally, we have used interaction terms of political institutions (electoral rules and form of government) with our inequality ratios with no difference in the results.

As pointed out in the Introduction, our empirical methodology rests on three key features. First, we construct our inequality variables with gross earnings and use median over mean gross earnings to proxy for the Meltzer-Richard ratio. Perotti (1996) uses the size of the middle class and, most notably, Rodriguez (1999) uses pre-tax median over mean earnings to proxy for the Meltzer-Richard ratio. Both papers reject the prediction that a higher Meltzer-Richard ratio is associated with less redistribution. Second, we control for all long run determinants of redistribution with country fixed effects. When we drop the rich and the poor inequality variables from our regressions, the coefficient for the Meltzer-Richard ratio enters insignificantly and in many cases with the wrong sign, with or without fixed effects. Therefore, what changes the sign and the significance of the estimated coefficient is the third feature of our empirical specification, namely that we control for the probability that the median voter may not be pivotal by introducing other dimensions of income inequality. In fact, in Section 4 we discuss how the rich and the poor can engage in activities that bring the size of redistribution closer to their most preferred size.

Most of the literature fails to find evidence in favor of the hypothesis that a higher Gini coefficient is associated with more redistribution. For example, see Alesina and Glaeser (2004), Persson and Tabellini (2003) and the references listed therein. Recently, Shelton (2007) tests carefully several prominent theories on the determinants of government expenditures and their composition, including the median voter theory. Using a random effects estimator, he estimates a close to nil impact of the Gini coefficient on transfers, evaluated at the mean value of political rights. He also finds no significant pattern in any other measure of redistribution.

A notable exception is Milanovic (2000) who argues that a careful construction of the relevant variables accounts for the link from more inequality (as proxied by the Gini coefficient) to higher redistribution. Given the non monotonic relationship between inequality and redistribution estimated in Tables 2 and 3, it is interesting to test the Gini coefficient in our sample. We use the factor Gini coefficient from Milanovic's (2000) accurate calculations. More often than not, we find no significant relationship between the Gini coefficient and the share of total public social expenditure in GDP, using various specifications with fixed effects. This result does not compare directly to Milanovic because we have fewer observations, but perhaps most importantly, because our measure of redistribution differs (see footnote 3 for more on this point).

3.2 Sensitivity Analysis

We consider a number of further robustness exercises. None of the robustness checks overturns our results, however, in some cases one of the three coefficients loses significance.

Influential Observations: Given the small number of observations and the correlation between the regressors, our results may be sensitive to the inclusion of some outliers. In Figures 2 we show plots from the regression of the first column in Table 2. These and similar graphs for the rest of the regressors show no obvious outliers. A formal procedure is to examine the standardized residuals from our regression. There are six observations that have standardized residuals above 2. Dropping in isolation or all together these six observations from our sample does not alter our results. As a further robustness exercise, we calculate each observation's Cook's distance, which is a measure of the total influence of each observation on the predicted values of redistribution. Loosely speaking, observations with a distance greater than 1 merit special consideration. In our model, the highest value is in the low .20s. Another robustness check is to slice the sample and for instance restrict the sample only to European or Anglo Saxon countries, or only to the 1990s. We find that the signs of the distributional coefficients remain intact, but sometimes lose significance because of the major loss in power.

Structure of Panel Errors and Country Trends: Thus far we have followed an agnostic strategy and considered robust clustered error (that is, errors robust to any type of correlation within country). In Table 4, we consider various correlation structures. Panel heteroscedasticity allows the variance of the error term to differ across panels. The contemporaneous correlation of the panel errors is addressed by time dummies that capture common trends. We also consider common cross-country autocorrelation of the error term and different autocorrelations across panels. Another concern is that country specific time trends to distribution and redistribution could drive our results. We augment our specification and add N country specific trends.

We undertake various estimation strategies, each of which imposes a different set of assumptions on the error term. In the first three columns of Table 4, we perform FGLS estimation. It is well known that FGLS estimation has two disadvantages. First, it requires strong assumptions on the structure of the standard errors. Second, as Beck and Katz (1995) show, it may severely underestimate the variability of the estimates because each round of iteration compounds the uncertainty over the coefficients. Therefore, in the next three columns we use Panel Corrected Standard Errors, which is a strategy more agnostic than FGLS. Finally, in the last two columns we consider Newey/West standard errors which correct for arbitrary heteroscedasticity and first and second order serial correlation. As Table 4 shows, our results remain strong for the coefficients of the rich and the poor, but the median to mean ratio loses significance. We have repeated these regressions using Pensions as the dependent variable and our results do not change.

Endogeneity and Dynamic GMM: The consistency of our fixed effects estimator requires the assumption that factors delegated to the error term, $\varepsilon_{i,t}$, and as a result affect redistribution in period t, remain conditionally uncorrelated with past, present and future values of our inequality variables. However, our inequality ratios may affect redistribution with a lag. An even more serious concern is that causation may run from current redistribution to the future distribution of income. Therefore, the unexplained part of redistribution in period t may correlate with *future values* of inequality. It is not easy to predict the possible direction of bias. We hypothesize that for the poor, our positive coefficient is downward biased, because current redistribution may cause moral hazard problems in the future and decrease the effort and the earnings of the poor. For the rich, moral hazard problems are less likely to exist because the rich do not benefit from redistribution. On the other hand, current redistribution could affect negatively the gross earnings of the rich, through dynamic distortions in their supply of labor. Therefore, one conjecture is that the negative coefficient for the rich in Tables 2 and 3 is downward biased.

To address these concerns we estimate the model in first differences. In the same exercise, we also control for the lag of the dependent variable to capture the persistence in the size of the welfare state. It is well known that with the lagged dependent variable included into the model, the OLS estimator is upward biased if the error term is positively autocorrelated. Also, the FE estimator is downward biased by construction, but the bias decreases as the number of periods grows. In the first two columns of Table 5 the coefficient of the lagged dependent variable is .80 for the FE technique and .88 for the POLS estimation. Therefore, the true value of the coefficient lies in between these two estimates or at least close to these when we take into account the uncertainty for the point estimates.

In Table 5, we examine the performance of our model under two estimation procedures, namely the difference GMM method of Arellano and Bond (1991) and the system GMM model of Blundell and Bond (1998).[11] The latter estimator differs from the former because it stacks together the equation in first differences and the levels equation in a system of equations. The additional assumption of the system GMM estimator is that changes in redistribution and inequality are orthogonal to the country fixed effect.

In Table 5, columns 3 and 4, we do not use the lags of the distributional variables as instruments. Therefore, the model is estimated under the assumption that past, current and future values of the inequality ratios correlate arbitrarily with the error term. In columns 5 and 6, we report estimates which use all lagged regressors as instruments, i.e. under the assumption that past inequality is exogenous, but current and future inequality is endogenous to the current error term. We have also used Pensions as the dependent variable and our results do not change.

The results in Table 5 show that the estimated coefficient for the lagged dependent variable in the System GMM method is close to the interval of true values (as implied by the FE and the POLS techniques). The Difference GMM method however appears to deliver downward biased estimates.[12] Focusing on the System GMM results (columns 3 and 5), the coefficient for the rich loses significance. This accords well with the intuition that redistribution may affect negatively the labor supply of the rich in the future. Our results for the coefficients of the median and the poor are robust and retain their sign and statistical significance.

4 Discussion

In the previous Section we showed a robust association between income inequality and redistribution. While the main focus of the paper is to report this novel relationship, in this Section we briefly associate our results to existing theoretical literature. We first define our one dollar, one vote equilibrium and explain why the empirical results are in line with this definition. Then we propose some mechanisms from previous literature that may explain the empirical one dollar, one vote result. We conclude

[11] While relaxing the strict exogeneity sounds promising, the problem is that too many instruments may severely overfit the model. Dynamic panel estimators are not expected to behave perfectly for small N panels such as ours (Roodman, 2006). For instance, one common problem is that J-tests lack power because the number of instruments exceeds the total number of groups. In this case, the Hansen/Sargan test for overidentification lacks power and tends to over-accept the null hypothesis, with p-values of 1.000 being common. Notwithstanding these caveats, we believe that a sound robustness practice is to consider different modeling strategies and examine the sensitivity of our results to the assumed specification.

[12] This result is in line with Blundell and Bond (1998) because inequality and redistribution are fairly persistent processes. Bobba and Coviello (2007) apply this argument in the democracy and education literature and show the downward bias of the Difference GMM estimator.

with the implications of our work for the US vs. Europe paradox of redistribution stated in the Introduction.

4.1 One Dollar, One Vote

We define one dollar, one vote as a political system in which aggregate outcomes tilt towards a specific group of citizens' preferences, as this group's income (relative to the mean income) increases.

Our empirical results in Section 3 are consistent with this definition. To see this start with the well known theoretical result that an individual's most preferred tax rate is inversely related to individual income, a result that finds considerable empirical support.[13] Alesina and La Ferrara (2005) have shown for the US that richer individuals (currently, but also in expectation), prefer less redistributive policies. In Appendix 2 we extend their exercise for our panel of OECD countries using data from the World Value Survey dataset. We also find, that as income increases, support for redistribution decreases, for every country and every time period in our sample.

Therefore, the poor prefer more redistribution than the median and the median more than the rich. If the initial equilibrium tax rate lies between the bliss points of the poor and the rich, the fact that redistribution increases in the income of the poor implies that as the poor become richer, aggregate spending tilts closer to their most preferred outcome. When median income rises relative to the mean, the median's net benefit from redistribution decreases. The one dollar, one vote definition holds trivially in the Meltzer and Richard (1981) model, since the median is always decisive. This system also holds for the median in an extended framework, where the median voter is not exclusively pivotal, but in which its decisiveness increases with its relative income, holding constant the relative political power of the other groups. In either framework, the median desires less redistribution and also has the political power to impose this lower level of taxes as an equilibrium outcome. Hence, aggregate spending goes down, in accordance with our empirical estimates. Finally, it is easy to see that if the rich desire less taxes and redistribution decreases in the rich to mean ratio, then the one dollar, one vote definition is also satisfied for the high income group, since the rich in general desire less (or no) taxes.

4.2 Explanations

If the median class is the pivotal voter, then the conditions in the Meltzer and Richard (1981) framework apply. Holding constant the size of the tax base, this theory implies no role for other indices of inequality, a hypothesis that our evidence clearly refutes. In order to shed light on the empirical

[13] See e.g. Romer (1975) and Roberts (1977) for the theoretical derivation.

evidence presented in Section 3, we need to introduce a more vivid role for the poor and the rich relative to the median voter model.

When the rich become richer, their most preferred tax rate decreases. As a result, the one dollar, one vote definition is satisfied if the rich's pivotal power also increases with income. In this case the rich group will impose lower taxes, shifting the aggregate outcome closer to its preferences. On theoretical grounds, the positive association between income and pivotal role of the rich can be explained by the ideas developed by Grossman and Helpman (1996). Because a richer rich class has more to lose from redistribution, the incentive for them to contribute in political parties and tilt policy in their favor is expected to increase with income. Benabou (1996) uses this general insight to introduce a non monotonic relationship between inequality and redistribution. Campante (2007) lays out a model that describes this behavior, with specific reference to campaign contributions.

Overall, when we contrast the micro facts with the macro outcomes, we see a one-to-one correspondence between preferences for redistribution and aggregate policies for the middle and the rich class: Both desire less redistribution as they become richer, and aggregate spending decreases in the earnings of these two groups. On the other hand, the micro and the macro facts go in the opposite direction for the poor: Although, we find no evidence that poor persons enjoy paying higher taxes when they get richer, in Section 3 we presented strong evidence at the macro level that a relatively more affluent poor class is associated with a broader welfare state. Because income is inversely related to the preference for redistribution, it follows that (i) the poor in general prefer less taxes when they get richer, but also (ii) the poor's bliss point always involves higher taxes than that of the other groups. The evidence at the macro level that redistribution increases in poor's gross earnings, implies that the second effect dominates the first. Or, that in response to a richer low-income group of citizens, public policy moves closer to the poor's most preferred tax system, increasing spending on welfare. This is the essence of our one dollar, one vote definition.

The question then becomes, why policy moves closer to the poor's bliss point as they get richer. There are various plausible mechanisms. One natural explanation is political participation. Rosenstone and Hansen (1993) and McCarty, Poole and Rosenthal (2006) have shown how in the US political participation and income are strongly and positively correlated. To verify that these results are not special to the US, in Appendix 2, we present evidence from the World Value Survey dataset. We proxy for political participation using 14 different indicators, ranging from signing petitions, to participating in demonstrations, boycotts and political parties. Our findings indicate that income is strongly and positively associated with political participation across all income spectrum for all countries and time periods in our macro-sample. The poor do not contribute in cash for campaigns in a relevant manner, however they do engage in activities that influence policy, and engage in them more as they become richer. A richer poor class demonstrates more, participates in political parties, voluntary political clubs and labor unions. If the stronger political voice of the poor dominates their decreased preference for redistribution, then public spending increases in poor's income. That is, point (ii) above dominates (i).[14]

The political participation explanation—especially for the poor—assumes that, policy is sufficiently elastic to rich's and poor's political participation, i.e. that the position of the decisive voter is not rigid. We also consider some alternative explanations that do not assume the reallocation of political power, and can rationalize why equilibrium policy moves closer to groups of citizens with higher income.

Our regressions include a proxy for the deadweight loss of taxation.[15] With a progressive system of taxes, however, the distortionary effects of labor income taxation are higher at the right tail of the income distribution. Therefore, a richer high-income class could be associated with less taxes if the median voter internalizes these increasing distortions when demanding redistribution.[16]

The social affinity theory developed by Kristov, Lindert and McClelland (1992), could potentially rationalize the effects we find at the left tail of the distribution. The authors develop a model in which political agents organize into pressure groups and devote resources to affect redistribution, as in Peltzman (1980). The equilibrium redistribution is determined in the following manner: one feels closer to people that have a similar level of income. The theory predicts a positive sign for the gap between the poor's and the middle's earnings coefficients in a regression of redistribution on this income gap. We note that the authors allow the interpretation that social affinity operates through social mobility as in Benabou and Ok (2001). Therefore, when the median income gets closer to the income of the poor, redistribution increases. This mechanism could be part of a model in which the median remains decisive and hence there is no need to consider reallocations of political power.

A third, plausible, explanation comes from relaxing the assumption of unidimensionality in the Meltzer and Richard (1981) model. In both sides of the Atlantic, modern welfare states include means-tested redistributive programs. Examples of such programs are the Medicaid and Food-Stamps in the

[14]Note that any model that collapses to maximizing a weighted average of citizens' utility, as for instance the utilitarian model or its positive version, the probabilistic voting model, cannot explain the poor's positive coefficient. As poor's income (or ability, with endogenous labor supply) increases, the Pareto weight attached to the poor class declines, and hence the planner or the politician tilts redistribution closer to the median's or the rich's most preferred policy. This result holds under fairly general conditions, and is the essence of utilitarian redistribution. See Hellwig (1986) for these conditions.

[15]Constructed by dividing the top corporate marginal tax rate by government's share on GDP. A higher value for this index denotes higher distortions per unit of government activity.

[16]Note that this explanation is close to the theory developed by Becker (1983) and Becker and Mulligan (2003).

US, family and solidarity allowances in many European countries, and parts of the medical insurance in France. The statutory nature of redistribution provides an explanation for our results because if transfers are declining with income, a richer poor class becomes less eligible for specific target transfers and a higher fraction of tax revenues can be appropriated by the middle class. If political power remains in the hands of the median class, then the median's increased benefit from receiving transfers will change the politico-economic trade-off in favor of higher taxes.

4.3 The US-Europe Difference in Redistribution

In the Introduction we referred to the striking divergence of the data from the theory to motivate the wide held belief that inequality does not matter for redistribution. In light of our results it is worthwhile to briefly revisit this issue. Consider the US versus a European aggregate of Finland, France, Netherlands and Sweden (two Continental welfare states and two Nordic).[17]

Between 1980 and 2001, European growth in redistribution exceeded US growth by 2.7%. The European aggregate redistributed approximately 26.4% of its GDP in 2001, while 23.8% in 1980. US redistribution rose from 13.3% in 1980 to 14.4% in 2001. Consider, as an illustrative example, the estimates in the first column of Table 3. For this period of time, the growth of the $y^{10}/\bar{y}$ ratio is -2% in Europe but almost -21% in the US. So, multiplying the difference, 18.8%, with the estimated coefficient for the poor in Table 3, leads to an approximately 15.6% change in redistribution. The growth of the $y^{50}/\bar{y}$ ratio is less than -1% in Europe but almost -12% in the US. Multiplying 11.2% with the estimated coefficient, leads to an approximately -15.2% change in redistribution. So according to these estimates, Europe increased redistribution relative to the US in 2001 relative to 1980 because the European poor did not become as poor as the American poor (both relative to their mean), but US increased redistribution relative to Europe because the US median voter became poorer. Finally, the growth of the $y^{90}/\bar{y}$ ratio is -3% in Europe but 0.5% in the US. Multiplying -3.5% with the estimated coefficient, leads to an approximately 3.9% change in redistribution. Hence, the fact that the rich became richer in the US and poorer in Europe explains the rest of the variation. Actually, these changes in inequality overpredict the difference in the growth of redistibution across the two countries. Of course these predictions are illustrative and they change as we vary the specification (usually they overpredict the difference), and as we consider or drop other countries from the sample. In any case, the main conclusion from of our analysis is that looking at different parts of the income distribution offers a convincing explanation of cross country differences in redistribution. A single summary statistic is unlikely to capture all these heterogeneous effects.

[17] Germany and Italy have missing observations and cannot be included in this calculation.

5 Conclusion

This paper uses various indices of income inequality to explain redistribution. Our results challenge the conventional view that inequality does not matter for redistribution. We show how the gross earnings of the rich, median and poor (relative to mean earnings), all matter for redistribution. This finding strongly suggests that the use of a single summary statistic to measure inequality may not be appropriate. A contribution of our paper is to explain why the literature has found scant evidence in favor of the politico-economic mechanism of Meltzer and Richard (1981) and to account for the "paradox of redistribution".

References

[1] Alesina, A. and G.-M. Angeletos, (2005), "Fairness and Redistribution." *American Economic Review,* 95: 960–980.

[2] Alesina, A. and E. Glaeser, (2004), *Fighting Poverty in the US and Europe: A World of Difference.* Oxford, UK: Oxford University Press.

[3] Alesina, A. and E. La Ferrara, (2005), "Preferences for Redistribution in the Land of Opportunities." *Journal of Public Economics,* 89: 897–931

[4] Allard, G. and P. Lindert, (2006), "Euro-Productivity and Euro-Job Since the 1960s: Which Institutions Really Mattered?", NBER Working Paper No. 12460.

[5] Atkinson, A. and A. Brandolini, (2001), "Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries As a Case Study." *Journal of Economic Literature,* 39: 771–799.

[6] 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: 277–297.

[7] Beck, N. and J. Katz, (1995), "What to Do (And Not to Do) With Time Series Cross Section Data." *American Political Science Review,* 89: 634–647.

[8] Becker, G., (1983), "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics,* 98(3): 371–400.

[9] Becker, G. and C. Mulligan, (2003), "Deadweight Costs and the Size of Government." *Journal of Law and Economics,* 46: 293–340.

[10] Benabou, R., (1996), "Inequality and Growth." in eds. B. Bernanke and J. Rotemberg, *NBER Macroeconomics Annual,* 11–74.

[11] Benabou, R. and E. Ok, (2001), "Social Mobility and the Demand for Redistribution: The POUM Hypothesis." *Quarterly Journal of Economics,* 116, 447–487.

[12] Benabou, R. and J. Tirole, (2006), "Belief in a Just World and Redistributive Politics." *Quarterly Journal of Economics,* 121: 699–746.

[13] Blundell, R. and S. Bond, (1998),"Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics,* 87: 115–143.

[14] Bobba, M. and D. Coviello, (2007), "Weak Instrument and Weak Identification, in Estimating the Effects of Education, on Democracy." *Economic Letters,* 96: 301–306.

[15] Campante, F., (2007), "Redistribution in a Model of Voting and Campaign Contributions." mimeo, Harvard University.

[16] Deininger, K. and L. Squire, (1996), "A New Dataset Measuring Income Inequality." *The World Bank Economic Review,* 10: 565–591.

[17] Easterly, W. and R. Levine, (1997), "Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics,* 111: 1203–1250.

[18] Grossman, G. and E. Helpman, (1996), "Electoral Competition and Special Interest Politics." *Review of Economic Studies,* 63: 265–286.

[19] Hellwig, M., (1986), "The Optimal Linear Tax Revisited." *Journal of Public Economics,* 31: 163–179.

[20] Heston A., R. Summers and B. Aten, (2006), Penn World Table Version 6.2, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

[21] Huber E., C. Ragin, J. Stephens, D. Brady and J. Beckfield, (2004), Comparative Welfare States Dataset, Northwestern University, University of North Carolina, Duke University and Indiana University.

[22] 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: 165–181.

[23] Kenworthy, L., (2001), "Wage-Setting Measures: A Survey and Assessment." *World Politics,* 54: 57–98.

[24] Kristov, L., P. Lindert and R. McClelland, (1992), "Pressure Groups and Redistribution." *Journal of Public Economics,* 48: 135–163.

[25] Liebman, J., (2001), "Redistribution in the Current U.S. Social Security System." NBER Working Paper No. 8625.

[26] Lind, J., (2005), "Why is There so Little Redistribution?" *Nordic Journal of Political Economy,* 31: 111–125.

[27] Lind, J., (2007), "Fractionalization and the Size of Government." *Journal of Public Economics,* 91: 51–76.

[28] McCarty, N., K. Poole and H. Rosenthal, (2006), *Polarized America: The Dance of Ideology and Unequal Riches.* MIT Press.

[29] Meltzer, A. and S. Richard, (1981), "A Rational Theory of the Size of the Government. *Journal of Political Economy,* 89: 914–927.

[30] 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.

[31] Mirrlees, J., (1971), "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies,* 38: 175–208.

[32] Moene, K. and M. Wallerstein, (2001), "Inequality, Social Insurance and Redistribution." *American Political Science Review,* 95: 859–874.

[33] Peltzman, S., (1980),"The Growth of Government." *Journal of Law and Economics,* 23(2): 209–287.

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

[35] Persson, T. and G. Tabellini, (2003), *The Economic Effects of Constitutions.* Cambridge, Mass: MIT Press.

[36] Piketty, T., 1995, "Social Mobility and Redistributive Politics." *Quarterly Journal of Economics,* 110(3): 551–84.

[37] Roberts, K., (1977), "Voting over Income Tax Schedules." *Journal of Public Economics,* 8: 329–340.

[38] Rodrik, D., (1998), "Why Do More Open Economies Have Bigger Governments?" *Journal of Political Economy,* 106(5): 997–1032.

[39] Rodriguez, F., (1999), "Does Distributional Skewness Lead to Redistribution? Evidence from the United States." *Economics and Politics,* 11: 171–199.

[40] Roeder, P., (2001), "Ethnolinguistic Fractionalization (*ELF*) Indices, 1961 and 1985". Downloaded 12/28/2006.

[41] Romer, T., (1975), "Individual Welfare, Majority Voting and the Properties of a Linear Income Tax." *Journal of Public Economics,* 4: 163–185.

[42] Roodman, D., (2006), "How to Do xtabond2: An Introduction to "Difference" and "System" GMM in STATA." Working Paper 103, Center for Global Development.

[43] Rosenstone, S. and J. Hansen, (1993), *Mobilization, Participation and Democracy in America.* New York, MacMillan.

[44] Shelton, C., (2007), "The Size and Composition of Government Expenditure." *Journal of Public Economics,* 91: 2230–2260.

[45] Tabellini, G., (2000), "A Positive Theory of Social Security." *Scandinavian Journal of Economics,* 102: 523–545.

[46] Tabellini, G., (2007), "Culture and Institutions: Economic Development in the Regions of Europe." IGIER mimeo.

[47] Wooldridge, J., (2002), *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Mass: MIT Press.

Appendix 1: Data Appendix

Country Level Data

Redistribution: This is the Total Gross Public Expenditure variable from OECD's database that we downloaded from the SOCX database. It includes expenditures on old age (pensions), survivors, incapacity related benefits, health, family, active labor market programs, unemployment, housing and other categories but not education. The data for the construction of redistribution net of health and pensions that we use in Table 2 is also available in the above link. Redistribution is expressed as a percentage of GDP. The variable "Tax Wedge" is taken from the Lindert-Allard (2006) dataset, and measures the sum of personal income tax and employee plus employer social security contributions together with any payroll tax less cash transfers, expressed as a percentage of labor costs

Distribution: Our distributional variables are from OECD's Labor Force Statistics database. We quote the description that OECD gives for the distributional variables: "This [...] contains gross earnings of full-time workers by earnings percentiles and mean earnings, in national currency units. The series are a mixture of hourly, daily, weekly, monthly, and annual earnings and are specified in the country notes. These data were first collected and used in the tables, charts, and analysis on earnings dispersion presented in various editions of the OECD Employment Outlook: 1993 (Chapter 5), 1996 (Chapter 3), 1997 (Chapter 2), 1998 (Chapter 2)".

Wage Coordination index: Taken from Kenworthy (2001). The coordination index takes values from 1 which denotes fragmented wage bargaining, a bargaining process confined mostly to large enterprises and plants, to 5 which denotes centralized bargaining by peak confederations or government imposition of wage schedules.

Old, Surplus, Voter Turnout, Left and Right Vote, Presidentialism and Single Member District Electoral System: These variables are taken from the LIS-CWS database (2004). Old measures the percentage of the population older than 65 years. The variable general government surplus is expressed as a fraction of GDP. Voter turnout refers to the percentage of the population that voted in the last elections. Left and Right Vote are expressed as fractions of total votes directed towards leftist and rightist parties respectively in the last elections. For the interaction of our distributional variables with institutions, we used two variables. The variable Presidentialism is a binary dummy and the Single Member District variable takes the value 0 under proportional representation, 1 under modified proportional representation and 2 under plurality systems. See the LIS-CWS database (2004) for the classification of the parties into left, central and right and the definition of the constitutions variables.

Openness, Growth, and GDP: These variables are taken from the Penn Tables, Version 6.2 of Heston, Summers and Aten (2006). Openness is defined as exports plus imports as a percent of GDP. Growth is the growth rate of real GDP per capita in constant 2000 prices. GDP is real GDP per capita in 1996 International Dollars (Laspeyers).

Standardized Unemployment Rate and Employment Ratio: The SUR variable is taken from OECD's database on Labor Market Statistics and is expressed in percentage points. The Employment Ratio is defined as the proportion of an economys working-age population that is employed and taken from the OECD's LFS database. It is also expressed in percentage points.

ELF: The ethnolinguistic fractionalization index is taken from Roeder (2001). The ELF index is defined as one minus the probability that two randomly chosen persons from a population belong to the same ethnic, linguistic or racial group. A higher ELF index denotes a more heterogeneous population.

Deadweight Loss: Constructed by dividing the top corporate marginal tax rate (taken from the World Tax Database) by government's share on GDP (from the Penn Tables, Version 6.2). A higher value for DWL denotes higher distortions per unit of government activity.

Individual Level Data: Taken from the World Values Survey

Support for Redistribution: The first question ($E035$) asks respondents to choose a number from 1 if they believe that "Incomes should be made more equal" to 10 if "We need larger income differences as incentives". The second question ($E146$) asks to rank from "Very Important" to "Not at all Important" the statement "How important it is to eliminate big income inequalities between citizens?". In both cases we recode the variables to take higher values for more support for redistribution. MOREEQ takes the value 1 if $11 - E035 > 5$ and 0 otherwise. ELIMINEQ takes the value 1 if $6 - E036 > 3$ and 0 otherwise. Our results are not sensitive to the construction of the cutoff.

Political Participation: We recode all variables such that higher values denote more political participation. Some variables in the WVS are binary and we recode non-binary variables into a binary form to ease the comparison. The 14 variables that we use are: Interest in Politics ($E023$; 1 if $E023 = 1, 2$), Participation Local Political Acts ($A069$; binary), Belong to Political Party ($A068$; binary), Join Boycotts ($E026$; 1 if $E026 = 1, 2$), Sign Petitions ($E025$; 1 if $E025 = 1, 2$), Participation in Lawful Demonstration ($E027$; 1 if $E027 = 1, 2$), Adherence to Unofficial Strike ($E028$; 1 if $E028 = 1, 2$), Politics Important in Life ($A004$; 1 if $A004 = 1, 2$), Discussion of Political Matters with Friends ($A062$; 1 if $A062 = 1, 2$), Unpaid Work for Political Parties ($A085$; binary), Unpaid Work for Local Political Acts ($A086$; binary), Unpaid Work for Labor Unions ($A084$; binary), Belong to Labor Union ($A067$;

binary), and Active in Labor Union ($A101$; 1 if $A101 = 2$).

Income and Other Socioeconomic Controls: Each respondent is placed in a country-specific income decile ($X047$). Female ($X001$) takes the value 1 if the respondent is female and 0 otherwise, Age ($X003$) measures the age of the respondent, Children ($X011$) take the value 1 if the respondent has children, Married ($X007$) takes the value 1 if the individual is married. High School ($X025$) takes the value 1 if the respondent has completed at maximum the equivalent to High School in America. College ($X025$) takes the value 1 if the individual has completed at least college education.

Appendix 2: Micro-Evidence on Income, Preferences for Redistribution and Political Participation

We use the World Value Survey (WVS) dataset to study the relationship between income, attitude for redistributive policies and political participation. We restrict our micro-sample to the sample of developed democratic OECD countries that we also use in the macro regressions. We use the Four-Wave Integrated Data File (2006), which encompasses data from 1981 to 2004. The data allows us to conduct panel estimations and hold constant country specific and time effects. Table 6 presents summary statistics and the Data Appendix contains details for our variables and links to the original sources.

Preferences for Redistribution

We measure support for redistribution with two variables. The first variable codes answers that range from 1 if respondents think that "Incomes should be made more equal" to 10 if they believe that "We need larger income differences as incentives". The second variable ranges from 1 when the respondent believes that it is "Very important to eliminate big income differences" to 5 if it is "Not at all important". To ease the interpretation we recode the variables in binary form and 1 means more support for redistribution and 0 less support for redistribution. Our results are not sensitive to the recoding procedure that we followed. We name the two variables MOREEQ and ELIMINEQ respectively.

The survey codes the income of the respondent in deciles, which are standardized at the country level. A higher value for the income variable means that the respondent belongs to a higher decile in the income distribution. We control for gender, age, the presence of children, marital status and education level. In our baseline specification, we assume that the support for redistribution of individual i living in country c at time t can be characterized by a latent variable

$$R^*_{ict} = \beta Y_{ict} + \gamma X_{ict} + c_i + \tau_t + \varepsilon_{ict} \tag{2}$$

where Y_{ict} denotes the income decile of the respondents, X_{ict} is the vector of controls, c_i and τ_t are country and year dummies and ε_{ict} is the error term. We do not observe R^*_{ict}, but a variable R_{ict} taking values 0 for low support and 1 for high support for redistribution. Assuming that the distribution of the error term is logistic, we estimate a logit model.

Table 7 shows how the support for redistribution declines with income. In the first two columns we use MOREEQ and ELIMINEQ respectively. Our point estimates suggest that the likelihood of finding someone who supports redistribution decreases by around 2.5 percentage points each time we move up one decile in the income distribution. For all countries in our sample, the marginal effect is significant at the 1% level, and its magnitude ranges from 1.5% to approximately 4%. In the last two columns of Table 7, we use MOREEQ, the dependent variable with the largest number of observations, and repeat the exercise in the restricted sample of Anglo Saxon countries in column 3 and Continental Europe countries in Column 4. We find that there is no specific group of countries that drives the results. We have also restricted the sample for specific time periods and also estimated ordered logit models without difference in the results.

Political Participation

We find 14 variables that measure participation in political activities. These are (i) Interest in Politics, (ii) Participation in Local Political Acts, (iii) Belong to Political Party, (iv) Join Boycotts, (v) Sign Petitions, (vi) Participation in Lawful Demonstration, (vii) Strikes, (viii) Politics Important in Life, (ix) Discussion of Political Matters with Friends, (x) Unpaid Work for Political Parties, (xi) Unpaid Work for Local Political Acts, (xii) Unpaid Work for Labor Unions, (xiii) Belong to Labor Union, (xiv) Active in Labor Union. Variables are recoded in a binary form such that 1 means higher political participation.

In our baseline specification, we assume that the level of political participation of individual i, living in country c at time t can be characterized by a latent variable

$$P^*_{ict} = \beta Y_{ict} + \gamma X_{ict} + c_i + \tau_t + \varepsilon_{ict} \tag{3}$$

where Y_{ict} denotes the income decile of the respondents, X_{ict} denotes the same vector of controls as in the previous Section, c_i and τ_t are country and year dummies and ε_{ict} is the error term. We do not observe P^*_{ict}, but instead a variable P_{ict} taking values 0 for low and 1 for high political participation.

We find strong evidence that multiple dimensions of political participation, ranging from discussing politics with friends to participating in local political acts, rise with income. The point estimates for all 14 variables are positive and with the exception of the coefficient for Joining an Unofficial Strike, they are all significant at 1% level. These results are robust when restricting to a limited number of countries and time periods. Our point estimates indicate that the likelihood of participating politically each time we move up one decile in the income distribution increases by 0.5 to 2 percentage points, depending on the variable. In Table 8, we present results for the five variables with the highest number of observations.

Table 1: Summary Statistics of Macro Level Data

Variable	Mean	Std. Dev.	Min	Max	Observations
Social Public Expenditure (RED)	21.59	5.78	10.20	35.10	118
RED net of Health and Pensions	8.75	3.70	2.40	17.90	118
Pensions	6.99	2.39	2.70	11.50	118
Health	5.85	.96	3.70	8.50	118
Tax Wedge	33.54	10.17	14.10	48.50	120
y^{90}/y^{mean}	1.57	0.11	1.32	1.89	90
y^{50}/y^{mean}	0.88	0.04	0.71	0.93	90
y^{10}/y^{mean}	0.54	0.09	0.30	0.70	90
GDP	19608	3984	9276	33905	135
Growth	2.11	1.79	-5.4	8.9	135
Share of Elderly Population	13.28	2.34	8.10	18.00	135
Deadweight Loss	2.28	.56	1.01	3.34	131
Voter Turnout	79.16	11.45	49.10	95.80	134
Openness	62.67	33.66	16.30	174.00	135
Unemployment Rate	7.22	3.21	0.30	16.40	127
Employment Ratio	65.27	6.91	51.80	80.50	120
Left Vote	35.10	14.09	0	56.00	135
Coordination Index	3.11	1.40	1	5	135

Notes to Table 1: These are summary statistics for the 3 year averaged sample. See the Appendix for the definition and the source of each variable.

Table 2: Baseline Specification

Dependent Variable	*RED*	*RED* net of Health and Pensions	Pensions	Health	Tax Wedge
	(1)	(2)	(3)	(4)	(5)
$y^{90}/\bar{y}$	-1.06	-1.71	-1.13	.29	-1.26
	(.23)***	(.52)***	(.45)**	(.42)	(.34)***
$y^{50}/\bar{y}$	-1.44	-1.22	-.94	-1.62	-.44
	(.63)**	(.64)*	(1.09)	(1.20)	(.63)
$y^{10}/\bar{y}$	1.01	1.59	1.47	-.07	1.02
	(.26)***	(.43)***	(.45)***	(.41)	(.37)**
GDP	-.07	-1.37	.78	.81	.29
	(.40)	(.73)*	(.55)	(.49)	(.37)
Growth	-.03	-.05	-.01	-.01	.00
	(.01)***	(.01)***	(.01)*	(.01)	(.01)
Share of Elderly	.31	.28	.53	-.07	.42
	(.13)**	(.25)	(.21)**	(.12)	(.18)**
DWL	-.15	-.03	-.18	-.21	-.17
	(.09)	(.18)	(.06)***	(.10)	(.04)***
R-squared ("Within")	.70	.54	.66	.64	.55
Observations	82	82	82	82	82

Notes to Table 2: All specifications include year and country specific fixed effects. Standard errors are clustered by country. Independent variables are defined in the text and the Appendix. Standard Errors are displayed in parentheses. *** denotes significance at 1%, ** significance at 5%, * significance at 10%. Approximate t-statistics used.

Table 3: Other Time Varying Controls

Dep. Var.: *RED*	(1)	(2)	(3)	(4)	(5)	(6)
$y^{90}/\bar{y}$	-1.11	-1.04	-.98	-1.15	-1.12	-1.24
	(.26)***	(.27)***	(.27)	(.21)***	(.22)***	(.45)**
$y^{50}/\bar{y}$	-1.36	-1.19	-1.22	-1.40	-1.42	-.73
	(.56)**	(.67)*	(.63)*	(.65)**	(.58)**	(.71)
$y^{10}/\bar{y}$	.83	.96	1.03	1.12	1.02	1.00
	(.19)***	(.30)***	(.34)**	(.29)***	(.27)***	(.33)**
GDP	.21	.11	-.14	-.08	-.04	.28
	(.61)	(.27)	(.38)	(.39)	(.40)	(.56)
Growth	-.03	-.02	-.03	-.03	-.03	-.03
	(.01)***	(.01)***	(.01)***	(.01)***	(.01)***	(.01)***
Share of Elderly	.26	.40	.35	.31	.32	.31
	(.13)*	(.23)	(.15)**	(.15)*	(.14)**	(.26)
DWL	-.10	-.17	-.14	-.14	-.13	-.08
	(.12)	(.09)*	(.10)	.10	(.10)	(.16)
Unemp. Rate	.09					.10
	(.11)					(.12)
Openness		.14				.09
		(.19)				(.29)
Left Vote			.02			-.05
			(.09)			(.08)
Voter Turnout				.31		.42
				(.41)		(.38)
Coordination					-.01	-.01
					(.01)	(.01)
R-squared ("Within")	.72	.71	.70	.70	.70	.73
Observations	82	82	74	81	82	73

Notes to Table 3: Dependent Variable is Total Social Public Expenditures in all columns. All specifications include year and country specific effects. Standard errors are clustered by countries. Independent variables are defined in the text and the Appendix. Standard Errors are displayed in parentheses. *** denotes significance at 1%, ** significance at 5%, * significance at 10%. Approximate t-statistics used.

Table 4: The Structure of Panel Errors and Country Specific Trends

	FGLS	FGLS	FGLS	PCSE	PCSE	PCSE	HAC	HAC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$y^{90}/\bar{y}$	-.71	-.77	-.76	-1.17	-1.27	-1.23	-1.17	-1.17
	(.32)**	(.33)**	(.29)***	(.39)***	(.40)***	(.37)***	(.64)*	(.68)*
$y^{50}/\bar{y}$	-.39	-.33	-.25	.08	.12	.36	.08	.08
	(.46)	(.46)	(.44)	(.57)	(.57)	(.57)	(.69)	(.70)
$y^{10}/\bar{y}$	.69	.70	.70	.56	.59	.49	.56	.56
	(.20)***	(.20)***	(.20)***	(.26)**	(.25)**	(.26)*	(.30)*	(.30)**
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country Specific Trends	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel Heteroscedasticity	Yes	Yes	Yes	Yes	Yes	Yes		
Autocorrelation	No	Common AR(1)	Panel AR(1)	No	Common AR(1)	Panel AR(1)		
Standard Errors							HAC(1)	HAC(2)
Observations	82	80	80	82	82	82	82	82

Notes to Table 4: Dependent variable is log of Total Public Social Expenditure. Independent variables are defined in the text and the Appendix. Standard Errors are displayed in parentheses. *** denotes significance at 1%, ** significance at 5%, * significance at 10%. Approximate t-statistics used. Common AR(1) is .12 in column 2 and .11 in column 5. All regressions include fixed effects, time effects, a country specific trend and the controls of the first column of Table 2. FGLS denotes feasible generalized least squares estimation, PCSE denotes panel-corrected standard errors and HAC heteroscedasticity and autocorrelation robust standard errors.

Table 5: Dynamic Regressions

	POLS	FE	SGMM	DGMM	SGMM	DGMM
	(1)	(2)	(3)	(4)	(5)	(6)
RED_{t-1}	.88	.80	.94	.61	.91	.74
	(.03)***	(.10)***	(.05)***	(.23)***	(.04)***	(.19)***
$y^{90}/\bar{y}$	.06	-.01	.15	-.02	.31	.01
	(.10)	(.33)	(.27)	(.42)	(.15)**	(.38)
$y^{50}/\bar{y}$	-.45	-.11	-.72	-3.44	-.55	-1.42
	(.22)**	(.51)	(.38)*	(1.00)***	(.27)**	(.67)**
$y^{10}/\bar{y}$	.15	.19	.28	.96	.15	.37
	(.09)*	(.26)	(.13)**	(.66)	(.08)*	(.46)
p-value			.36	.16	.41	.21
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Lag Instruments			All but Inequality	All but Inequality	All	All
Observations	77	77	77	63	77	63

Notes to Table 5: Dependent variable is log of Total Public Social Expenditure. Independent variables are defined in the text and the Appendix. POLS stands for Panel Ordinary Least Squares, FE for fixed effects estimation, DGMM for difference Generalized Method of Moments as in Arellano and Bond (1991), and SGMM for System Generalized Method of Moments as in Blundell and Bond (1998). All specifications include the controls of the first column of Table 2. Standard Errors are displayed in parentheses. *** denotes significance at 1%, ** significance at 5%, * significance at 10%. Approximate t-statistics used. The row p-value refers to the null of the hypothesis "H_0: No Third Order Autocorrelation in the Original Residuals". The first column does not include country dummies, while the second includes. In all other columns country effects are differenced out. The matrix of instruments is "collapsed" (see Roodman, 2006). When Lag Instruments indicate 'All', deep lags of all variables are used as instruments, when it indicates 'All but Inequality', then the three ratios that refer to the income distribution are not used as instruments. Third and higher order lags are used as instruments.

Table 6: Summary Statistics Individual Level Data

Variable	Mean	Std. Dev.	Min	Max	Observations
MOREEQ	.44	.50	0	1	37434
ELIMINEQ	.60	.49	0	1	10215
Interest in Politics	.54	.50	0	1	40372
Discuss with Friends	.15	.36	0	1	50077
Politics Important	.44	.50	0	1	39822
Demonstrations	.22	.42	0	1	46551
Sign Petitions	.63	.48	0	1	48623
Income	5.31	2.68	1	10	41915
Female	.53	.50	0	1	50418
Married	.61	.49	0	1	50257
Age	44.92	16.93	15	97	50273
Children	.74	.44	0	1	50491
High School	.75	.43	0	1	24253
College	.14	.35	0	1	24253

Notes to Table 6: These are summary statistics for the variables taken from the World Values Survey. See the Appendix for the definition of each variable.

Table 7: Support for Redistribution

	(1)	(2)	(3)	(4)
Income	-.02	-.02	-.02	-.02
	(.00)***	(.00)***	(.00)***	(.00)***
Female	.05	.03	.04	.08
	(.01)***	(.01)***	(.01)***	(.01)***
Married	-.02	-.03	-.00	-.07
	(.01)***	(.01)**	(.01)	(.02)***
Age	.00	.00	-.00	.00
	(.00)	(.00)***	(.00)	(.00)***
Children	.00	.03	-.02	.02
	(.01)	(.02)*	(.01)*	(.02)
No High School	.04	.03	.04	.01
	(.01)***	(.02)	(.02)***	(.03)
College	-.02	.01	.04	-.08
	(.02)	(.03)	(.02)*	(.03)**
Time Effects	Yes	Yes	Yes	Yes
Country Effects	Yes	Yes	Yes	Yes
Standard Errors	Clustered	Clustered	Clustered	Clustered
Observations	17837	8213	8924	5205

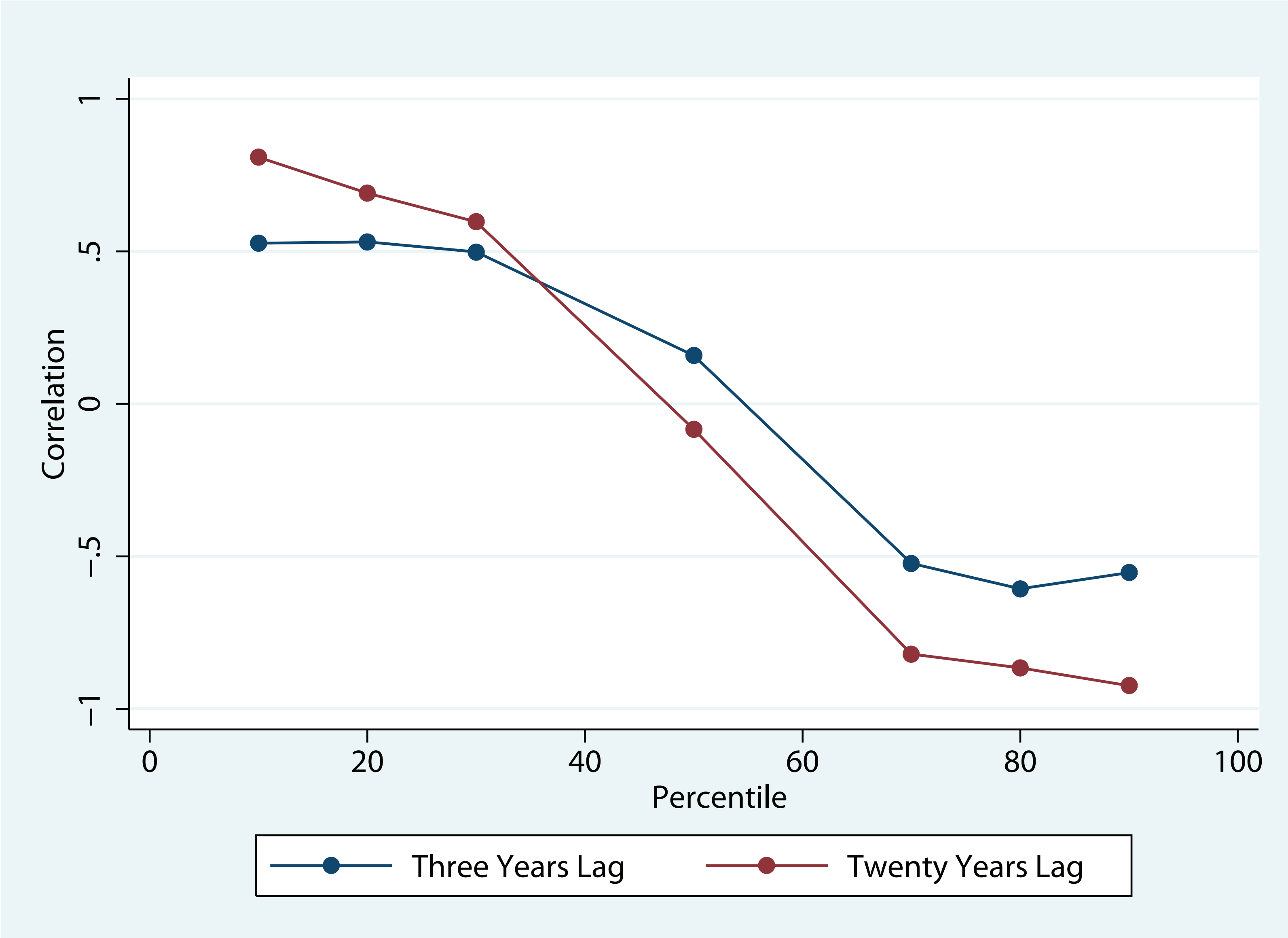
Notes to Table 7: Dependent variable is MOREEQ in columns 1, 3 and 4 and ELIMINEQ in column 2. Column 3 displays the same regression as in column 1 for a restricted sample of Anglo Saxon countries and column 4, for Continental Europe countries. Independent variables are defined in the text and the Appendix. Standard Errors are displayed in parentheses. Clustered errors are at the town size level. *** denotes significance at 1%, ** significance at 5%, * significance at 10%.

Table 8: Political Participation

	(1)	(2)	(3)	(4)	(5)
Income	.02	.01	.01	.01	.02
	(.00)***	(.00)***	(.00)***	(.00)***	(.00)***
Female	-.11	.-05	-.03	-.06	.03
	(.01)***	(.00)***	(.01)***	(.01)***	(.00)***
Married	.02	-.00	.03	-.03	.01
	(.01)**	(.00)	(.01)***	(.01)***	(.01)
Age	.00	.00	.00	-.00	-.00
	(.00)***	(.00)***	(.00)***	(.00)**	(.00)***
Children	-.04	-.02	-.04	.01	.03
	(.01)***	(.01)**	(.01)***	(.01)	(.01)***
No High School	-.19	-.06	-.13	-.13	-.13
	(.01)***	(.01)***	(.01)***	(.01)***	(.01)***
College	.03	.04	.01	.05	.02
	(.02)	(.01)***	(.04)	(.01)***	(.02)
Time Effects	Yes	Yes	Yes	Yes	Yes
Country Effects	Yes	Yes	Yes	Yes	Yes
Observations	18959	20520	19719	19165	20294

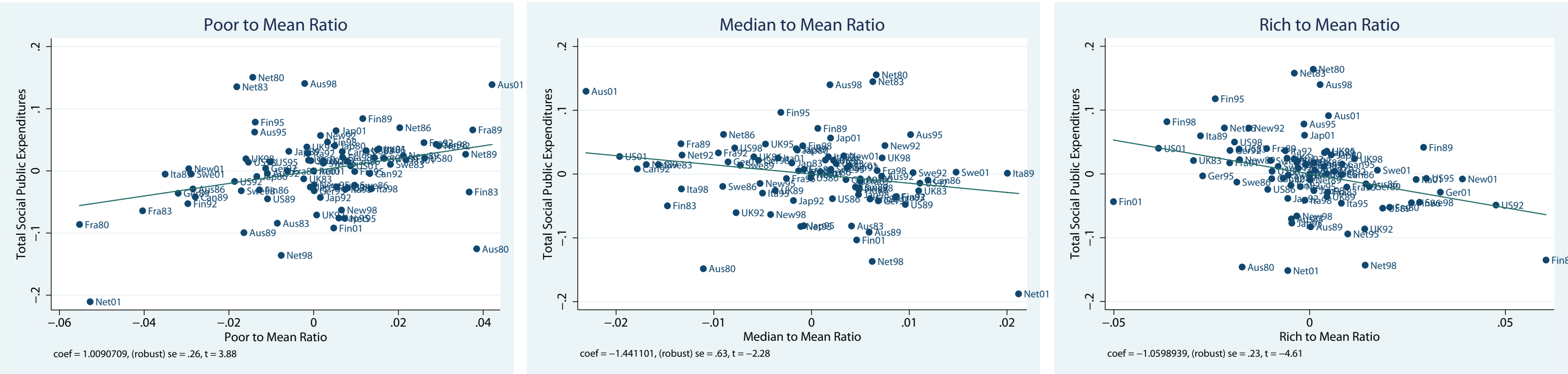
Notes to Table 8: Dependent variable for each column is (1) Interest in Politics, (2) Discuss Politics with Friends, (3) Politics Important in Life, (4) Attend Lawful Demonstrations, (5) Sign Petitions. Independent variables are defined in the text and the Appendix. Standard Errors are displayed in parentheses. Clustered errors are at the town size level. *** denotes significance at 1%, ** significance at 5%, * significance at 10%.

Figure 1: Correlations of Redistribution with Lagged Percentile to Mean Ratios of Gross Earnings



Notes: The horizontal axis measures the percentile of the income distribution. The vertical axis shows the unconditional correlation between redistribution and the gross earnings of every percentile relative to mean gross earnings. For instance, the correlation between redistribution and the 10/mean ratio of gross earnings is about .50, when the latter is measured with a three year lag. The correlation is calculated for the pooled sample. Correlations for the 40th and the 60th percentile are not displayed because of the scarcity of this data.

Figure 2: Redistribution and Poor, Middle, Rich Income



Notes: The figure shows the residuals from the conditional relationship between redistribution and the log of the ratios $y^{10}/\bar{y}$, $y^{50}/\bar{y}$ and $y^{90}/\bar{y}$ as estimated in the first column of Table 2.