Income, Risk and Voting: Demand for Redistribution and Left Parties in Industrialized Democracies*

David Rueda†
Daniel Stegmueller‡

This version: September, 2019

Abstract

A significant literature in political economy has recently focused on the relationship between income and risk and redistribution preferences. However, it remains unclear whether the interplay between material position and preferences has any influence on political behavior. In this paper we argue that redistribution preferences are indeed a most significant channel shaping vote choice. We test our theoretical claims with data from Western Europe and the US and show that voting for redistributive parties is highly dependent on individual levels of demand for redistribution. The poor and those exposed to more risk are more supportive of redistribution and, in turn, these redistribution preferences make them more likely to vote for redistributive parties. Our analysis goes beyond previous research by explicitly studying this preference mechanism. We disentangle the direct and indirect effects of income and risk (as well as other factors) to obtain estimates of their effects on voting through preferences.

*Rueda and Stegmueller's research was supported by the National Research Foundation of Korea (NRF-2017S1A3A2066657)

†University of Oxford, `david.rueda@politics.ox.ac.uk`
‡Duke University, `daniel.stegmueller@duke.edu`

I. Introduction

Most analysts would agree that an individual's relative income (i.e., whether she is rich or poor) and her exposure to risk (i.e., whether she will be rich or poor tomorrow) affect her redistribution preferences. But why should we care about redistribution preferences in the first place? We argue that the (often implicit) model behind much of comparative politics and political economy starts with redistribution preferences. These redistribution preferences affect how individuals behave politically and their behavior in turn affects the strategies of political parties and the policies of governments. In this paper, we will focus on perhaps the most momentous potential consequence of redistribution preferences: voting.

Inequality and redistribution have seen a resurgence in academic interest in recent times. This is particularly the case in the US, where Bartels (2009) has shown the spectacular increase in inequality over the past 35 years to be the product of policy choices in a political system dominated by partisanship and particularly receptive to the preferences of the wealthy. Hacker and Pierson (2011) coincide not only in the appreciation of the attention that policy-makers pay to the rich but also about the fact that politics is the main factor behind inequality ("American politics did it").

The connection between inequality and political behavior, however, remains unclear. A number of observers would deny that income and inequality are significant determinants of voting.[1] Some analysts would agree that an individual's present and future income affects her political behavior,[2] but they would not necessarily agree on the reasons why this is the case. This paper's analysis addresses one of the implications of most arguments about the importance of economic circumstances to political outcomes. If income and risk matter to individual political behavior, it seems reasonable to assume that they do so through their influence on preferences for redistribution. These redistribution preferences may (or may not) then be reflected on party positions and, eventually, government policy.

While a voluminous political economy literature has emerged on the influence of income and risk on preferences, we know much less about whether these preferences do in fact affect political behavior at all. Most political economy arguments start from the assumption that an individual's position in the income distribution determines her preferences for redistribution. The most popular version of this approach is the theoretical model proposed by Romer (1975) and developed by Meltzer and Richard (1981). And there is some evidence supporting the argument that relative income and risk influence preferences for redistribution. A relative income effect is found in the US by, among others, Gilens (2005), McCarty et al. (2008), and Page and Jacobs (2009). Using comparative data, Bean and Papadakis (1998), Finseraas (2009), and Shayo (2009) (again, among others) find similar effects. But the idea that material self-

[1] See, for example, Green et al. (2004) and Lewis-Beck (2009) or, more recently, Achen and Bartels (2016).

[2] There is an influential literature in political science on how pocketbook issues and class (both closely related to income) influence voting. See Downs (1957), Key (1966) or Fiorina (1981) on pocketbook issues and Lipset (1983), Evans (1999) or Brooks and Manza (1997) on class.

interest determines redistribution preferences should not be simply limited to a measure of present income, but should also include risks affecting future income. Exposure to risk, whether measured as occupational unemployment or as skill specificity, is shown to affect preferences for redistribution (as insurance) by Iversen and Soskice (2001) and Rehm (2009, 2016).

Do these preferences translate into political behavior? As we will show below, redistribution preferences (which are in turn affected by relative income and exposure to risk) are an important factor affecting voting in both Western Europe and the US. Our approach adds to those prevalent in the literature in three important ways. First, we provide an argument and convincing evidence that income and risk are in fact significant determinants of voting. Second, most political economy models link individual income (or exposure to risk) to policy outcomes making the essential assumption that there is a relationship between preferences and voting. We specify explicitly the theoretical mechanisms that determine preferences and party choice, and test them empirically. Third, much of the recent debate about the lack of redistributive policies in industrialized democracies has centered around the perception that second-dimension issues are disproportionately important to the poor. Perhaps the most well-known example of this is the contention that cultural, religious and social values outweigh economic concerns for the American working class in some states (see Frank 2004 and, more recently, Hersh and Nall 2016). The implication of these arguments is that the solution to the puzzle affecting (the lack of) redistribution in industrialized democracies concerns demand. We show in this paper that this may not be the case. We find the poor and those exposed to risk to be uniformly in favor of redistribution and therefore uniformly more likely to vote for redistributive parties. The puzzle of redistribution may have more to do with supply (such as party platforms or the effects of electoral institutions) than with demand.

II. Argument

Our theoretical argument proceeds in two stages. First, we address the formation of preferences for redistribution, explaining why income and risk are important determinants of demand for redistribution. Second, we detail the influence of redistribution preferences on voting choices. We argue that those who are supportive of redistribution will be more likely to vote for redistributive parties.

II.A. Inequality and redistribution preferences

The first step in our argument involves the relationship between individual levels of income and redistribution preferences. Political economy approaches that start from the assumption that an individual's position in the income distribution determines her preferences for redistribution are often inspired by the theoretical model proposed by Romer (1975) and developed by Meltzer and Richard (1981). To recapitulate very briefly, the RMR model assumes that the preferences of the median voter determine government policy and that the median voter seeks to maximize current income. If there are no deadweight costs to redistribution, all voters with incomes below the mean maximize their utility by imposing a 100% tax rate. Conversely, all voters with incomes above the mean prefer a tax rate of zero. When there are distortionary costs to taxation, the RMR model implies that, by increasing the distance between the median and the mean incomes, more inequality should be associated with more redistribution.

While it is the case that the rich support redistribution less than the poor almost everywhere, the strength of this relationship is hardly consistent (Dion 2010; Dion and Birchfield 2010; Beramendi and Rehm 2016). A reason for this is that, as mentioned above, the material self-interested factors affecting redistribution preferences should not be limited to a measure of present income. If material self-interest is defined inter-temporally, the more direct effects of contemporary relative income (as in Romer 1975 and Meltzer and Richard 1981) should be complemented by arguments about about social insurance and risk (as in Sinn 1995; Moene and Wallerstein 2003; Iversen and Soskice 2001; Rehm 2009; Mares 2003), and about social mobility and life-cycle profiles (Rueda and Stegmueller 2017, 2019; Alesina and Giuliano 2011; Haider and Solon 2006; Benabou and Ok 2001).

Arguments about the importance of insurance are most relevant to our focus in this paper. They emphasize the importance of risk in determining redistribution and insurance preferences. In this vein, Rehm (2009, 2016) argues that, while income captures redistribution preferences, occupation characteristics capture risk exposure and insurance motivations. In a highly influential article, Iversen and Soskice (2001) argue that exposure to risk is inversely related to the portability of individual skills. While we agree with Iversen and Soskice that individual expected utility (across a range of possible labor market stages) is a key factor in determining redistribution preferences, we do not highlight the difference between general and specific skills here. Instead, we will use Rehm's occupational unemployment as our proxy for exposure to risk.

As in the Meltzer-Richard model, our argument implies that a rise in income will reduce the demand for distribution. It also implies that the immediate pocketbook consequences of inequality are fully contained in the individual income distance changes produced by this inequality shift. In other words, the tax and transfer consequences of inequality (and their effects on individual demands for redistribution) are picked up by individual income changes. As in Rehm (2009, 2016), we then argue that while income captures redistribution preferences, occupational unemployment captures risk exposure and insurance motivations. The higher the risk an individual is exposed to, the more supportive of redistribution she will be.

II.B. Redistribution preferences and vote choice

In the second stage of our argument, we argue for the relevance of redistribution preferences to voting.[3] We therefore follow a well-established literature on the relationship between economic considerations and political behavior. As mentioned above, most political economy arguments start from the assumption that an individual's redistribution preferences affect her political choices (see Romer 1975 and Meltzer and Richard 1981). The literatures on economic voting and class voting are based on similar arguments. Like authors in the economic voting tradition (e.g., Duch and Stevenson 2008), our argument posits that there is a relationship between an individual's economic interests and her likelihood to vote for a particular party. Class voting analyses (e.g., Evans and de Graaf 2013 and Evans 1999) emphasize the effects of socio-economic cleavages on political preferences, but their focus on occupational factors is largely compatible with our arguments. Our approach is also related to a recent literature that emphasizes risks and skills as determinants of preferences. While this literature associates unemployment vulnerability with skill profiles (e.g, Cusack et al. 2006), we highlight the direct effects of redistribution preferences (regardless of skills).

Like the traditional economic voting literature (Downs 1957) we conceive of voters as instrumental rational actors. Individuals will vote following a comparison of what they gain or lose from the policies proposed by each party. In the words of Duch and Stevenson, we assume that "voters rationally derive expected utilities for competing political parties and that these determine their vote choice" (2008: 9). As in the pioneering work of Kramer (1971) and Fair (1978), we consider that economic well-being (and therefore redistribution and insurance) is a significant factor affecting a voter's utility function.

A substantial literature debates the issue of how economic considerations enter a citizen's vote choice function. Two main approaches can be distinguished, one emphasizing *sanctioning* and the other focusing on *selection*. The *sanctioning model* is characterized by the consideration that voters are narrowly retrospective and mostly motivated by punishing or rewarding incum-

[3] As mentioned above, an important literature posits that, in some cases, the poor are diverted from the pursuit of their material self-interest. Perhaps the most well-known example of these arguments is the contention that second-dimension issues (particularly cultural and social ones) outweigh economic ones for the American working class. Frank (2004) and the critique in Bartels (2006) are good illustrations of this debate, but so is the emphasis on cosmopolitanism as a determinant of vote in Gelman et al. (2008). More comparatively, the important contribution in Shayo (2009) to the political economy of identity formation follows a similar logic. Shayo's theoretical model emphasizes two identity dimensions: economic class and nationality. As a result of status differences, the poor are more likely than the rich to identify with the nation rather than their class in high inequality countries. Because they take group interests into account, moreover, the poor who identify with the nation are less supportive of redistribution than the poor who identify with their class. While not denying that moral and cultural issues are important to voting Democrat in the US, we emphasize the importance of redistribution preferences. McCarty et al. find that income is an extraordinarily good predictor of partisanship and voting even among conservative Christians in the US (2008: 100-101). In the same vein, we will show below that the influence of income and risk through redistribution preferences are a powerful predictor of voting even when controlling for the influence of other channels of influence (such as values).

bents (see the classic works of Kramer 1971, Key 1966 and Fiorina 1981). Focusing on moral hazard, i.e., the risk of rent-seeking by incumbents if not punished for bad economic outcomes, Barro (1973) and Ferejohn (1986) also belong within this tradition. The *selection/competency model* argues that voters gather more information to assess the likely economic outcomes associated with competing political alternatives. Downs (1957) and Stigler (1973) are classical examples of this approach but we would argue that this is also the understanding of voting underlying Meltzer and Richard (1981) and subsequent political economy treatments of redistribution and voting (Persson and Tabellini 2000). While not incompatible with sanctioning, our argument more clearly implies a selection logic. We propose that individuals who are in favor of redistribution and insurance will identify the party more likely to promote equality and therefore be more likely to vote for it.

More specifically, in our analysis we consider voting to be a discrete choice. By this we mean a decision made over a set of exclusive and exhaustive choices (see Duch and Stevenson 2008: 39). Each voting choice (i.e., the parties a voter can select) offers some utility with regards to the voter's redistribution preferences. It is the contribution of these individual redistributive preferences to the voting choice that matters to the main focus of our paper, but our approach can be described in more general terms. Like Alvarez et al. (2000: 240), we assume that each individual obtains some utility from each party, and that the individual votes for the party offering the highest utility. The utility of each party is understood to be a function of a set of systematic components (specific to the voter, to the party and to the election) and a random disturbance. The parameters in these random utility models are often estimated with multinomial probit techniques using distance variables as the predictors. These variables reflect the spatial distance between a respondent's position on an issue (in our case, redistribution) and the respondent's view of each party's position on the same issue (for examples of this approach see Alvarez and Nagler 1995, 1998). In our analysis of American data, we use an explicit measure for redistributive distance (we provide the details below). In our analysis of European data, we lack information on the respondent's views of each party's position and we use party manifesto information on party positions instead.[4] In both cases, we explore individual vote choice as an unobserved vector of probabilities associated to the redistributive positions of different parties.

The intuition linking redistributive preferences to voting choice explained above is pretty straightforward, but it has arguably not received enough attention in the existing comparative political economy literature. This is also the case in the American politics literature (McCarty et al. 2008). Two clear illustrations of this are major works on partisan identification and voting by Green et al. (2004) and Lewis-Beck (2009). Both analyses underplay the importance of income (and, even more so, its connection to redistribution preferences). The more recent contribution by Achen and Bartels (2016) could also be added here and our paper challenges

[4]We sacrifice an explicit measure of spatial distance distance on redistribution here to maximize the coverage of countries and years in our analysis.

their argument that political loyalties (typically acquired in childhood) and economic myopia determine political behavior. To the extent that income and redistribution preferences are considered in this literature, it is through the prism of "class voting." But this approach is quite distinct from the political economy arguments that we present in this paper.

The equilibrium in most political economy models is achieved by individuals deriving their preferences over optimal fiscal policy based on their income position (or their occupational or labor market position), which are then "aggregated into an economywide policy via the collective choice mechanism in place" (Drazen 2000: 312). Thus, the two central concepts are citizens' redistribution preferences (or ideal points) and vote choices (the collective choice mechanism). The traditional modes of empirical analysis have then been (i) to explore the influence of income on voting and (ii) to relate income to economy-wide outcomes, such as spending (see, e.g., the summaries of empirical research in Persson and Tabellini 2000 and Mueller 2003). This, however, simply assumes that our central argument – the relationship between preferences and voting – is indeed the mechanism at work. This paper's contribution is to specify explicitly the theoretical mechanisms that determine preferences and party choice, and to test them empirically.[5]

II.C. *Defining the mediating role of preferences*

To make transparent how we think about these relationships conceptually, we explicitly state our hypothesized mechanism: redistribution preferences are a mediating variable (Pearl et al. 2017: 75) between income/risk and vote choice. The aim of our analysis is thus to evaluate the pathway linking income and risk to vote choice *through* preferences (termed the *indirect effect*) and separate it from other possible channels. Take, for example, income. As illustrated in Figure I, this amounts to separating the indirect path (*income* $\longrightarrow$ *preferences* $\longrightarrow$ *vote choice*) from the direct path (*income* $\longrightarrow$ *vote*), which represents non-preference relationships. The total effect of income on vote choice is given by the sum of both paths.

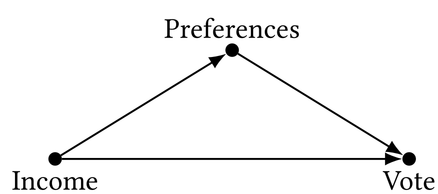

Figure I
Graph illustrating pathways between income (risk), preferences, and vote choice.

One should keep in mind that the terms developed in this subsection are independent

[5]The approach most similar to ours is that of Brooks and Brady (1999), who argue that income shapes voting behavior indirectly (by affecting evaluations of social welfare and government size).

of the specific statistical estimation strategy (to be implemented later). Mediation is not a 'method.' The existence of a mediating relationship is the property of a population. It is defined counterfactually, and thus studying it requires making counterfactual assumptions.

We use potential outcomes notation to express our causal quantities of interest. We hasten to add that using causally defined quantities does not automatically imply that resulting estimates are causal. The (many) limits of observational data analysis still apply.[6] Rather, in our view (and that of e.g, VanderWeele and Vansteelandt 2009 or Imai et al. 2010) the key benefit of clearly defining mechanisms in a potential outcomes framework is that it lays bare the identifying assumptions needed. We state these key assumptions explicitly and conduct robustness tests and sensitivity analyses to see how our results hold up if they are violated.

Let individual i ($i = 1, \ldots, N$) receive some level of income, w_i and face some level of risk/occupational unemployment, z_i. Our individual prefers a certain level of redistribution, which is a function of her income and occupational unemployment risk, which we write as $R_i(z_i, w_i|c_i)$. Possibly confounding variables (individual and contextual characteristics) are denoted by c_i. At election time she casts her vote based on her redistribution preferences and on a number of other factors. We write this vote function as $V_i(z_i, w_i, R_i(z_i, w_i|c_i)|c_i)$. Again, we condition on a set of possible confounders, c_i. Note that income and occupational unemployment risk appear twice: as factors changing preferences (which in turn shape vote choice) and as factors directly shaping vote choice (via possibly infinitely many other possible channels).

To understand the role of, for example, income, examine a (counterfactual) shift in income from w_i to w_i'. Holding everything else constant, the *total unit effect* of income on vote choice is given by (we omit possible confounders for clarity):

$$TE \equiv V_i(z_i, w_i, R_i(z_i, w_i)) - V_i(z_i, w_i', R_i(z_i, w_i')) \tag{1}$$

This is the expected difference in the probability of voting for a redistributive party as a result of changing income. It results from the combination of the systematic effects of changing preferences and all other factors, which are not relevant to our argument.

To trace *how* income shapes voting *via* preferences it is not enough to look at disparate sets of regression coefficients (of, say, income on preferences, and preferences on voting). Rather, we need to explicitly state our hypothesized mechanism. We define this *indirect effect* (Robins and Greenland 1992; Pearl 2001) as:

$$IE \equiv V_i(z_i, w_i, R_i(z_i, w_i)) - V_i(z_i, w_i, R_i(z_i, w_i')). \tag{2}$$

[6] But note that even if we had access to a randomized treatment of income or risk, the decomposition developed below has elements of "observational" inference that require careful analysis. As Keele (2015) puts it (in his aptly titled article), there is "no 'gold standard' method for the identification of causal mediation effects. In particular, mediation effects will always have the character of estimates from observational data since they are generally subject to a specific form of confounding."

This is the effect a change in income has on vote choice via redistribution preferences *only*. By fixing income and only changing preferences, we isolate our preference mechanism and eliminate the impact of competing mechanisms (Imai et al. 2011: 769). In other words, it is a formal (counterfactual) expression of our hypothesized income–preference nexus *net* of alternative channels (such as, for example, second dimension concerns).

The remaining effect of changes in income on vote choice not transmitted via preferences is termed the *direct effect* and represents how income affects vote choice in ways that are not considered in our model (i.e., all mechanisms other than redistribution preferences):

$$DE \equiv V_i(z_i, w_i, R_i(z_i, w_i)) - V_i(z_i, w_i', R_i(z_i, w_i)). \tag{3}$$

The previous discussion lays out the definition of our key quantities and is independent of the specific statistical model used to estimate it (cf. Pearl 2001; Imai et al. 2010). Its value lies not only in stating clearly what we want to know, but also in making explicit the central identifying assumptions needed to estimate these quantities (VanderWeele and Vansteelandt 2009; VanderWeele 2010; Imai et al. 2011).[7] The first is the standard assumption that, after conditioning on included observables, there are no unobserved confounders that change with treatment (e.g., income) and affect vote choice (V_i) or preferences (R_i). The second assumption concerns the mediating variable, namely redistribution preferences. It requires that no unobserved confounders affect both V_i and R_i after conditioning on observables c_i.

In our empirical application, as in any analysis having to rely on observational data, we accept that these conditions are likely to be violated to some degree. We therefore accompany our estimates with sensitivity analyses to gauge how increasingly severe violations of these identifying assumptions influence our results. It is also clear that the set of selected covariates plays an important role in securing credible inference.[8] In our analysis we include model extensions where we select confounders from a high-dimensional vector of controls employing new insights from the econometrics and machine learning literature.

We describe the statistical model used to estimate the quantities described above in Section III. Let us emphasize again that this setup provides a rather strict test of our hypotheses. We test if income and risk systematically shape the vote via redistribution preferences while allowing for (an unspecified number of) other channels by which risk and income could be linked to vote choice.

[7] The usual assumptions of standard regression models still apply. What we focus our discussion on are additional assumptions needed to decompose different mechanisms.

[8] Note that above we conditioned on one common set of confounders. It might be necessary to use different sets of covariates to deconfound different parts of the model. We examine this issue more closely in one of our robustness tests.

III. Data and Statistical Specification

III.A. Data sources

For our Western European analysis, we use data from eight waves of the European Social Survey (ESS), collected between September 2002 and June 2017. It is a large scale multi-country survey administered bi-annually in European countries starting in 2002. Its target population are all individuals aged 15 or over, residing in private households (regardless of nationality, language, citizenship or legal status). The ESS provides a measure of income that is applied consistently over countries and survey waves, and which provides enough detail for us to construct a usable measure of an individuals' income distance to the national mean. We select countries who participated in at least two rounds. For each election between 1999 and 2016, we match the corresponding waves from the ESS. If multiple waves were available, we use the one closest to the last election. We also eliminate surveys that were conducted in months that include an election (and therefore may contain voting choices for different elections depending on the respondent's interview date).[9] Table A1.1 in the appendix shows survey fieldwork periods and election dates for waves included in our analysis. In practical terms, this means the analysis matches macro-micro data for 55 elections held in 14 countries, namely Austria, Belgium, Germany, Denmark, Spain, Finland, Great Britain, Ireland, Italy, the Netherlands, Norway, Portugal, Sweden, and Switzerland. For our analysis of the United States, we use data from the American National Election Study (ANES) Time Series surveys administered to a sample of a cross-section of eligible voters in the US. We select surveys starting in 1982 (when our redistribution preference measure becomes available) and ending in 2016. They are available bi-annually until 2004 after which they are conducted in four-year intervals.

We exclude individuals with missing responses on control variables, as well as individuals with missing responses on both vote choice and preference questions. This leaves us with 67,191 individuals in Western Europe and 16,103 individuals in the United States. We impute missing values on the remaining variables (vote choice, preferences, income, and risk) as part of our statistical model using a fully Bayesian imputation strategy (Ibrahim et al. 2005). See appendix A2 for more details (and a comparison with results based on listwise deletion and non-parametric multiple imputation).

[9] The influence of redistribution preferences is the main focus in this paper's analysis of voting. For this reason, it is of paramount importance that the voting data coincides with the redistribution preferences data. As explained in more detail below, respondents are asked about the parties they voted for in the previous national election. At the time of the survey, these elections have taken place in the past while redistribution preferences are measured in the present. It is important therefore to restrict the analysis to ESS waves when this coincidence of data is reasonable. The same considerations apply to measures of relative income and risk, which are part of the redistribution preferences estimation. This also requires special attention to when the surveys were actually conducted.

III.B. Measures

Vote choice Our main dependent variable is an individual's choice to vote for a redistributive party. Recall from our theoretical intuitions above that we model voting as a discrete choice influenced by the distance between an individual's redistribution preferences and the redistributive positions of the parties she can vote for. This approach requires us to define whether a party is redistributive or not. In the US context, this is straightforward enough. Our dependent variable translates into a respondent choosing the Democratic Party (which consistently offers relatively more redistributive policy positions) over the Republican alternative.

In the Western European multi-party context, this issue is more complex. On the one hand, we could use a party 'label' as the indicator of redistributive position. In this approach, a 'left' party would be considered a redistributive party by virtue of its ideology and its commitments to historically meaningful groups of voters. The existence of stable ideological and historical connections between parties and some social groups "not only creates easily identifiable choices for citizens, it also makes it easier for parties to seek out their probable supporters and mobilize them at election time" (Powell 1982: 116). To the extent that party labels are used as information shortcuts by voters to capture a party's redistributive position, this is an attractive strategy. In the analysis below, we classify parties as 'left' if their party family (as recorded by the Comparative Manifesto Project, CMP) is either socialist/social democratic or communist.[10]

Labels, ideology and history, however, are not enough. Elections need to be contested and they inevitably revolve around issues, like redistribution, that give political meaning to partisan attachments and social divisions (Dalton 2002: 195). Moreover, in our analysis of Western Europe, simply classifying parties based on their label might not constitute an accurate operationalization of the concept of redistributive voting, since country- as well as election-specific factors influence parties' position on redistribution.

We therefore construct an alternative dependent variable based on of how much redistribution a party proposes in its electoral platform (Stegmueller 2013).[11] Using data from the Comparative Manifesto Project (Budge et al. 2001) and its 2016 update (Volkens et al. 2016), we calculate the extent to which parties favor state involvement in the economy—a measure of redistributive politics proposed by Benoit and Laver (2006, 2007).[12] It is calculated from parties' statements on multiple economic topics (represented by "quasi sentences" in the CMP data set), which are combined into a measure of a party's policy position as the balance of negative (N) to positive (P) statements[13] following Lowe et al. (2011): $\theta = \log \frac{N+.5}{P+.5}$. Parties can occupy any

[10] For a more detailed review of party families, see Mair and Mudde (1998).

[11] For an alternative approach, see Huber (2017), who identifies the location of parties in the left-right redistributive continuum as the mean of the redistributive preferences for voters of that party.

[12] One should note that using the CMP's simple "left-right" measure is misleading, since it carries surplus meaning which is not related to redistribution, such as positions on "traditional morality" (Huber and Stanig 2008).

[13] Positive statements include those referring to market regulation, economic planning, protectionism, controlled economy, nationalization, welfare, education, and labor groups. Negative statements refer to free market economy, incentives, (against) protectionism, economic orthodoxy, and (against) welfare.

position on this scale, but more extreme positions need considerable more relative emphasis, yielding a magnitude scaling of policy positions.[14] This yields interval level information on the redistributive policies of almost all European parties, where smaller values indicate a more pro-redistributive position.[15]

For the following analyses, we create a binary variable indicating if a party favors redistributive policies. We classify a party as redistributive if it takes a policy position below (or 'to the left of') the country-election specific redistribution policy mean, and thus proposing more redistribution than the (hypothetical) average party. Taking the mean as reference point is the preferred strategy, since the interval level measure of party policy does not imply that zero is a centrist position (cf. Lowe et al. 2011: 131). It also makes clear that this reference point changes (endogenously) with each new election in each country.

The distinction between 'redistributive' and 'left' parties turns out to be a significant one. Let's take the redistributive position of Spanish parties from 2002 to 2016 as an example. Using our first classification of redistributive parties based on the 'left' party label (socialist/social democratic or communist families), Labour, PSOE and Podemos (but not its related regional parties like En Marea) are considered redistributive throughout the period under analysis. It is clear, however, that when using the second classification, which is based on parties occupying more redistributive positions than the country-election mean, both Labour and PSOE are not considered redistributive in the 2015 elections. In Spain, the more redistributive positions of Izquierda Unida, Podemos and related regional parties pull down the country-election mean, which makes PSOE propose less redistribution than the (hypothetical) average party in that election.

Preferences For our measure of redistribution preferences in the United States, we follow Ashok et al. (2015) and use an item containing the following statement: "Some people feel that the government in Washington should see to it that every person has a job and a good standard of living. Others think the government should just let each person get ahead on their own." Respondents are then asked to place themselves on a 7-point scale with labeled end-points, ranging from "Government see to job and good standard of living" to "Government let each person get ahead on their own". The distribution of responses is shown in panel A of Figure II. In Western Europe our measure is an item commonly used in individual level research on preferences (e.g., Rehm 2009). It elicits a respondent's support for the statement "the government should take measures to reduce differences in income levels" measured on a 5 point agree-disagree scale with labeled answer categories ("Strongly agree" to "Strongly disagree"). Panel B of Figure II shows a histogram for our pooled European sample. We reverse

[14]Lowe et al. (2011) add a small constant (.5) to prevent problems with low numbers of quasi sentences.

[15]Some small, extreme parties are not represented in the data set, since the CMP contains no information on their position. An example is the National Democratic Party (NPD) in Germany, a nationalistic, extreme right party. However, the number of survey respondents that choose those parties is generally negligible.

the scale such that higher values represent support for redistribution.[16] It shows that Western Europe is characterized by a rather high level of popular support for redistribution. More than two thirds of ESS respondents either agree or strongly agree with the statement that the government should take measure to reduce income differences. Explicit opposition is much less widespread.

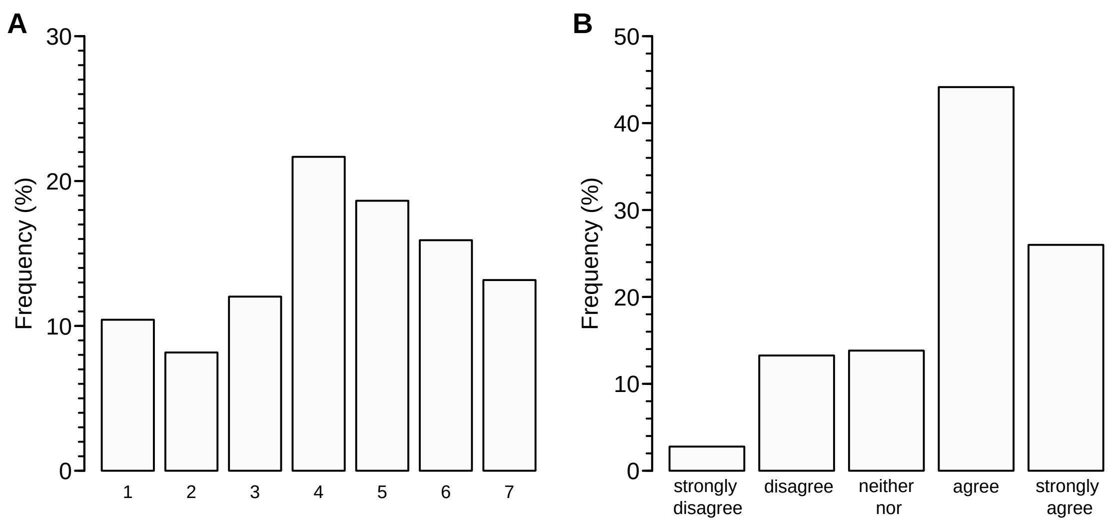

Figure II
Distribution of redistribution preferences in the United States and Western Europe.

In the theory section, we explained how the parameters in discrete choice voting models are often estimated using distance variables as the predictors. In the American analysis below, we have a measure explicitly capturing the spatial distance between a respondent's position on redistribution (as described above) and the respondent's view of each party's position on the same issue. In the analysis of Western European data, we use redistribution preferences as a predictor of redistributive party voting.[17]

A preliminary illustration of the influence of redistribution preferences on voting Before we delve into the full-fledged analysis decomposing the effects of income and risk, we start with a simpler question: are redistribution preferences related to individuals vote choices at all? To illustrate this relationship, we estimate models for our Western European and US sample in which we relate the probability of voting for a "left" or redistributive party (as defined above)

[16] All descriptive results are weighted for survey design characteristics.

[17] We should mention that, for the theoretical reasons outlined in the previous section of this paper, the dependent variable in our analysis needs to be voting choice (and not the actual redistributive position of different parties). We come back to this issue in the robustness section.

to the preferences over redistributive policies held by individuals.[18]

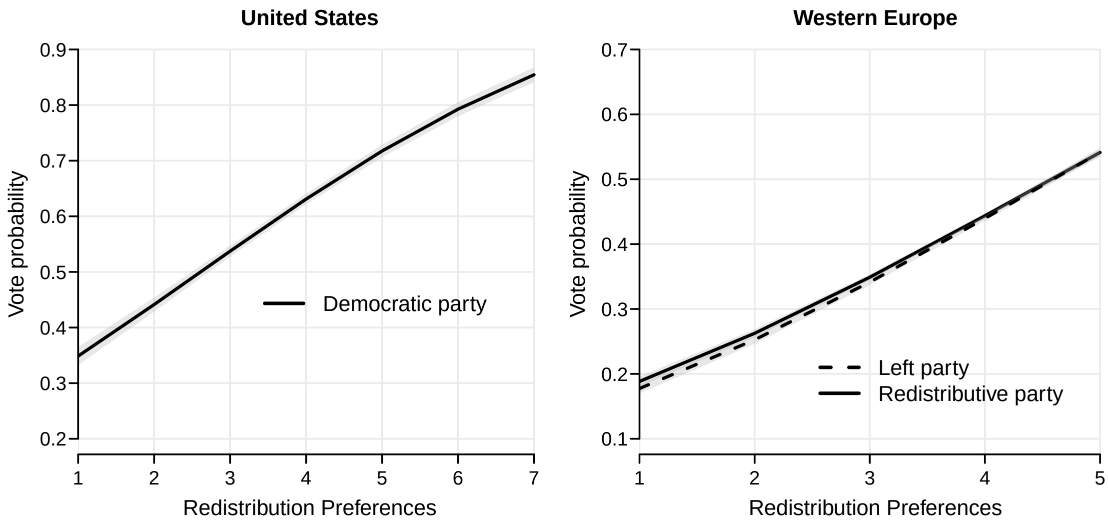

Figure III
Redistribution preferences and vote choice in the US and Western Europe

In Figure III, we plot average predicted probability of voting Democrat (in the US case) and voting for a party labelled as "left" or defined as redistributive based on its stated policy position (in the Western European case) as a function of preferences for redistribution.[19] As suggested in Figure II, the mean for the redistribution preferences question in our ANES sample is 3.6 (the standard deviation is 1.8). The mean in our ESS sample is around 3.5 (the standard deviation is 1). The left panel in Figure III therefore shows that a change in the demand for redistribution from around 2.5 to 4.5 (about a standard deviation around the mean), increases the likelihood of voting Democrat from around 50% to around 65% for an American respondent. A similar change using the Western European data (from 3 to 4, a standard deviation around the mean of 4), increases the likelihood of voting for a redistributive or 'left' party from around 35% to 45%. These are substantively important effects that can easily affect the outcome of elections. Do they hold up in an analysis decomposing the effects of income and risk?

Income distance Our central measure of an individual's material position is the distance between the income of respondents and the mean income in their country (at the time of the survey). In other words, we calculate income distance as a respondent's income minus the

[18]We specify probit models with state/country and year fixed effects and heteroscedasticity-consistent standard errors and calculate average predicted probabilities over the observed range of redistribution preferences.

[19]Note that we use respondents' self placement for the US in this preliminary analysis, while later models will use the spatial distance between a respondent's position on redistribution and the respondent's view of each party's position on the same issue.

country-year income mean.[20] The ANES captures income using an item asking a respondent to place his or her family's total market income in one of at least 22 income bands with boundaries varying throughout the years.[21] To create a measure of income that closely represents our theoretical concept, income distance, we follow the American Politics literature and transform income bands into their midpoints (e.g., Hout 2004).[22] We impute the open-ended top income category by assuming that the upper tail of the income distribution follows a Pareto distribution (e.g., Kestenbaum 1976, Kopczuk et al. 2010). Finally, for each respondent, we calculate the distance between her assigned and the national mean income in a given year.

In our European analysis, we rely on the same strategy. The ESS captures income by asking respondents to place their total net household income into a number of income bins giving yearly, monthly, or weekly figures.[23] We transform these categories into (country-year-specific) mid-points and impute the open-ended top income category from the Pareto distribution. The purchasing power of a certain amount of income varies across the countries included in our analysis. Simply put, it could be argued that the meaning of being Euro 10,000 below the mean is different in Sweden than in the United Kingdom.[24] Thus, for each country and each year, we convert a country's currency into PPP-adjusted constant 2010 US dollars. Finally, we calculate the distance of a respondent's income to the country-year mean.

Figure IV plots the distribution of income distance in the United States and Western Europe. The widely recognized difference in income inequality between Western Europe and the United States can be identified as two visible features in the figure. First, the income distribution in the US seems to be more right skewed (with a longer right tail of richer than average individuals) than the one in Europe. Second, the percentage of respondents around the mean seems higher in Europe that in the US.

[20]This represents a simple centering, which leaves the *distribution* of incomes unchanged. However, it takes into account that mean incomes differ over countries. For example, in 2004, the mean income (after PPP adjustment) in Sweden is 32,721, while in Austria it is 36,122. Note that using untransformed income yields the same pattern of substantive results.

[21]The exact question wording is: "Please [...] tell me the letter of the income group that includes the income of all members of your family living here in [year] before taxes. This figure should include salaries, wages, pensions, dividends, interest, and all other income." The wording varies slightly between phone and face-to-face interviews, but respondents face the same income bands.

[22]For example, this means that the third income category in 2000 ($10,000 to $14,999) becomes mid-point 12,500, while the third-to-last category ($185,000 to $194,999) becomes 190,000.

[23]It thus differs from the ANES in that only post-tax income is available. The exact question wording is: "Using this card, if you add up the income from all sources, which letter describes your household's total net income? If you don't know the exact figure, please give an estimate. Use the part of the card that you know best: weekly, monthly or annual income." The wording of this question between 2008 and 2012 varies, but the meaning remains the same. In these surveys, "after tax and compulsory deductions" replaces "net." From 2002 to 2006 the ESS used 12 income bands common to all countries, while starting in 2008 it used 10, based on each country's income deciles.

[24]And more importantly, it could be argued that the bulk of rich or poor people would be concentrated in the wealthiest (or most unequal) countries, therefore distorting our results.

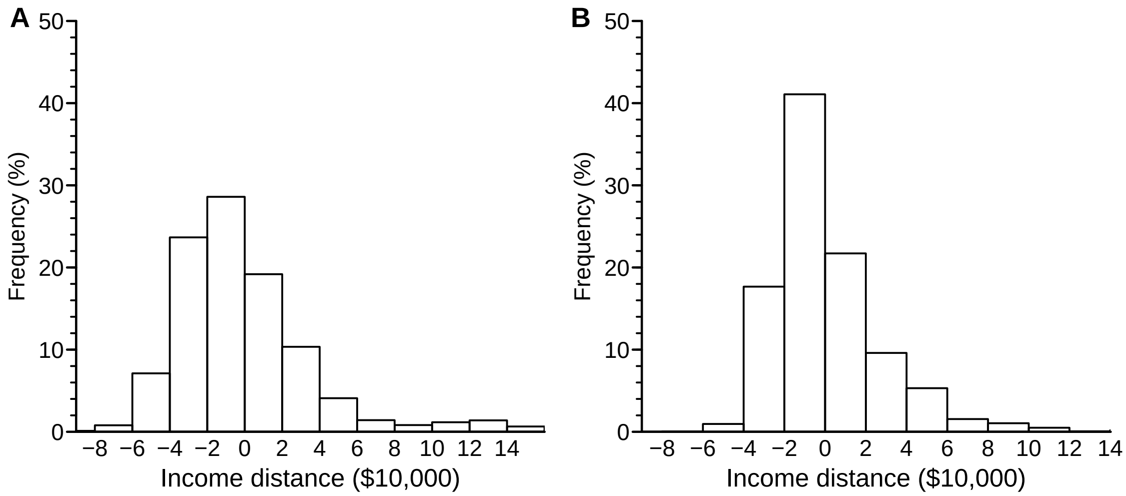

FIGURE IV
Distribution of income distance in the United states (A) and Western Europe (B).

Occupational risk To operationalize a respondent's exposure to occupational unemployment risk we follow the suggestion of Rehm (2005) and use the unemployment rate of his or her occupation.[25] In our US sample, we calculate occupational unemployment rates for 70 Census occupational categories from the Current Population Survey (CPS). To ensure comparability over time we use a constant classification scheme based on an aggregated occupational classification system in the 1990 Census (cf. Meyer and Osborne 2005), which we matched to the 70 occupation categories in the ANES.[26] An occupation's unemployment rate is the share of unemployed persons among the total economically active labor force (following the BLS definition), which we calculate separately for each decade from 1982 to 2016. All our calculations adjust for sample inclusion probabilities using survey design weights. Panel A of Figure V shows the distribution of occupational unemployment in our American sample and illustrates the spread of unemployment risk. Some occupations have a rather low risk of unemployment (< 2%). This group includes, lawyers, judges, engineers and architects, but also some plant and systems operators. At the other end of the distribution are occupations such as workers in freight handling and (non-managerial) farm occupations, which regularly face unemployment risk of more than 10%. Some workers with higher skills, such as mechanics and repairers, are unemployed at medium risk rates (of about 5 to 6%). And a seizable group of individuals in traditional "middle class" occupations, such as specialized lab technicians, engineers, urban planners and architects, face relatively low unemployment risk (ranging from 2 to 4%).

[25]In our main analyses we do not distinguish between unemployment rates for men and women. Both male and female occupational unemployment rates are highly correlated with average rates (0.97 / 0.94 for men / women in the US; 0.97 / 0.95 in Western Europe).

[26]ANES categories are based on a simplified version of 1980 Census occupations. The crosswalk from these categories to the aggregated scheme proposed by Meyer and Osborne is almost completely 1:1.

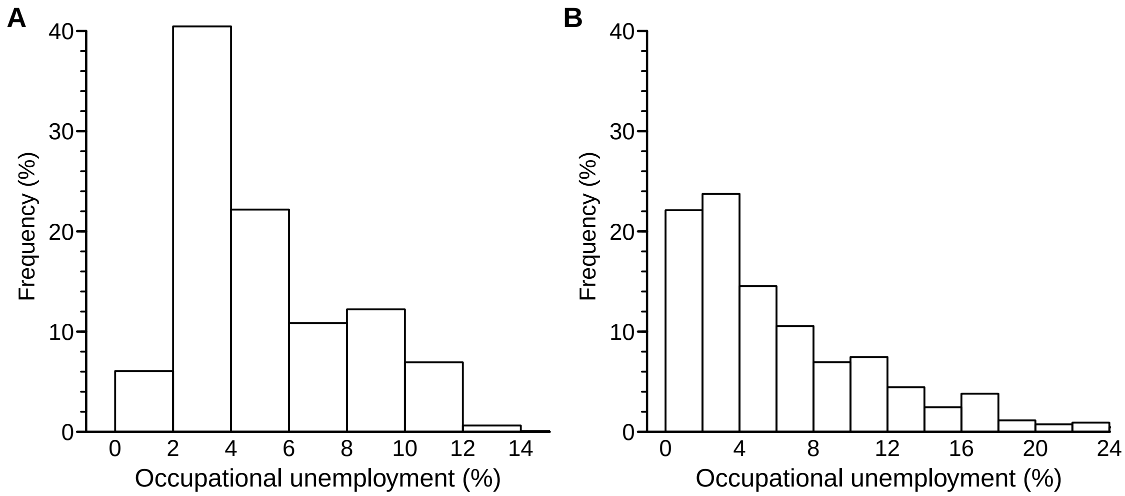

Figure V
Occupational unemployment risk in the United states (A) and Western Europe (B).

In our Western European analysis, we similarly calculate occupational unemployment rates from the European Union Statistics on Income and Living Conditions (EU-SILC), which provides high quality, statistically harmonized microdata of the European population. For each country we calculate the share of unemployed persons among the economically active workforce for 26 occupations based on the 2-digit aggregation of the International Standard Classification of Occupations (available both in EU-SILC and ESS). Due to the shorter time span covered by the ESS, we do not use decade-specific values in our main results.[27] Our calculations adjust for sample inclusion probabilities using survey design weights.

Panel B of Figure V plots the distribution of occupational unemployment in the Western European sample. Compared to the US, while the average risk of unemployment is quite comparable (5.1% in the US and 6.3% in Western Europe), the 90th quantile of the risk distribution is much higher in Europe (14% versus 9%). Occupational groups with particularly high levels of risk are laborers in construction, mining, agriculture and fisheries, as well as some building trade workers. The histogram also reveals the existence of a seizable group with very low levels of risk (< 2%), which is dominated by corporate and general managers, as well as health professionals and teachers.

Individual characteristics The models below include a number of individual characteristics to adjust for observable differences between individuals. We refrain from including a large set of variables, since many are arguably post-treatment (to income and occupational risk). We include age (in years), gender (an indicator for female), years of schooling, labor force status,

[27] But we do provide a robustness test with decade-specific occupational unemployment rates, which also capture pre- and post-crisis trends. See details below.

and household size. We explore the impact of other variables in a robustness section.

III.C. Statistical specification

We now describe how we model individuals' vote choices and how they are shaped by (endogenous) redistribution preferences. Let V_{ijt} represent the observed vote choice of individual i ($i = 1, \ldots, n_{jt}$) in geographical unit j ($j = 1, \ldots, J$) at time point (survey year) t ($t = 1, \ldots, T$). When analyzing Western Europe, the geographical units are countries, while in the US analysis they are states.

In a decision theoretic formulation, an individual will vote for a party if the utility derived from that choice, V^*_{ijt}, exceeds that of the alternative. In our setting we have a simplified choice set (redistributive vs. non-redistributive, Democrat vs. Republican), so that we observe $V_{ijt} = 1$ if $V^*_{ijt} > 0$ (and zero otherwise). Our measures of preferences are the categorical survey items described above and denoted by R_{ijt}. For simplicity, we treat them as continuous.[28]

We want to model the role of income distance and occupational risk in shaping preferences and how preferences themselves influence vote choice. Thus we jointly estimate the following two equations:

$$R_{ijt} = \beta_1 w_{ijt} + \beta_2 z_{ijt} + x'_{ijt}\delta^R + \xi_{jt} + \epsilon^R_{ijt} \tag{4}$$

$$V^*_{ijt} = \alpha R_{ijt} + \gamma_1 w_{ijt} + \gamma_2 z_{ijt} + x'_{ijt}\delta^V + \theta\xi_{jt} + \epsilon^V_{ijt} \tag{5}$$

Here, redistribution preferences are a function of income distance, w_{ijt} and occupational risk, z_{ijt}, captured by their respective β coefficients. Preferences then enter the vote choice equation, with their effect captured by α. In the Western European case they enter as R_{ijt}, while in our US analysis, they are the relative squared distance between a respondent's preferred policy position and the (perceived) position of each party, $R^*_{ijt} = (R_{ijt} - R^P_{ijt})^2$. Income distance and occupational risk also enter the vote choice equation directly (in addition to impacting it via changing preferences); their role is captured by the two γ coefficients. Both equations also include a vector of individual controls, x_{ijt}, with associated coefficients δ^R and δ^V, respectively.

The reader may have noted that our specification does not include country-level variables. In order to adjust for macro-level confounders, we include fixed effects for both geographic units (countries, states) and survey-year, denoted by ξ_{jt}. Note, that this amounts to including $J \times T$ effects compared to the $J + T$ effects that would be included in a specification that treats geography and time as separable (i.e., the usual state and time fixed effects setup). Thus, we model *within*-state and within-country changes, while accounting for state- and country-specific election contexts. We implement these as the Bayesian version of the classical 'fixed effects' model. We allow covariates to be related to group-specific effects by employing the

[28]Using a more complex latent variable model for ordinal outcomes does not make a substantive difference to our results.

Chamberlain-Mundlak device (an orthogonal projection of ξ_{jt} on covariate group averages; cf. Mundlak 1978; Chamberlain 1982).[29] Country/state-level unobservables are allowed to affect preferences and vote choices differently due to the scale factor θ in equation (5).[30] Since our focus in this paper are micro-level relationships, this model exhaustively adjusts for country- or state-level confounders as well as for election-time confounders and we do not include country-level controls.

Finally, residuals ϵ_V and ϵ_R are both zero-mean normally distributed. While the variance of ϵ_R is freely estimated, the variance of ϵ_V is fixed to one to identify the probit equation. With estimates from our joint preference and vote model in hand, we can calculate the direct and indirect (counterfactual) effects specified in equations (2) and (3). Appendix A3 shows how these are derived from our model estimates.

Estimation We estimate our model using MCMC sampling. This allows us to obtain the full posterior distribution of not just the model parameters, but also all derived quantities (such as indirect effects) and sensitivity simulations. We assign uninformative priors to all model parameters.[31] All identification is classical. In our results tables we display means of the posterior distribution as 'estimates,' together with posterior standard deviations, which can be thought of as the Bayesian equivalent to classical standard errors.

IV. Results

IV.A. Western European sample

Table I shows estimates and derived quantities from our model for the European analysis. In panel (A) we show the relevant parameter estimates from equations (4) and (5), omitting other estimates for reasons of space. We divide the results in Table I into two columns, the first one uses the 'left' label definition of voting for a redistributive party and the second one uses the definition of voting for parties that are more redistributive than the country-election mean. The distance of a respondent's income to the country average has the expected negative impact on preferences and is clearly statistically different from zero. The same is true for its coefficient in the vote choice equations. Higher income is associated with a lower likelihood of voting both for a left party and for a redistributive party (the coefficient being substantively higher when looking at parties that are more redistributive than the country-election mean). We find similar relationships regarding occupational risk. Individuals belonging to occupational groups with

[29] See Rendon (2012) for an extended discussion of fixed and random effects in a Bayesian context.

[30] We can test if effects do indeed differ between preferences and choices by testing if $\theta = 1$. Such a specification is rejected.

[31] More explicitly, we set all regression-type parameters to be *a priori* normally distributed mean zero with a standard deviation of 10. For the free variances in the model we use vague inverse Gamma priors (Spiegelhalter et al. 1997), $IG(0.001, 0.001)$. Given our large sample, the data clearly dominate these prior choices.

higher unemployment rates have a stronger preference for redistributive policies. Occupational risk in turn has a positive and significant effect on voting for both left and redistributive parties (although in this case the effect is larger for parties with the 'left' label). It should be noted that this effect of occupational risk is *net* of income by construction (we orthogonalize both variables). The last parameter of interest in equation (5) is α, the effect of endogenous preferences on the probability of choosing a 'left' party or one proposing redistributive polices (relative to other parties in a particular election). We find clear evidence for a strong link between a respondent's preferences and her party choice.

TABLE I
Model results for left and redistributive party choice. Western European sample.

	Left party		Redistributive party	
(A) Coefficient estimates				
	Preferences	Vote	Preferences	Vote
Income distance	−0.149 (0.004)	−0.034 (0.006)	−0.149 (0.004)	−0.082 (0.006)
Occupational risk	0.034 (0.005)	0.055 (0.008)	0.034 (0.005)	0.029 (0.008)
Preferences		0.203 (0.006)		0.290 (0.006)
(B) Effect decomposition				
	Income	Risk	Income	Risk
Indirect effect	−0.979 (0.052)	0.240 (0.038)	−1.578 (0.064)	0.374 (0.059)
Direct effect	−1.137 (0.215)	1.870 (0.266)	−3.038 (0.238)	1.097 (0.285)
Proportion	0.467 (0.051)	0.115 (0.022)	0.343 (0.020)	0.264 (0.067)
(C) Sensitivity analysis for IE [a]				
	Income	Risk	Income	Risk
$\rho \in \{0, \ldots, 0.3\}$	−0.264 (0.029)	0.059 (0.011)	−0.264 (0.029)	0.059 (0.011)
$\rho \in \{-0.3, \ldots, 0\}$	−1.555 (0.068)	0.342 (0.054)	−1.555 (0.068)	0.342 (0.054)

Note: Estimates (posterior means) with posterior standard deviations in parentheses. Based on 20,000 MCMC samples. Units in panel (A) are probit; units in panels (B) and (C) are differences in probabilities (in % pts). All continuous inputs are standardized to have mean zero and unit variance. Categorical inputs are mean zero. N=60,739.

a Sensitivity analysis for mediator-outcome confounding, simulated over 100-point grid $\rho \in \{a, \ldots, b\}$. Displayed results are averages over 100 simulations. Based on 5,000 MCMC samples.

We now turn to a quantitative assessment of how much income and risk shape vote choice *via* preferences (and by how much they do not), by calculating the quantities described in equations (2) and (3). Results are shown in Panel (B) of Table I. Note that the metric of both IE and DE is the difference in probability of voting for a 'left' or redistributive party. We find that both income and risk significantly shape choices via preferences. Looking at our 'left' label classification of redistributive parties, a standard deviation (SD) increase in income distance decreases redistributive party choice via preferences by 0.98 (±0.05) percentage points, while its effect on vote choice that is due to factors other than preferences is 1.14 (±0.22) points. Redistribution preferences therefore account for 47% of the total effect of income that we observe. A standard deviation increase in occupational unemployment risk increases the probability of voting for 'left' parties via its effect on redistribution preferences by about 0.24 (±0.04) percentage points, accounting for 12% of the total effect. Its corresponding effects not due to redistribution preferences is 1.87 (±0.27) points. When focusing on the definition of voting for parties that are more redistributive than the country-election mean, an increase in income distance decreases redistributive party choice via preferences by 1.58 (±0.06) percentage points, while its effect on vote choice that is due to factors other than preferences is 3.04 (±0.24) points. Thus, when considering redistributive parties, redistribution preferences account for 34% of the total effect of income that we observe. A standard deviation increase in occupational unemployment risk increases the probability of voting for a redistributive parties via its effect on redistribution preferences by 0.37 (±0.06) percentage points. It thus accounts for a much larger share (26%) of the total effect compared to classifying parties by label. Its corresponding effect not due to redistribution preferences is 1.10 (±0.29) percentage points.

What these results tell us is that (i) both income distance and risk shape preferences, (ii) income distance and risk in turn affect the probability of voting for a party offering redistributive policies, and (iii) that there also is a direct effect of income and risk on vote choice (not due to their effect on preferences).

The value of defining (natural) indirect effects is that it makes explicit the assumptions needed to estimate them. As we mentioned above, these assumptions are unlikely to be met in an observational analysis such as ours. Thus, the best available strategy is to assess the robustness of our results by conducting sensitivity analyses (VanderWeele 2010; Imai et al. 2010). In panel (C) of Table I we show the results of several sensitivity analyses, where we average over 100 increasingly extreme levels of unobserved confounders affecting both preferences and vote choice. The empirical implication of an unobserved confounder affecting both observed values of the mediator and potential outcomes is a correlation between residuals (Imai et al. 2010: 61); in our context $\rho(\epsilon_1, \epsilon_2)$.[32] Simulating this correlation, we evaluate our indirect effect estimates over a 100-point grid with successively increasing values of $\rho \in \{-r, \ldots, 0\}$ and $\rho \in \{0, \ldots, r\}$, where the limit r is chosen to represent ranges of possible correlations. We use 5,000 Monte Carlo samples (obtained from the posterior distribution of each parameter) to account for estimation uncertainty. We present both averages of estimated indirect effects (putting equal weights on all possible levels of confounding) as well as posterior standard deviations of the distribution of indirect effect estimates.[33]

Panel (C) show how accounting for possible confounding affects our (average) results. To give an idea of the substantive magnitude of the level of confounding we are simulating: we

[32] See model equations 4 and 5

[33] Our results should be understood as the average of mediated effects over all levels of possible correlations simultaneously. In the appendix we calculate mediated effects for larger ranges of possible correlations.

use a range from no correlation to 0.3, which is roughly the relationship (generally recognized as strong) between education and income observed in our data (0.29 in Western Europe and 0.30 in the US). When accounting for confounders that induce a positive correlation between preferences and vote choice, our indirect effect estimates are reduced: we find that the effect of income via preferences is now about one fifth of its previous size, although it is still clearly statistically different from zero. The size of the indirect effect of occupational risk (not due to income) is similarly reduced and also remains statistically different from zero. In contrast, accounting for confounders that induce a negative correlation between preferences and choices yields indirect effect estimates that are (even) larger than those obtained assuming a correlation of zero. This is particularly the case for the indirect effect of income, which increases by almost 0.5 percentage points.

IV.B. United States sample

Table II shows estimates and derived quantities from our American analysis. The structure of the results is the same as in Table I, but in this case there is only one definition of redistributive voting (voting for the Democratic Party). Note that in this analysis the measure of preferences is explicitly about the distance between a respondent's position on redistribution and the respondent's perception of each party's position on the same issue. In panel (A), as was the case in the European analysis, we find the distance of a respondent's income to the country average to have the expected negative impact on preferences and on vote choice (both clearly statistically different from zero). We also find that occupational risk significantly impacts redistribution preferences, but its coefficient in the vote equation is indistinguishable from zero. As in our analysis of European data, the effect of occupational risk is *net* of income by construction. Again, the last parameter of interest in panel (A) is the effect of endogenous preferences on the probability of voting for the Democratic Party.

Panel (B) of Table II shows that income significantly shapes choices via preferences. A standard deviation increase in income distance decreases redistributive party choice via preferences by 2.73 (±0.25) percentage points, while its effect on vote choice that is due to factors other than preferences is 3.00 (±0.48) points. Redistribution preferences therefore account for about 48% of the total effect of income that we observe. An equal sized increase in occupational risk increases redistributive party choice via preferences by 1.02 (±0.28) percentage points. The effect of risk via other channels is small and indistinguishable from zero. Thus the structure of occupational risks in the United States impacts individuals' party choice almost completely via its effect on preferences for redistribution.

The results for the American analysis therefore convey a very similar message to those for the European one: (i) both income distance and risk shape preferences, (ii) income distance and risk in turn affect the probability of voting for a party offering redistributive policies, and (iii) that there also is a direct effect of income (but not of risk) on vote choice (not due to their

TABLE II
Model results for Democratic vote choice. United States sample.

(A) Coefficient estimates				
	Preferences		Vote	
Income distance	−0.161	(0.014)	−0.090	(0.014)
Occupational risk	0.062	(0.017)	0.004	(0.018)
Preferences			0.503	(0.019)
(B) Effect decomposition				
	Income		Risk	
Indirect effect	−2.728	(0.250)	1.024	(0.283)
Direct effect	−2.996	(0.479)	0.142	(0.588)
Proportion	0.479	(0.052)	0.876	(0.405)[a]
(C) Sensitivity analysis for IE [b]				
	Income		Risk	
$\rho \in \{0, \ldots, 0.3\}$	−1.870	(0.177)	0.701	(0.177)
$\rho \in \{-0.3, \ldots, 0\}$	−2.874	(0.253)	1.088	(0.253)

Note: Estimates (posterior means) with posterior standard deviations in parentheses. Based on 20,000 MCMC samples. Units in panel (A) are probit; units in panels (B) and (C) are differences in probabilities (in % pts). All continuous inputs are standardized to have mean zero and unit variance. Categorical inputs are mean zero. N=13,930.

a Calculated from trimmed posterior distribution (due to non-normality of ratio).

b Sensitivity analysis for mediator-outcome confounding, simulated over 100-point grid with $\rho \in \{a, \ldots, b\}$. Displayed results are averages over 100 simulations. Based on 5,000 MCMC samples.

effect on preferences).

Panel (C) shows average estimates over a range of omitted variables confounding the preferences-vote relationship. When simulating omitted confounders that induce a positive correlation between preferences and choices, we find that the indirect effect of income is reduced by 0.86 percentage points, while the indirect effect estimate for risk is reduced by 0.12 points. As in our European analysis, we find that if omitted confounders suppress the correlation between both equations, indirect effect estimates for both income and risk would be somewhat larger in magnitude (although the change is comparatively smaller). In sum, our simulation suggests that preferences are a significant channel (in both the substantive and statistical sense) linking income and risk to Democratic party choice.

IV.C. Robustness checks

We also conduct a number of robustness checks dealing with omitted variables affecting the relationship between income, risk and preferences. In order not to display a wealth of specifications, we group some of them together. In the first test, displayed in panel (A) of Table III, we include variables capturing distinct economic characteristics of a respondent: whether he or she is a member of a trade union, unemployed, or (in the European analysis only) holding a fixed-term contract. While much of this paper focuses on the effects of risk exposure measured as occupational unemployment, these are alternative risk-related factors that could possibly affect our results. Like Cusack et al. (2006) the unemployment variable can be interpreted as a measure of realized risk, while union membership and temporary contracts capture labor market 'dualization' that protects 'insiders' instead of 'outsiders' (see (Rueda 2007)).

Our second test, displayed in panel (B), is designed to capture socio-cultural characteristics: religiosity and distinct preferences of individuals living in high-density, urban areas (see, for example, Cho et al. 2006).[34] We include indicator variables for the two dominant religious groups in Western Europe as well as a variable capturing the frequency with which a respondent attends religious services.[35] We also include an indicator equal to one if the respondent lives in a major city or its outskirts and suburbs.

These specifications produce rather similar results. Adjusting for union membership, type of contract and unemployment status, as well as religion and urban density, leads to reduced indirect effect of income distance and risk on vote choice channeled via redistribution preferences. This holds for both Western Europe (where the dependent variable is redistributive party choice) and the United States. However, since direct estimates are also affected, the share of the total effect of income on vote choice due to preferences remains virtually unchanged (within ± 2 percentage points). The role of occupational unemployment risk (as well as its mediated proportion) in Western Europe is similarly unchanged. In the US sample, the indirect effect of risk is reduced by between 0.2 and 0.3 percentage points. The direct effect of risk on vote choice (previously estimated close to zero) now has a negative sign with a large posterior standard deviation, keeping it statistically indistinguishable from zero. Consequently, when adjusting for additional covariates the proportion of the risk effect mediated by preferences is even closer to 1 than previously estimated.

We use the appendix to provide additional robustness tests. In Table A2.1 we present results related to the treatment of missing values, including listwise deletion and non-parametric multiple imputation. Our measure of vote choice is based on respondents recalling their vote

[34] As argued by Rodden (2010: 322), it is clear that individuals sort themselves into neighborhoods with similar demographic, occupational, income, and ultimately political preferences. Since it has direct effect on both preferences and choices, urban location is therefore not included in our main model.

[35] Note that, as has been argued by Stegmueller (2013), religion also affects *economic* preferences and choices. We include it here to capture possible non-economic ("second dimension") considerations.

Table III
Robustness checks. Effect decomposition estimates with posterior standard deviation in parentheses.

	Western Europe		United States	
	Income	Risk	Income	Risk
(A) Economic variables[a]				
Indirect effect	−1.486	0.323	−2.632	0.794
	(0.061)	(0.056)	(0.248)	(0.282)
Direct effect	−2.876	1.020	−3.007	−0.147
	(0.239)	(0.292)	(0.486)	(0.592)
(B) Cultural variables[b]				
Indirect effect	−1.372	0.287	−2.458	0.697
	(0.061)	(0.057)	(0.243)	(0.275)
Direct effect	−2.877	0.974	−2.978	−0.530
	(0.251)	(0.312)	(0.488)	(0.604)
(C) Immigration preferences[c]				
Indirect effect	−1.613	0.425	−2.441	0.914
	(0.064)	(0.057)	(0.231)	(0.245)
Direct effect	−3.028	1.096	−2.609	0.112
	(0.236)	(0.291)	(0.443)	(0.511)

Note: Posterior means with posterior standard deviations in parentheses. Based on 16,000 MCMC samples. Western Europe LHS variable is social democratic party vote.

a Indicators for unemployment, union membership, limited contract (the latter only in WE).

b Religion (Catholic, Protestant, else) and frequency of church attendance. Indicator for living in urban area.

c EU: 1 dim factor, US: admit immigrants. See appendix A5

cast in the last election. This introduces the possibility of recall bias, be it due to simple errors in memory or respondents' tendency to 'adjust' their reported choices to contemporary preferences. In addition, due to the different timing of elections in each country, some respondents have to recall choices made further back in time. To get a sense of how this issue might affect our results, we compare party preferences based on the recall of *past* choices with stated *contemporary* party preferences at the time of each survey in appendix A4. Another potential concern with our analysis is that our measure of preferences conflates 'true' preferences for redistributive policies with other non-economic attitudes that are nonetheless related to the provision of social insurance. One such concern relates to attitudes towards immigrants and immigration. In appendix A5, we change our preference equation to allow for the possibility that individuals' stated redistribution preferences partly depend on their anti-immigration attitudes. We also do not explicitly include the role of social class in our main analyses but we consider social class an important factor structuring the distribution of both income and risk. We address this issue in detail in appendix A6. Finally, due to sample size constraints, our measure of occupational unemployment rates in our Western European sample is constant over time. To further investigate the robustness of our results, we calculate a two-period measure of occupational risk splitting the EU-SILC samples at 2008 in appendix A7.

IV.D. Relaxing model assumptions

In this subsection we change the model more radically, in order deal with issues of omitted confounders (in both mediator and vote equations) and functional form dependence. We employ a post double-selection strategy (Belloni et al. 2017: e.g.,), which aims to account for confounders in a flexible way by allowing for nonlinear functional forms (using cubic splines) and higher order interactions among covariates. This produces a high dimensional vector of (over 100) covariate terms, out of which a subset is selected using standard Machine Learning tools (such as the LASSO).[36] Importantly, this approach goes beyond being a causally naive model selection tool by jointly estimating a set of equations: the outcome equation, where vote choice is the dependent variable, and "selection" equations, where the dependent variables are income and risk (and preferences). This strategy explicitly accounts for the logic of omitted variable bias: confounders that matter in the selection stage are kept in the model even if they have only moderate weight in the outcome equation. The selected set of covariates is used in a post-LASSO estimation step to account for relevant confounders (see appendix A8 for more details). The resulting estimator is robust even under (moderate) selection mistakes (Chernozhukov et al. 2015; Belloni et al. 2014). The LASSO step selecting the control set from all equations is responsible for this robustness property. It finds controls whose omission leads to "large" omitted variable bias and includes them in the model. Any variables that are not included are therefore at most mildly associated to the treatment and the outcome, which decidedly limits the scope of omitted variable bias (Chernozhukov et al. 2015).

Table IV shows direct and indirect effect estimates. In panel (A) we apply the DSE to the income, risk and choice part of our model, while in panel (B) we additionally apply it to the preference equation. One key finding emerges from this exercise. We find the indirect effect estimates of income distance and risk to be somewhat reduced, but only in the Western European sample. Furthermore, since the direct effect of income on vote choice is also impacted, the proportion of the total effect on voting due to preferences sees far less change. For income, it is reduced by about 1 percentage point in Europe while it increases by about 3 percentage points in the United States. For risk, the proportion mediated decreases in Western Europe (by about 4% pts), while in the US it increases to one. In line with our result in the previous section, this suggest that the role of risk in Americans party choice is almost fully driven by redistributive policy considerations. In both cases these results obtain wether we select

[36]The key is to transform this system of equations into one that represents a predictive relationship (where the application of machine learning tools such as the LASSO make sense).

Table IV
Post-LASSO double selection results. Effect decomposition estimates with posterior standard deviation in parentheses.

	Western Europe		United States	
	Income	Risk	Income	Risk
(A) Exposure-outcome confounders				
Indirect effect	−1.380	0.433	−2.769	1.207
	(0.076)	(0.051)	(0.303)	(0.275)
Direct effect	−2.844	1.422	−2.781	0.209
	(0.256)	(0.251)	(0.642)	(0.647)
(B) Exposure-mediator-outcome confounders				
Indirect effect	−1.381	0.432	−2.769	1.207
	(0.075)	(0.051)	(0.303)	(0.275)
Direct effect	−2.833	1.436	−2.781	0.209
	(0.254)	(0.252)	(0.642)	(0.647)

Note: Includes control set $\hat{\mathcal{W}}$ obtained using double selection (using the root-LASSO) on a predictive system of treatment and outcome (and, in panel B, preference) equations. See appendix A8 for more details. For the purpose of variable selection, we treat the voice choice equation as a linear probability model. Western Europe LHS variable is redistributive party vote.

confounders affecting the income/risk-vote or the income/risk-preferences-vote relationship.

In sum, these results, while different in quantitative magnitude, strengthen the findings we obtained in our specification reported in Tables I and II since they seem not to be driven by functional form assumptions or a specific (linear-additive) combination of covariates.

V. Conclusion

We will not restate in this conclusion the theoretical arguments and empirical findings provided in the previous sections. But we will suggest a potential connection to two topics of increasing relevance to political science: the electoral decline of main left parties and the reemergence of populism. The key question for this paper is the extent to which redistribution preferences are part of the explanation for the decline of main left parties and the upsurge in populism. We have shown in the preceding pages that redistribution preferences matter significantly to voting. One plausible explanation is that in recent times, left parties have become less redistributive and populist parties more redistributive–and therefore more attractive to parts of the traditional core constituency of the left. On the one hand left parties face a “progressive dilemma” (Kymlicka 2015): maintaining a comprehensive welfare state in increasingly multicultural society without losing public support. On the other hand, some populist right parties have taken up “welfare chauvinism” as a way to appeal to poor voters (see, for example, De Koster et al. 2013).[37] Given the contemporary nature of these developments and the consequent lack of systematic data, we leave this topic to future work.

Two important issues are not fully resolved by our voting analysis. The first one concerns the calculus of winning and or losing votes from core and non-core constituencies, the second one concerns voting turnout itself. Regarding the first, we have explored in this paper both the factors affecting the demand for redistribution and the connection to individual voting. A more explicit connection between our "bottom up" political economy approach and the general comparative politics literature on dynamic party competition[38] is a productive area for future research. A particularly relevant element in these arguments is the distinction between core and non-core voters. As Downs (1957) recognized long ago, political parties are torn between the incentives to seek the support of pivotal, centrist voters and the need to cater to the interests of core supporters. Dixit and Londregan (1996) express the logic for parties to cater to the interest of swing voters in the center, noting that groups "that are densely represented at the center will be the beneficiaries of redistributive politics," whereas individuals closer to the extremes "will be written off by one party and taken for granted by the other" (1996: 1143). Others (see, for example, Cox and McCubbins 1986) argue that political parties have strong incentives to mobilize core voters by targeting benefits to those groups. Having clarified in this paper how the redistributive preferences of individuals are connected to voting, a necessary next step is to explore the momentous consequences of the strategic choices regarding core and non-core potential voters for redistributive outcomes.

We will point out, finally, that while a transfer of votes to populist alternatives is a possible explanation for the decline of main Left parties, the implications of the connection between preferences and voter turnout are also important. In our analysis of voting, like in much of the related literature, we excluded individuals who did not turn out to vote. We argued that a full model combining turnout and party choice is a lot more complex than simply including abstention as another "party" (see, e.g., Adams et al. 2006 for a more sophisticated approach). This issue has far-reaching implications. Let's assume the reader accepts the main arguments in this paper (that voters' preferences for redistribution are a function of their relative income and of risk) and also the conceptualization of parties as strategic actors responding to voter preferences. As argued by Pontusson and Rueda (2010), among others, in this theoretical framework the extent to which income inequality is associated with political inequality conditions party responses to voter preferences. The issue of income skew in voter turnout is therefore central to the relationship between redistribution preferences and voting. As Meltzer and Richard (1981) themselves recognize, their prediction that inequality will be associated with more redistribution rests on the unrealistic assumption that all income earners

[37] It is, however, also possible that redistribution preferences are becoming less relevant for voting left through time, while other preferences (potentially unrelated to redistribution) are making populist parties more attractive to the traditional core constituency of the left.

[38] See, for example, Stimson et al. (1995); Adams et al. (2004, 2009); Adams and Somer-Topcu (2009).

vote. Under any other circumstance, we are required to distinguish between the income of the median voter and the median income (Nelson 1999; Barnes 2013). To the extent that political inequality rises with income inequality (Leighley 1995; Schlozman et al. 2012), the effect of increasing income inequality on redistribution preferences and then voting might well be offset by a decline in electoral turnout among low-income and/or high-risk citizens. This is once more a topic to which these conclusions cannot dedicate the attention it deserves, and a potentially productive area for future research.

References

Achen, C. H. and L. M. Bartels (2016). *Democracy for Realists: Why Elections Do Not Produce Responsive Government.* Princeton University Press.

Adams, J., M. Clark, L. Ezrow, and G. Glasgow (2004). Understanding change and stability in party ideologies: Do parties respond to public opinion or to past election results? *British Journal of Political Science 34*(04), 589–610.

Adams, J., J. Dow, and S. Merrill (2006). The political consequences of alienation-based and indifference-based voter abstention: Applications to presidential elections. *Political Behavior 28*(1), 65–86.

Adams, J., A. B. Haupt, and H. Stoll (2009). What moves parties? the role of public opinion and global economic conditions in western europe. *Comparative Political Studies 42*(5), 611–639.

Adams, J. and Z. Somer-Topcu (2009). Policy adjustment by parties in response to rival parties' policy shifts: Spatial theory and the dynamics of party competition in twenty-five post-war democracies. *British Journal of Political Science 39*(04), 825–846.

Alesina, A. and P. Giuliano (2011). Preferences for Redistribution. In J. Benhabib, A. Bisin, and M. O. Jackson (Eds.), *Handbook of Social Economics*, pp. 93–131. San Diego: North-Holland.

Alvarez, R. M. and J. Nagler (1995). Voter choice in 1992: economics, issues and anger. *American Journal of Political Science 39*, 714–744.

Alvarez, R. M. and J. Nagler (1998). Economics, entitlements, and social issues: Voter choice in the 1996 presidential election. *American Journal of Political Science*, 1349–1363.

Alvarez, R. M., J. Nagler, and J. R. Willette (2000). Measuring the relative impact of issues and the economy in democratic elections. *Electoral Studies 19*(2), 237–253.

Andersen, J. G. and T. Bjørklund (1990). Structural changes and new cleavages: The progress parties in denmark and norway. *Acta Sociologica 33*(3), 195–217.

Ashok, V., I. Kuziemko, and E. Washington (2015). Support for redistribution in an age of rising inequality: New stylized facts and some tentative explanations. *Brookings Papers on Economic Activity*, 367–405.

Barnes, L. (2013). Does median voter income matter? the effects of inequality and turnout on government spending. *Political Studies 61*(1), 82–100.

Barro, R. (1973). The control of politicians: An economic model. *Public Choice 14*(1), 19–42.

Bartels, L. M. (2006). What's the Matter with What's the Matter with Kansas? *Quarterly Journal of Political Science 1*(1), 201–226.

Bartels, L. M. (2009). *Unequal democracy: The political economy of the new gilded age.* Princeton University Press.

Bean, C. and E. Papadakis (1998). A comparison of mass attitudes towards the welfare state in different institutional regimes, 1985–1990. *International Journal of Public Opinion Research 10*(3), 211–236.

Belloni, A., V. Chernozhukov, I. Fernández-Val, and C. Hansen (2017). Program evaluation and causal inference with high-dimensional data. *Econometrica 85*(1), 233–298.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). Inference on treatment effects after selection amongst high-dimensional controls. *Review of Economic Studies 81*, 608–650.

Belloni, A., V. Chernozhukov, and C. B. Hansen (2013). Inference for high-dimensional sparse econometric models. In D. Acemoglu, M. Arellano, and E. Dekel (Eds.), *Advances in Economics and Econometrics: Tenth World Congress*, Volume 3, pp. 245–295. Cambridge: Cambridge University Press.

Benabou, R. and E. A. Ok (2001). Social Mobility and the Demand for Redistribution: The Poum Hypothesis. *The Quarterly Journal of Economics 116*(2), 447–487.

Benoit, K. and M. Laver (2006). *Party Policy in Modern Democracies.* Oxon: Routledge.

Benoit, K. and M. Laver (2007). Estimating party policy positions: Comparing expert surveys and hand-coded content analysis. *Electoral Studies 26*, 90–107.

Beramendi, P. and P. Rehm (2016). Who gives, who gains? progressivity and preferences. *Comparative Political Studies 49*(4), 529–563.

Brooks, C. and D. Brady (1999). Income, economic voting, and long-term political change in the u.s., 1952–1996. *Social Forces 77*(4), 1339–1374.

Brooks, C. and J. Manza (1997). The social and ideological bases of middle-class political realignment in the united states, 1972 to 1992. *American Sociological Review 62*(2), 191–208.

Budge, I., H.-D. Klingemann, A. Volkens, J. Bara, E. Tanenbaum, R. C. Fording, D. J. Hearl, H. M. Kim, M. McDonald, and S. Mendez (2001). *Mapping Policy Preferences. Estimates for Parties, Electors, and Governments 1945-1998.* Oxford: Oxford University Press.

Burgette, L. F. and J. P. Reiter (2010). Multiple imputation for missing data via sequential regression trees. *American Journal of Epidemiology 172*(9), 1070–1076.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of Econometrics 18*(1), 5–46.

Chernozhukov, V., C. Hansen, and M. Spindler (2015). Valid post-selection and post-regularization inference: An elementary, general approach. *Annual Review of Economics 7*(1), 649–688.

Cho, W. K. T., J. G. Gimpel, and J. J. Dyck (2006). Residential Concentration, Political Socialization, and Voter Turnout. *The Journal of Politics 68*(01), 156–167.

Cox, G. W. and M. D. McCubbins (1986). Electoral politics as a redistributive game. *Journal of Politics 48*, 370–89.

Cusack, T., T. Iversen, and P. Rehm (2006). Risks at work: The demand and supply sides of government redistribution. *Oxford Review Of Economic Policy 22*(3), 365–389.

Dalton, R. J. (2002). *Citizen Politics: Public Opinion and Political Parties in Advanced Western Democracies.* Chatham House Publishers.

De Koster, W., P. Achterberg, and J. Van der Waal (2013). The new right and the welfare state: The electoral relevance of welfare chauvinism and welfare populism in the netherlands. *International Political Science Review 34*(1), 3–20.

Dion, M. (2010). When is it rational to redistribute? a cross-national examination of attitudes toward redistribution. In *delivery at the 2010 summer meeting of the Society of Political Methodology, University of Iowa, Iowa City, IA*, pp. 22–24.

Dion, M. L. and V. Birchfield (2010). Economic development, income inequality, and preferences for redistribution1. *International Studies Quarterly 54*(2), 315–334.

Dixit, A. and J. Londregan (1996). The determinants of success of special interests in redistributive politics. *Journal of politics 58*, 1132–1155.

Doove, L. L., S. Van Buuren, and E. Dusseldorp (2014). Recursive partitioning for missing data imputation in the presence of interaction effects. *Computational Statistics & Data Analysis 72*, 92–104.

Downs, A. (1957). *An economic theory of democracy.* New York: Harper.

Drazen, A. (2000). *Political Economy in Macroeconomics.* New Jersey: Princeton University Press.

Duch, R. and R. Stevenson (2008). *The Economic Vote: How Political and Economic Institutions Condition Election Results.* Political Economy of Institutions and Decisions. Cambridge University Press.

Evans, G. (1999). *The end of class politics?: class voting in comparative context.* Oxford University Press.

Evans, G. and N. D. de Graaf (Eds.) (2013). *Political choice matters: explaining the strength of class and religious cleavages in cross-national perspective.* Oxford: Oxford University Press.

Fair, R. C. (1978). The effect of economic events on votes for president. *The Review of Economics and Statistics 60*(2), 159–173.

Faist, T. (1994). Immigration, integration and the ethnicization of politics. *European Journal of Political Research 25*(4), 439–459.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice 50*(1-3), 5–25.

Finseraas, H. (2009). Income Inequality and Demand for Redistribution: A Multilevel Analysis of European Public Opinion. *Scandinavian Political Studies 32*(1), 94–119.

Fiorina, M. P. (1981). *Retrospective voting in American national elections.* Yale University Press.

Frank, T. (2004). *What's the matter with Kansas? How conservatives won the heart of America.* New York: Metropolitan Books.

Gelman, A., G. King, and C. Liu (1998). Not asked and not answered: Multiple imputation for multiple surveys. *Journal of the American Statistical Association 93*(443), 846–857.

Gelman, A., D. K. Park, B. Shor, J. Bafumi, and J. Cortina (2008). *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do.* Princeton: Princeton University Press.

Gilens, M. (2005). Inequality and Democratic Responsiveness. *Public Opinion Quarterly 69*(5), 778–796.

Green, D. P., B. Palmquist, and E. Schickler (2004). *Partisan hearts and minds: Political parties and the social identities of voters.* Yale University Press.

Hacker, J. S. and P. Pierson (2011). *Winner-Take-All Politics: How Washington Made the Rich Richer–and Turned Its Back on the Middle Class.* Simon and Schuster.

Haider, S. and G. Solon (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review 96*(4), 1308–1320.

Hersh, E. D. and C. Nall (2016). The primacy of race in the geography of income-based voting: New evidence from public voting records. *American Journal of Political Science 60*(2), 289–303.

Hout, M. (2004). Getting the most out of the GSS income measures. GSS Methodological Report 101.

Huber, J. D. (2017). *Exclusion by Elections. Inequality, Ethnic Identity, and Democracy.* Cambridge: Cambridge University Press.

Huber, J. D. and P. Stanig (2008). Voting polarization on redistribution across democracies. Manuscript.

Ibrahim, J. G., M.-H. Chen, S. R. Lipsitz, and A. H. Herring (2005). Missing-data methods for generalized linear models: A comparative review. *Journal of the American Statistical Association 100*(469), 332–346.

Imai, K., L. Keele, and D. Tingley (2010). A general approach to causal mediation analysis. *Psychological Methods 15*(4), 309–334.

Imai, K., L. Keele, D. Tingley, and T. Yamamoto (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review 105*(4), 765–789.

Imai, K., L. Keele, and T. Yamamoto (2010). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science 25*(1), 51–71.

Iversen, T. and D. Soskice (2001). An asset theory of social policy preferences. *American Political Science Review 95*(4), 875–893.

Keele, L. (2015). Causal mediation analysis: Warning! assumptions ahead. *American Journal of Evaluation 36*(4), 500–513.

Kestenbaum, B. (1976). Evaluating ssas current procedures for estimating untaxed wages. *American Statistical Association Proceedings of the Social Statistics Section, Part 2*, 461–465.

Key, V. O. (1966). *The responsible electorate.* New York: Vintage Books.

Kopczuk, W., E. Saez, and J. Song (2010). Earnings inequality and mobility in the united states: Evidence from social security data since 1937. *The Quarterly Journal of Economics 125*(1), 91–128.

Kramer, G. H. (1971). Short-term fluctuations in U.S. voting behavior, 1896–1964. *American Political Science Review 65*(3), 131–143.

Kymlicka, W. (2015). Solidarity in diverse societies: beyond neoliberal multiculturalism and welfare chauvinism. *Comparative Migration Studies 3*(1), 17.

Leeb, H. and B. M. Pötscher (2008). Can one estimate the unconditional distribution of post-model-selection estimators? *Econometric Theory 24*(02), 338–376.

Leighley, J. E. (1995). Attitudes, opportunities and incentives: A field essay on political participation. *Political Research Quarterly*, 181–209.

Lewis-Beck, M. S. (2009). *The American voter revisited.* Ann Arbor: University of Michigan Press.

Lipset, S. M. (1983). *Political Man. The Social Bases of Politics.* London: Heinemann.

Lowe, W., K. Benoit, S. Mikhaylov, and M. Laver (2011). Scaling Policy Preferences From Coded Political Texts. *Legislative Studies Quarterly 36*(1), 123–155.

Mair, P. and C. Mudde (1998). The party family and its study. *Annual Review of Political Science 1*(1), 211–229.

Mares, I. (2003). *The Politics of Social Risk.* New York: Cambridge University Press.

McCarty, N., K. T. Poole, and H. Rosenthal (2008). *Polarized America: The Dance of Ideology and Unequal Riches.* The MIT Press.

Meltzer, A. H. and S. F. Richard (1981). A Rational Theory of the Size of Government. *Journal of Political Economy 89*(5), 914–927.

Meyer, P. B. and A. N. Osborne (2005). Proposed category system for 1960–2000 census occupations. Working Paper 383, Bureau of Labor Statistics, U.S Department of Labor.

Moene, K. O. and M. Wallerstein (2003). Earnings Inequality and Welfare Spending: A Disaggregated Analysis. *World Politics 55*(4), 485–516.

Mueller, D. C. (2003). *Public choice III.* Cambridge: Cambridge University Press.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica 46*(1), 69–85.

Muthen, B. O. and T. Asparouhov (2015). Causal effects in mediation modeling: An introduction with applications to latent variables. *Structural Equation Modeling: A Multidisciplinary Journal 22*(1), 12–23.

Nelson, P. (1999). Redistribution and the income of the median voter. *Public Choice 98*(1/2), 187–194.

Page, B. I. and L. Jacobs (2009). *Class War? What Americans Really Think about Economic Inequality.* Chicago: University of Chicago Press.

Pearl, J. (2001). Direct and indirect effects. Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence.

Pearl, J., M. Glymour, and N. P. Jewell (2017). *Causal Inference in Statistics. A Primer.* Chichester: Wiley.

Persson, T. and G. Tabellini (2000). *Political Economics. Explaining Economic Policy.* Cambridge: MIT Press.

Pontusson, J. and D. Rueda (2010). The politics of inequality: Voter mobilization and left parties in advanced industrial states. *Comparative Political Studies 43*(6), 675–705.

Powell, G. B. (1982). *Contemporary democracies.* Harvard University Press.

Ratkovic, M. and D. Tingley (2017). Sparse estimation and uncertainty with application to subgroup analysis. *Political Analysis 25*(1), 1–40.

Rehm, P. (2005). Citizen support for the welfare state. WZB discussion paper 2005-02.

Rehm, P. (2009). Risks and Redistribution: An Individual-Level Analysis. *Comparative Political Studies 42*(7), 855–881.

Rehm, P. (2016). *Risk inequality and welfare states: social policy preferences, development, and dynamics.* Cambridge University Press.

Rendon, S. R. (2012). Fixed and random effects in classical and bayesian regression. *Oxford Bulletin of Economics and Statistics 75*(3), 460–476.

Robins, J. M. (2003). Semantics of causal dag models and the identification of direct and indirect effects. In P. Green, N. Hjort, and S. Richardson (Eds.), *Highly structured stochastic systems*, pp. 70–81. New York: Oxford University Press.

Robins, J. M. and S. Greenland (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology 3*, 143–155.

Rodden, J. (2010). The Geographic Distribution of Political Preferences. *Annual Review of Political Science 13*(1), 321–340.

Romer, T. (1975). Individual welfare, majority voting and the properties of a linear income tax. *Journal of Public Economics 7*, 163–188.

Rubin, D. (1996). Multiple Imputation After 18+ Years. *Journal of the American Statistical Association 91*(434), 473–489.

Rueda, D. (2007). *Social Democracy Inside Out: Partisanship and Labor Market Policy in Industrialized Democracies.* Oxford: Oxford University Press.

Rueda, D. and D. Stegmueller (2017). The effects of income expectations on redistribution preferences in western europe. Paper presented at the Workshop on Interdisciplinary Approaches to Inequality and Mobility, Duke University, April 1 2017.

Rueda, D. and D. Stegmueller (2019). *Who Wants What? Redistribution Preferences in Comparative Perspective.* Cambridge University Press.

Schlozman, K. L., S. Verba, and H. E. Brady (2012). *The unheavenly chorus: Unequal political voice and the broken promise of American democracy.* Princeton University Press.

Searle, S. R. (1971). *Linear Models.* New York: Wiley.

Shayo, M. (2009). A model of social identity with an application to political economy: Nation, class, and redistribution. *American Political Science Review 103*(2), 147–174.

Sinn, H.-W. (1995). A theory of the welfare state. *Scandinavian Journal of Economics 97*(4), 495–526.

Spiegelhalter, D. J., A. Thomas, N. Best, and W. R. Gilks (1997). *BUGS: Bayesian Inference Using Gibbs Sampling. Manual.* Cambridge: Medical Research Council Biostatistics Unit.

Stegmueller, D. (2013). Religion and redistributive voting in western europe. *The Journal of Politics 75*(4), 1064–1076.

Stekhoven, D. J. and P. Bühlmann (2011). Missforest. non-parametric missing value imputation for mixed-type data. *Bioinformatics 28*(1), 112–118.

Stigler, G. J. (1973). General economic conditions and national elections. *The American Economic Review 63*(2), 160–167.

Stimson, J. A., M. B. MacKuen, and R. S. Erikson (1995). Dynamic representation. *American Political Science Review 89*(03), 543–565.

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society B 58*(1), 267–288.

VanderWeele, T. and S. Vansteelandt (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface 2*, 457–468.

VanderWeele, T. J. (2010). Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology 21*(4), 540.

Volkens, A., P. Lehmann, T. Matthiess, N. Merz, and S. Regel (2016). The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR). Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Appendices

A1. Descriptive information

Table A1.1 on the following page shows countries and elections included in our Western European sample.

Table A1.1

Countries and election included in Western European analysis.

	Survey years								Fieldwork	Election dates
	2	4	6	8	10	12	14	16		
Austria	**x**	x	**x**				x	x	02.02.03-30.09.03; 06.01.05-25.04.05; 18.07.07-05.11.07; 14.10.14-05.05.15;19.09.16-28.12.16	24.11.02, 1.10.06, 28.09.08, 29.09.13, 15.10.17
Belgium	**x**	**x**	x	**x**	**x**	x	*x*	x	01.10.02-30.04.03; 04.10.04-31.01.05; 23.10.06-19.02.07; 13.11.08-20.03.09; 11.10.10-06.05.11; 10.09.12-24.12.12; 10.09.14-01.02.15; 14.09.16-31.01.17	13.6.99, 18.5.03, 10.6.07, 13.06.10, 25.05.14; 26.05.19
Switzerland	**x**	**x**	x	**x**	x	**x**	x	**x**	09.09.02-08.02.03; 15.09.04-28.02.05; 24.08.06-02.04.07; 30.08.08-17.04.09; 02.10.10-23.03.11; 01.09.12-22.04.13; 29.08.14-20.02.15; 01.09.16-02.03.17	24.10.99, 19.10.03, 21.10.07, 23.10.11; 18.10.15
Germany	**x**	x	**x**	x	**x**	x	x	x	20.11.02-16.05.03; 26.08.04-16.01.05; 01.09.06-15.01.07; 27.08.08-31.01.09; 15.09.10-03.02.11; 06.09.12-22.01.13; 18.08.14-05.02.15; 23.08.16-26.03.17	27.9.98, 22.9.02, 18.9.05, 27.09.09, 22.09.13; 24.09.17
Denmark	**x**	x	**x**	**x**	x	**x**	x		28.10.02-19.06.03; 20.11.02-16.05.03; 19.09.06-02.05.07; 01.09.08-11.01.09; 20.09.10-31.01.11; 10.01.13-24.04.13;12.09.14-17.02.15	20.11.01, 8.2.05, 13.11.07, 15.09.11; 18.06.15
Spain	**x**	**x**	x	**x**	x	**x**	x	**x**	19.11.02-20.02.03; 27.09.04-31.01.05; 25.10.06-04.03.07; 05.09.08-31.01.09; 11.04.11-24.07.11; 23.01.13-14.05.13; 22.01.15-25.06.15; 16.02.17-23.06.17	12.3.00, 14.3.04, 9.3.08, 20.11.11; 20.12.15
Finland	**x**	**x**	x	**x**	x	**x**	x	*x*	09.09.02-10.12.02; 20.09.04-17.12.04; 18.09.06-20.12.06; 19.09.08-05.02.09; 13.09.10-30.12.10; 03.09.12-02.02.13;03.09.14-09.02.15;15.09.16-08.03.17	21.3.99, 16.3.03, 18.3.07, 17.04.11; 19.04.15
France	**x**	x	x	**x**	x	**x**	x	x	15.09.03-15.12.03; 27.11.04-04.03.05; 19.09.06-07.04.07; 28.09.08-31.01.09; 15.10.10-06.04.11; 08.02.13-30.06.13; 31.10.14-03.03.15; 10.11.16-11.03.17	25.5./1.6.97, 9./16.6.02; 10./17.6.07; 22.04.12;23.04.17
United Kingdom	**x**	x	**x**	x	**x**	x	x	x	24.09.02-04.02.03; 27.09.04-16.03.05; 05.09.06-14.01.07; 01.09.08-19.01.09; 31.08.10-28.02.11; 01.09.12-07.02.13; 01.09.14-25.02.15 + 02.10.15-07.12.15; 01.09.16-20.03.17	7.6.01, 5.5.05, 6.5.10, 7.5.15, 8.6.17
Ireland	**x**	x	x		**x**	x	x	**x**	11.12.02-12.04.03; 18.01.05-20.06.05; 14.09.06-31.08.07; 20.09.11-31.01.12; 15.10.12-09.02.13; 04.09.14-31.01.15; 25.11.16-08.05.17	6.6.97, 17.5.02, 24.5.07, 25.02.11; 26.02.16
Italy	**x**					**x**		x	13.01.03-30.06.03; 01.06.13-20.12.13; 11.09.17-19.11.17	13.5.01, 9.04.06, 13.04.08, 24.02.13; 4.3.18
Luxembourg	**x**	**x**							14.04.03-14.08.03; 13.09.04-26.01.05	13.6.99, 13.6.04
Netherlands	**x**	**x**	x	**x**	**x**	x	x	x	01.09.02-24.02.03; 11.09.04-19.02.05; 16.09.06-18.03.07; 08.09.08-28.06.09; 27.09.10-02.04.11; 28.08.12-30.03.13; 08.09.14-15.01.15; 01.09.16-31.01.17	6.5.98, 15.5.02, 22.1.03, 22.11.06, 9.06.10, 12.9.12; 15.3.17
Norway	**x**	x	**x**	x	**x**	x	**x**	x	16.09.02-17.01.03; 15.09.04-15.01.05; 21.08.06-19.12.06; 25.08.08-20.01.09; 09.09.10-15.02.11; 14.08.12-08.02.13; 20.08.14-08.01.15; 22.08.16-17.01.17	10.9.01, 12.9.05, 13.9.09, 8.9.13; 11.09.17
Portugal	**x**	x	**x**	x	**x**	**x**	x	**x**	26.09.02-20.01.03; 15.10.04-17.03.05; 12.10.06-28.02.07; 09.10.08-08.03.09; 11.10.10-23.03.11; 24.10.12-20.03.13; 02.02.15-30.11.15; 20.10.16-15.06.17	10.10.99, 17.3.02, 20.2.05, 27.9.09, 5.6.11, 4.10.15
Sweden	**x**	x	**x**	x	**x**	x	x	**x**	23.09.02-20.12.02; 29.09.04-19.01.05; 21.09.06-03.02.07, 15.09.08-03.02.09; 27.09.10-01.03.11; 01.10.12-05.05.13; 01.08.14-30.01.15; 26.08.16-10.02.17	20.9.98, 15.9.02, 17.9.06, 19.9.10, 14.9.14

Symbols: x: country participated in ESS; **x**: survey matched to previous election and included in analysis; *x*: survey matched but not included due to missing party manifesto coding.

A2. Imputation of missing covariates

In this subsection we present details how we impute missing values of our key variables income, occupational risk, vote choice and redistribution preferences. The latter two variables are endogenous in the model and thus can be imputed at every step of the MCMC sampler from their respective model equations. For income and risk we use auxiliary equations to predict missing values. We thus implement a fully Bayesian model-based imputation approach (cf. Ibrahim et al. 2005) which properly takes into account the uncertainty arising from using imputed data. Furthermore, we make use of auxiliary information available to us by including auxiliary variables that are predictive of missingness but not part of the main model (Rubin 1996: 481).

Denote by $\mathbf{d} = d_{ik}$ our rectangular data matrix, with i indexing individuals and k indexing variables. Due to missing observations on our central right-hand-side variables (income and occupational risk), we partition $\mathbf{d}$ into observed and missing values, $\mathbf{d} = (\mathbf{d}^{obs}, \mathbf{d}^{mis})$ and create a binary indicator $\mathbf{m} = (m_{ij})$ such that $m_{ij} = 0$ if d_{ij} is observed and $m_{ij} = 0$ if d_{ij} is unobserved. Denote unknown model parameters by $\boldsymbol{\beta}$ (in the main model) and $\boldsymbol{\theta}$ (in the imputation model). The joint model likelihood for the full data is given by

$$f(\mathbf{d}, \mathbf{m}|\boldsymbol{\beta}, \boldsymbol{\theta}) = f(\mathbf{d}^{obs}, \mathbf{d}^{mis}, \mathbf{m}|\boldsymbol{\beta}, \boldsymbol{\theta}), \tag{A2.1}$$

which cannot be evaluated because it depends on missing information. However, we can obtain the *marginal* distribution of the data by integrating out missing data

$$f(\mathbf{d}^{obs}, \mathbf{m}|\boldsymbol{\beta}, \boldsymbol{\theta}) = \int f(\mathbf{d}^{obs}, \mathbf{d}^{mis}, \mathbf{m}, \boldsymbol{\beta}, \boldsymbol{\theta}) d\mathbf{d}^{mis}. \tag{A2.2}$$

Under some mild conditional independence assumptions, we can factorize our joint model as follows

$$f(\mathbf{d}^{obs}, \mathbf{d}^{mis}, \mathbf{m}|\boldsymbol{\beta}, \boldsymbol{\theta}) = f(\mathbf{m}|\mathbf{d}^{obs}, \mathbf{d}^{mis}, \boldsymbol{\theta}) f(\mathbf{d}^{obs}, \mathbf{d}^{mis}|\boldsymbol{\beta}). \tag{A2.3}$$

Here $f(\mathbf{d}^{obs}, \mathbf{d}^{mis}|\boldsymbol{\beta})$ is the same likelihood we would have specified if all data had been observed. The missing data mechanism for missing income and risk information is represented by $f(\mathbf{m}|\mathbf{d}^{obs}, \mathbf{d}^{mis}, \boldsymbol{\theta})$, which models the probability of not observing income or risk as a function of (observed and/or unobserved) covariates.

Under MAR (conditional on covariates), $f(\mathbf{m}|\mathbf{d}^{obs}, \mathbf{d}^{mis}, \boldsymbol{\theta})$ simplifies to $f(\mathbf{m}|\mathbf{d}^{obs}, \boldsymbol{\theta})$ so that

$$f(\mathbf{d}^{obs}, \mathbf{m}|\boldsymbol{\beta}, \boldsymbol{\theta}) = f(\mathbf{m}|\mathbf{d}^{obs}, \boldsymbol{\theta}) \int f(\mathbf{d}^{obs}, \mathbf{d}^{mis}|\boldsymbol{\theta}) d\mathbf{d}^{mis} \tag{A2.4}$$

$$= f(\mathbf{m}, \mathbf{d}^{obs}, \boldsymbol{\theta}) f(\mathbf{d}^{obs}|\boldsymbol{\beta}). \tag{A2.5}$$

By parameterizing the imputation model as linear additive equations (with separable errors), we arrive at two additional equations (for risk and income).

$$\mathbf{w} = \boldsymbol{\upsilon}'\boldsymbol{\delta}_w + \boldsymbol{\epsilon}_w \tag{A2.6}$$

$$\mathbf{z} = \boldsymbol{\upsilon}'\boldsymbol{\delta}_z + \boldsymbol{\epsilon}_z \tag{A2.7}$$

Equations (A2.6) and (A2.7) provide regression imputations, where $\mathbf{w}$ and $\mathbf{z}$ are vectors of income and risk, where missing values are predicted as a function of covariates in matrix $\boldsymbol{\upsilon}$ with associated regression weights $\boldsymbol{\delta}$. We include in $\boldsymbol{\upsilon}$ age, education, as well as their interaction, an indicator for gender, and information on household size and occupation (indicators for professional and routine manual occupations). We adjoin these equations to our main model. Jointly estimating imputation and outcome equations via MCMC and integrating over the posterior parameter distribution of imputed values yields model estimates that appropriately take into account the uncertainty caused by missing observations. Figure A2.1 compares the distribution of our central variables under listwise deletion and averaged over 1000 draws from the posterior (imputed) distribution.

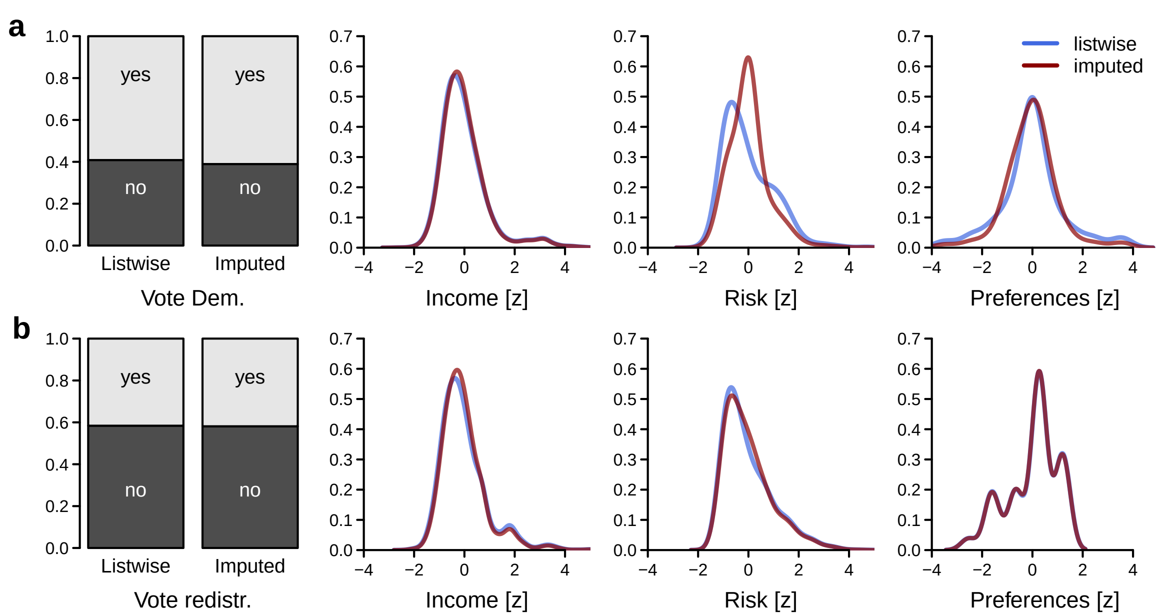

FIGURE A2.1
Distribution of listwise and imputed covariates in (a) United States and (b) Western Europe.

Table A2.1 shows two robustness tests that document that this treatment of missing values produces results comparable to alternative strategies. Panel (A) shows that our results also obtain when simply using listwise deletion of missing values.

Our treatment of missing values is ‘parametric’, in the sense that we specify a mean model and a set of covariates based on which we impute missing observations as part of the MCMC sampler. An alternative is a ‘non-parametric’ approach where we impute our data in a pre-processing step using a flexible machine learning algorithm. We generate 10 imputed data sets, where we fill-in missing continuous and categorical values values using classification and regression trees (Doove et al. 2014). One of their key advantage for the purpose of imputation is that they accommodate higher-order interactions among covariates and allow for nonlinear functional forms (Burgette and Reiter 2010). Panel (B) shows posterior summaries averaged over 10 imputed data sets. We find our results not materially changed.

Table A2.1
Further robustness checks for treatment of missing values

	Western Europe		United States	
	Income	Risk	Income	Risk
(A) Listwise deletion[a]				
Indirect effect	−1.486	0.323	−2.632	0.794
	(0.061)	(0.056)	(0.248)	(0.282)
Direct effect	−2.876	1.020	−3.007	−0.147
	(0.239)	(0.292)	(0.486)	(0.592)
(B) Nonparametric multiple imputation[b]				
Indirect effect	−1.484	0.489	−2.568	0.313
	(0.061)	(0.058)	(0.192)	(0.157)
Direct effect	−2.507	1.039	−2.509	0.345
	(0.220)	(0.276)	(0.458)	(0.529)

a Listwise deletion based on missing values of all covariates (i.e., no model based imputation).

b Multiple imputation pre-processing of all missing values using a non-parametric strategy (using Classification and Regression Trees). Entries are posterior summary of 10 multiply imputed data sets.

A3. Calculation of natural direct and indirect effects

This section describes how we calculate direct and total indirect effects (Robins and Greenland 1992; Robins 2003) of equations (3) and (2) from our model estimates obtained from equations (4) and (5). This follows directly from applying Pearl's mediation formula (Pearl 2001).[39] Write our model in simplified form with one covariate of interest (treatment), x_i, a mediating variable (preferences), m_i, and confounders, c_i. We estimate the following system of equations:

$$\text{probit}(y_i) = \beta_0 + \beta_1 x_i + \lambda m_i + \beta_2 c_i \tag{A3.1}$$

$$m_i = \gamma_0 + \gamma_1 x_i + \gamma_2 c_i + \epsilon_{2i}. \tag{A3.2}$$

with

$$\epsilon_{2i} \sim N(0, \sigma^2_{\epsilon_2}) \tag{A3.3}$$

Since our dependent variable is binary, $\text{probit}(y_i)$ is the probability of obtaining a positive response (voting Democrat), defined as

$$P(Y_i = 1|m, x, c) = \int_{-\infty}^{\text{probit}(y_i)} f(z; 0, 1)\partial z = \Phi(\text{probit}(y_i)) \tag{A3.4}$$

where $f(z; 0, 1)$ is the standard normal density, and Φ is the CDF of the standard normal distribution.

Take the general expression used in the formulas for direct and indirect effects (eq. (3) and (2)), $E(Y(x, M(x'))|C = c)$. As these quantities are not expressed conditional on M, we need to integrate over M:[40]

$$E(Y(x, M(x'))|C = c) = \int_{-\infty}^{\infty} E(Y|C = c, X = x, M = m) \times f(M|C = c, X = x')\partial M \tag{A3.5}$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\text{probit}(y_i)} f(z; 0, 1)\partial z \times f(M; \gamma_0 + \gamma_1 x' + \gamma_2 c, \sigma^2_{\epsilon_2})\partial M \tag{A3.6}$$

$$= \int_{-\infty}^{\text{probit}(x,x')} f(z; 0, 1)\partial z. \tag{A3.7}$$

Here, $\text{probit}(x, x')$ is given by:

$$\text{probit}(x, x') = [\beta_0 + \beta_1 x + \beta_2 c + \lambda(\gamma_0 + \gamma_1 x' + \gamma_2 c)]/\sqrt{var(x)} \tag{A3.8}$$

[39] Imai et al. (2010, 2011) call these average causal mediated effects for the treated and average direct effects for the control, while Pearl (2001) calls them total natural indirect effects and pure natural direct effects. See Imai et al. (2010) and Muthen and Asparouhov (2015) for an extended discussion on their computation.

[40] The last equality is obtained by variable transformation and a change of order of integration.

where the variance $var(x)$ is given by

$$var(x) = \lambda^2 \sigma^2_{\epsilon_2} + 1. \tag{A3.9}$$

Indirect effect Denote two values of a treatment by x and x' (e.g., low vs. high income). The (natural) indirect effect (eq. 2) is :

$$E[(Y(x', M(x')) - Y(x', M(x))|C] = \tag{A3.10}$$
$$\int_{-\infty}^{\infty} E[Y|C = c, X = x', M = m] \times f(M|C = c, X = x')\partial M \tag{A3.11}$$
$$- \int_{-\infty}^{\infty} E[Y|C = c, X = x', M = m] \times f(M|C = c, X = x)\partial M. \tag{A3.12}$$

Expressed in terms of equation A3.4 we calculate the indirect effect as:

$$\Phi(\text{probit}(x', x')) - \Phi(\text{probit}(x', x)) \tag{A3.13}$$

This is equivalent to the formula given in Imai et al. 2010, appendix F.

Direct effect The (natural) direct effect (eq. 3) is

$$E[Y(x', M(x)) - Y(x, M(x))|C] = \tag{A3.14}$$
$$\int_{-\infty}^{\infty} (E[Y|C = c, X = x', M = m] - E[Y|C = c, X = x, M = m]) \times f(M|C = c, X = x)\partial M. \tag{A3.15}$$

Expressed in terms of equation A3.4 it is calculated as:

$$\Phi(\text{probit}(x', x)) - \Phi(\text{probit}(x, x)). \tag{A3.16}$$

Substantive interpretation of these quantities rests on a number of assumptions. We discuss these and conduct sensitivity analyses (Imai et al. 2010).

Exposure mediator interaction The model outlined above does not include a possible interaction between the treatment and preferences. This can be handled straightforwardly by changing equation (A3.1) to

$$\text{probit}(y_i) = \beta_0 + \beta_1 x_i + \beta_2 x_i m_i + \lambda m_i + \beta_2 c_i$$

and adapting the calculation of NIE and NDE accordingly. However, in our application, we find no evidence of the existence of a mediator-outcome interaction (for both income and risk).

A4. Recall vote in the European Social Survey

Our measure of vote choice is based on respondents recalling their vote cast in the last election. This introduces the possibility of recall bias, be it due to simple errors in memory or respondents' tendency to 'adjust' their reported choices to contemporary preferences. In addition, due to the different timing of elections in each country, some respondents have to recall choices made further back in time.

To get a sense of how this issue might affect our results, we compare party preferences based on the recall of *past* choices with stated *contemporary* party preferences at the time of each survey. The ESS contains an item that invites respondents to express to which party they feel particularly close to. We classify each respondent's closest party as either redistributive or non-redistributive. Table A4.1 shows the distribution of respondents recalling having voted for a redistributive party in the last election against feeling close to a redistributive party at the time of the ESS interview. It reveals that about 94 percent of all responses are consistent between the two methods of assessment. 95.5% of respondents who recall having voted for a redistributive party also feel close to one currently. Conversely, 93.5% of respondent who recall not having voted for a redistributive party presently feel close to a non-redistributive party.

Table A4.1
Distribution of redistributive party vote based on recall and party closeness. Row percentages, cell percentages in brackets.

		Feel close: redistr. party	
		no	yes
Vote recall: redistr. party	no	93.5 [51.4]	6.5 [3.6]
	yes	4.5 [2.0]	95.5 [43.0]

Note: Closeness based on item prompting respondents to identify the party they feel closest to (at time of survey). Respondents identifying no close party are excluded. Vote recall refers to last election. Cramer's ϕ = 0.89, χ^2 (1df) p = 0.000.

In Table A4.2, we show direct and indirect effects estimates using both vote recall (as in the main text) and party closeness as left-hand-side variables. For our key indirect effect estimates, we find that both estimates for income and risk are slightly larger when using a respondent's closest party. Conversely, direct effect estimates are smaller, little in the case of income, much more so in the case of occupational risk.

Table A4.2
Comparison of estimates using vote recall and closest party as outcome

	Vote recall		Party closeness	
	Income	Risk	Income	Risk
Indirect effect	−1.578	0.374	−1.926	0.451
	(0.064)	(0.059)	(0.072)	(0.070)
Direct effect	−3.038	1.097	−2.753	0.896
	(0.238)	(0.285)	(0.286)	(0.375)

A5. Immigration

One possible concern with our analysis is that our measure of preferences conflates 'true' preferences for redistributive policies with other non-economic attitudes that are nonetheless related to the provision of social insurance. One such concern relates to attitudes towards immigrants and immigration. The increasing levels of immigration in Western Europe and the US (and their politization) have been associated with growing concerns about competition by the native populations (see, for example, Andersen and Bjørklund 1990, Faist 1994 or De Koster et al. 2013). This process, often referred to as "welfare chauvinism," can usefully be described as "the fear among groups in the native population (and settled immigrants) that certain new immigrant groups take away jobs, housing and social services" (Faist 1994: 440). To address this concern, we estimate models that explicitly allow (measured) immigration attitudes to shape individuals' redistribution preferences.

We change our preference equation to allow for the possibility that individuals' stated redistribution preferences partly depend on their anti-immigration attitudes:

$$R_{ijt} = \beta_1 w_{ijt} + \beta_2 z_{ijt} + x'_{ijt}\delta^R + \xi_{jt} + \phi'_i\kappa + \epsilon^R_{ijt}. \tag{A5.1}$$

where, ϕ_i is a low-dimensional vector of (anti-) immigration preferences and κ are coefficients capturing how much they affect redistribution preferences. We estimate ϕ from a measurement system for K observed immigration items M_{ik} $(k = 1, \ldots, K)$ available in each ESS survey:

$$M_i = \Lambda\phi_i + \omega_i, \quad \phi \sim N(0, I_D) \tag{A5.2}$$

Here, M_i is a $K \times 1$ vector of observed immigration attitudes, ϕ_i is a length-D vector of latent factors, Λ is a $K \times D$ matrix of "factor loadings", and ω_i is a length-K vector of residuals. We assume residuals are independently normally distributed with freely estimated variances. The latent factors are distributed normal with and their variance is fixed to 1 in order to identify the scale of the model. Note that (A5.2) does not include intercepts, therefore we transform every M_k to have mean zero before estimating the model.

In the application below we specify models with latent factor dimensions of $D = 1$ and $D = 2$. In the latter case we set $Cov(\phi_1, \phi_2) = 0$ for ease of interpretation. We adjoin the system in (A5.2) to our main model. By jointly estimating both we ensure that the uncertainty in estimating ϕ is taken into account (avoiding errors-in-variables bias in capturing the impact of immigration attitudes on stated redistribution preferences).

Table A5.1
Adjusting for immigration attitudes as confounders of stated redistribution preferences in the European Social Survey.

	Income	Risk
(A) 1-factor model [a]		
Indirect effect	−1.613	0.425
	(0.064)	(0.057)
Direct effect	−3.028	1.096
	(0.236)	(0.291)
(B) 2-factor model [b]		
Indirect effect	−1.625	0.454
	(0.065)	(0.058)
Direct effect	−3.027	1.094
	(0.239)	(0.288)

a $D = 1$. Items included: country's cultural life undermined by immigrants, immigrants make country worse place to live, immigration bad for country's economy. $\kappa_1 = 0.052$ (0.005).
b $D = 2$. Additional items: admissions of immigrants from different race/ethnic groups, same race/ethnic group, poor countries outside EU. $\kappa_1 = 0.021$ (0.005), $\kappa_2 = -0.054$ (0.005).

Table A5.1 shows direct and indirect effect estimates for income and risk. Panel (A) shows results for a model where ϕ is one-dimensional and summarizes survey items probing to what extent an individual agrees with the statements that (i) his/her country's cultural life is undermined by immigrants, (ii) immigrants make the country a worse place to live, (iii) immigration is bad for his/her country's economy. We find that this summary measure of attitudes towards immigration impacts preferences significantly: κ is estimated as 0.052 with a standard error of 0.005. However, even when accounting for immigration attitudes being part of individuals' stated preferences, we find clear evidence for the indirect effect of both income and risk on vote choice. The effect of income on choices via preferences is 1.6 (±0.06) percentage points, such that the preference channel alone accounts for 35% of the total effect of income. For occupational risk the corresponding figure is 0.43 (±0.06) percentage points accounting for 28% percent of the total risk effect. In Panel (B) we move to a model where ϕ is two-dimensional including three additional survey items capturing attitudes towards immigration on the second dimension: if his/her country should admit more immigrants from different race/ethnic groups; from the same race/ethnic group; from poor countries outside of the European Union. In this model, the first dimension is still positively related to redistribution preferences, $\kappa = 0.021 \pm 0.005$. The second, "anti-admission", dimension is strongly negatively related to stated preferences ($\kappa = -0.054 \pm 0.005$. Our findings regarding the role of preferences in linking income and risk to vote choice remain virtually unchanged under this more flexible specification.

Our United States sample contains more limited information on attitudes towards immigration. One survey item available at a reasonable number of waves probes respondents' attitudes towards increasing or decreasing the number of admitted immigrants.[41] This item is not available in ANES surveys conducted between 1982 and 1990 and in 2002. Since responses are missing for complete yearly cross-sections in the data (as opposed to a subset of rows due to item non-response), we treat their occurrence as missing-at-random conditional on covariates (Gelman et al. 1998). We fill in missing immigration attitudes from a set of observable characteristics using chained random forests (with 500 trees and 10 variables available at each split) as proposed by Stekhoven and Bühlmann (2011). The set of covariates includes age, gender, income and education, race, marital status, social class, labor market status, union membership, urban/rural location, household size, church attendance, and political interest and using a tree-based algorithm allows for flexible interactions of these.

Table A5.2 shows direct and indirect effect estimates for income and risk when also adjusting for respondents' immigration attitudes.

[41] "Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same, decreased a little, decreased a lot?"

Table A5.2
Adjusting for immigration attitudes as confounders of stated redistribution preferences in the ANES.

	Income	Risk
Indirect effect	−2.441	0.914
	(0.231)	(0.245)
Direct effect	−2.609	0.112
	(0.443)	(0.511)

Note: Missing immigration attitudes imputed using chained random forests.

A6. Social class

As we have argued in the main text, we do not explicitly include the role of social class in our analyses. We consider social class an important factor structuring the distribution of both income and risk. Panel (A) of Table A6.1 adds some empirical evidence to this view. It shows means and interquartile ranges for income and risk for social classes in Western Europe using the detailed 9-class version of the European Socio-Economic Classification. The clear relationship between class and income and risk is readily apparent. Moving down the class ladder, average income decreases monotonically, with the income of lower sales, service, and technical as well as routine manual occupations being clearly below the (country-election specific) mean. The same monotonic pattern holds for risk: lower classes face increasing levels of risk. The average risk of unemployment for a lower supervisor is more than twice as high as that of a lower manager or professional.

Table A6.1
Distribution of income and risk in different social classes.

	Income distance		Occupational risk	
(A) Means and IQRs				
	Mean	IQR	Mean	IQR
1: Higher managers / professionals	1.870	3.860	0.026	0.028
2: Lower managers / professionals	0.690	2.930	0.031	0.028
3: Intermediate occupations	0.185	2.440	0.052	0.045
4/5: Small employers & self-employed	0.076	2.740	0.069	0.083
6: Lower supervisors / technicians	−0.060	2.490	0.070	0.079
7: Lower sales and service	−0.512	2.240	0.076	0.065
8: Lower technical	−0.682	1.780	0.099	0.098
9: Routine	−0.959	1.950	0.116	0.078
(B) ANOVA				
Intra-class correlation	0.11		0.34	
	[0.05, 0.35]		[0.18,0.68]	

Note: European Socio-Economic Classification. ICC confidence confidence calculated following Searle (1971).

However, beyond this aggregate patterns lies considerable variability. Looking at the width of the interquartile ranges makes clear that there is large heterogeneity over individuals *within* classes for both income and risk. As an analysis of variance decomposition in panel (B) shows, only 11% of the total variation of income is due to social class. The corresponding number for occupational risk is 34%. Figure A6.1 makes this point in graphical form. It plots density estimates for income and risk separately for nine social classes. It is readily apparent that the differences in income and risk between individuals determine the overall shape of the distribution (even though class does structure the location and variance of income and risk profiles). For example, while a seizable share of the income of lower managers and professionals is $25,000 above the median, a large share (about less than half) have incomes at the mean or lower. This internal heterogeneity is visible for other classes as well. The differences between class-specific risk distributions are somewhat more marked, but we find seizable overlap between very dissimilar classes (e.g., compare the overlapping area of class 1 and class 9).

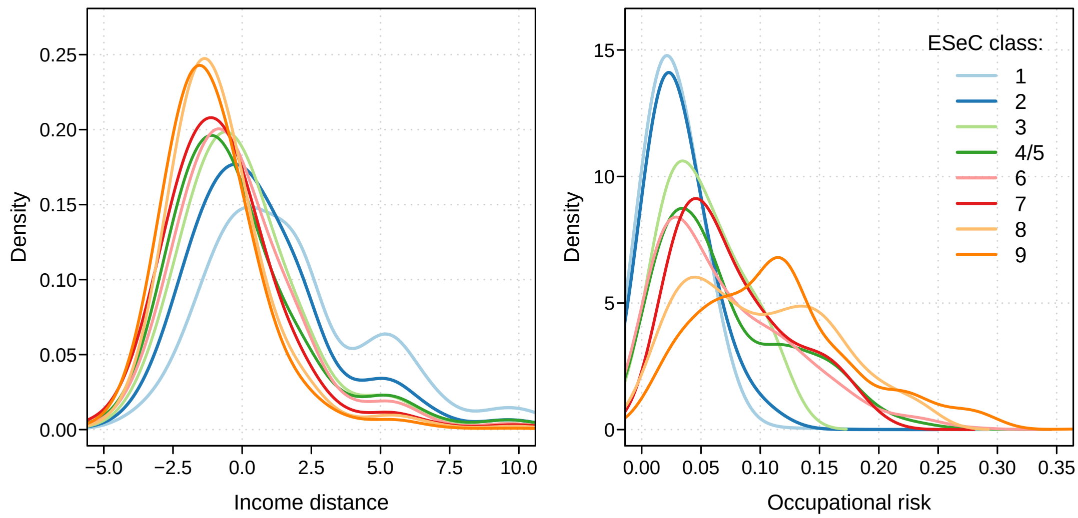

FIGURE A6.1
Density estimates of income (left) and risk (right) in nine social classes in Western Europe.

In conclusion, we accept that social class is an important variable structuring individuals' live chances, which includes, among others, their income and exposure to labor market risks. However, in this paper, we prefer to make use of the full distribution of risk and income, which represents many sources of heterogeneity (e.g., ability) other than class.

A7. Occupational unemployment rates

Due to sample size constraints, our measure of occupational unemployment rates in our Western European sample is constant over time. While this minimizes errors-in-variables problems, one possible concern is that it fails to capture changes in the structure of occupational risks induced by the great recession. To investigate the robustness of our results, we calculate a two-period measure of occupational risk splitting the EU-SILC samples at 2008. This captures the different patterns in risk in the early 2000s and the 2010s. The correlation between the two periods is 0.76.

Table A7.1 shows direct and indirect effects of income and risk on the vote for redistributive parties. We find that our core results remain substantively similar. Estimates for income are quantitatively close to those reported in Table I. The indirect effect estimate for risk is reduced

Table A7.1
Direct and indirect effects of income and risk using a two-period measure of occupational unemployment rates.

	Income	Risk
Indirect effect	−1.590	0.272
	(0.067)	(0.060)
Direct effect	−3.132	0.675
	(0.242)	(0.304)

Note Western European Sample. LHS variable: redistributive party vote. Occupational risk measure calculated before and after 2008.

by 0.1 points. However, its direct effect is similarly reduced. Thus, the overall proportion of the impact of risk mediated by preferences remains substantively unchanged. Given these results, we opt to rely on the more precisely estimated time-constant unemployment measures in the main text.

A8. Post-double-selection LASSO estimation

To relax our modeling assumptions, we report robustness test (3) in Table III that builds on the double-post-selection strategy proposed by Belloni et al. (Belloni et al. 2013, 2017). Specifically, this model setup aims to reduce the possible impact of omitted variable bias by accounting for a large number of confounders in the most flexible way possible. This can be achieved by moving beyond restricting confounders to be linear and additive, and instead considering a flexible, unrestricted (non-parametric) function. This leads to the formulation of the following partially linear model (we omit FEs and subscripts for grouping structures for notational parsimony)

$$V_i^* = \alpha R_i + \gamma D_i + g(x_i) + e_i, \quad E(e_i | D_i, x_i = 0) \tag{A8.1}$$

Here, V_i^* is the vote propensity of each respondent and $D_i = \{w_i, z_i\}$ are the "treatments" income and occupational risk. The function $g(x_i)$ captures the possibly high-dimensional and nonlinear influence of confounders. The utility of this specification as a robustness test stems from the fact that it imposes no a priori restriction on the functional form of confounding variables. A second key ingredient in a model capturing biases due to omitted variables is the relationship between the treatment(s) and confounders. Therefore, we consider the following auxiliary treatment equations

$$D_i = m(x_i) + u_i, \quad E(u_i|x_i = 0) \tag{A8.2}$$

which relates treatment to a set of covariates x_i. The function $m(x_i)$ summarizes the confounding effect and creates omitted variable bias.

The next step is to create approximations to both $g(\cdot)$ and $m(\cdot)$ by including a large number (p) of control terms $q_i = P(x_i) \in \mathbb{R}^p$. These control terms can be transforms of covariates, higher order interaction terms, etc. Even with an initially limited set of variables, the number of control terms can grow large, say $p > 200$. To limit the number of estimated coefficients, we assume that g and m are approximately sparse (Belloni et al. 2013) and can be modeled using s non-zero coefficients (with $s \ll p$) selected using regularization techniques, such as the LASSO (see Tibshirani 1996; see Ratkovic and Tingley 2017 for a recent exposition in a political science context):

$$V_i^* = \alpha R_i + \gamma D_i + q_i \kappa_{g0} + r_{gi} + e_i \tag{A8.3}$$

$$D_i = q_i \kappa_{m0} + r_{mi} + u_i \tag{A8.4}$$

Here, κ_{g0}, and κ_{m0} are coefficient vectors for the selected covariates and r_{gi} and r_{mi} are approximation errors.

However, before proceeding we need to consider the problem that variable selection techniques, such as the LASSO, are intended for prediction, not inference. In fact, a "naive" application of variable selection, where one keeps only the significant q variables in equation (A8.3) fails. It relies on perfect model selection and can lead to biased inferences and misleading confidence intervals (see Leeb and Pötscher 2008). Thus, we express our problem as one of prediction by substituting the auxiliary treatment equation (A8.4) for D_i in equation (A8.3) yielding a reduced form equation so that now *both* equations in this system are amenable to high-dimensional selection techniques.

Note that using this two equation setup is also necessary to guard against variable selection errors. To see this, consider the consequence of applying variable selection techniques to the vote equation only. In trying to predict V with q, an algorithm (such as LASSO) will favor variables with large coefficients but will ignore those of intermediate impact. However, omitted variables that are strongly related to one or both of the treatments can lead to large omitted variable bias in the estimate of γ even when the size of their coefficient in the outcome equation is moderate. The Post-double selection estimator suggested by Belloni et al. (2013) addresses this problem, by basing selection on *both* reduced form equations. Let $\hat{\mathcal{W}}_1$ be the set of controls selected by LASSO of V_i on q_i; and let $\hat{\mathcal{W}}_2$ be the set of controls selected by LASSO of D_i on u_i. Then, the set of control variables, $\hat{\mathcal{W}}$, used in our analysis reported in specification (3) of Table III is constructed by $\hat{\mathcal{W}} = \hat{\mathcal{W}}_1 \cup \hat{\mathcal{W}}_2$. Note that this strategy is robust to moderate selection mistakes. (Belloni et al. 2014).[42]

Responsible for the usefulness of this robustness check is the indirect LASSO step selecting the D-control set. It finds controls whose omission leads to “large” omitted variable bias and includes them in the model. Any variables that are not included (“omitted”) are therefore at most mildly associated to D_i and V_i^*, which decidedly limits the scope of omitted variable bias (Chernozhukov et al. 2015).

[42] For a very general discussion see Belloni et al. (2017).