Class Voting and Economic Policy Preferences: A Predictive Modelling Approach

Mustafa Kaba *

Abstract

Economic policy preferences are often thought to be less class-based over time due to improved living standards. Yet, there is no clear way to measure the degree of class distinctiveness in economic preferences and how it varies over time and space. In this paper, I introduce a novel measure of class distinctiveness in economic policy preferences based on predictive modeling and estimate it for 18 European countries at three different points in time. I then validate this new measure and explore its implications for class-based voting.

JEL Codes: C38, C51, C52, D72

Keywords: class divisions, economic preferences, party positions, predictive modeling, multilevel modeling

*Max Planck Institute for Research on Collective Goods, Bonn, `kaba@coll.mpg.de`. I thank David K. Levine, Arthur Schram, Geoffrey Evans, Daniela Iorio, Cemal Eren Arbatlı and Serkant Adigüzel for their helpful comments and suggestions.

1 Introduction

In the wake of rising extremism and an unprecedented populist surge, the discussion of social classes seems to have regained strength in political debate. One implication of this recent trend is that it is now even more critical to understand the relationship between classes and voting decisions of people. The recent political events such as Brexit, Donald Trump's victory in the 2016 US presidential election, and AfD's recent entry into the German parliament have been argued to be related to class politics (Guiso et al., 2017). Moreover, the scholarly research has documented that the preferences of the working class have become better-aligned with promises of the far-right parties (Oesch and Rennwald, 2018; Spies, 2013). These events and the related scientific evidence suggest that classes still play an important role in today's politics. Accordingly, a useful concept for gauging the relationship between social classes and voting decisions has been found in *class voting*.

The class voting refers to the tendency of individuals in a given social class to vote for a particular political party rather than the alternatives, compared to voters in other classes (Arzheimer et al., 2016). While much debated due to problems in conceptualization and measurement, both class and class voting have received substantial attention from the scholarly world. Several studies have attempted to measure the changes in the class-vote linkage over time (Evans, 2000, 1999). More recent studies on class voting, on the other hand, have also focused on the mechanisms that cause variations in the extent of class voting across countries and over time (Jansen et al., 2013; Evans and Tilley, 2012a,b). One of these mechanisms is the *blurring of class divisions in terms of economic preferences*. Nevertheless, although widely accepted, no empirical evidence has yet been presented for either the blurring of class divisions or its relationship with class voting. I therefore first develop a measure that captures the extent of class divisions in economic preferences and test whether blurring of divisions has ensued. Second, I test the statistical relationship between class distinctiveness in economic preferences and class voting.

The previous literature on class voting has put forward two main explanations for the variations in class voting: the supply-side and the demand-side. The former concerns the range of political choices offered by political parties to voters. The main hypothesis of this strand of the literature is that if parties offer similar policies on the economic policy dimension, then voters are less likely to base their voting decisions on their class membership. These studies rely on the assumption of voter responsiveness to party programs. The typical argument is as follows: when voters are responsive to party programs but the political parties have converged to the same economic programs, then the voting decisions are aligned with class membership to a lesser extent. Studies in this strand of literature commonly measure program convergence by parties on the economic dimension over time using the Manifesto Project or expert surveys. They test whether the convergence or polarization on the economic dimension by parties is associated with the strength of class voting. Evans and Tilley (2012a), Evans and Tilley (2012b) provide evidence from Great Britain in favor of this hypothesis. Evans and De Graaf (2013) study the same matter in a cross-country perspective. In this article, I do not focus on the supply-side argument, which is already well-established.

The demand-side approach, on the other hand, concerns the class structure of society and political preferences of distinct socio-economic classes. The central hypothesis puts forward that when the class structure becomes more diffused, or more specifically, when classes become less distinct in terms of their economic preferences, we should expect to observe that the voting choices are less based on class membership. Throughout this paper, I use the term *class distinctiveness* to refer to the extent to which we can distinguish between distinct classes using their economic preferences. Accordingly, a lower class distinctiveness implies the blurring of class divisions.

The hypothesis of class divisions becoming increasingly indistinct is not new. The diffusion of class structure and the resulting blurring of class divisions in economic preferences are usually thought to arise from the transition to a post-industrial society with more educated people, higher living standards, and more social mobility (Arzheimer et al., 2016). Early studies such as Inglehart and Rabier (1986) and Clark and Lipset (1991) had already announced the death of class as being of interest in the study of electoral politics. The same hypothesis have also mentioned in more recent studies and furthermore has managed to make its way to the standard comparative politics textbooks such as *Citizen Politics* by Dalton (2013).

Although the demand-side hypothesis has been put forward in several studies, to the best of my knowledge, no empirical evidence has yet been presented on either the blurring of class divisions or on the nature of its relationship with class voting. The only study that provides such evidence is by Evans and Tilley (2017), who show that class distinctiveness has been stable over time in Britain.[1] Along these lines, Evans and Tilley (2012a) point out this gap

[1]There are, however, studies that document evidence for the poor performance of socio-economic classes to predict economic preferences at a single time point. A recent example is Yagci et al. (2020) who show

in the literature as follows: "Conclusions concerning the impact of blurring and fracturing of the class structure on political choice have usually been inferred, retrospectively, from an observed decline in class voting, rather than measured independently and then used to account for such declines. [...] We did not, however, test directly whether a decline in the effects of class position on values and preferences can account for a decline in the effect of class on party choice."

In this paper, I fill this gap in the class voting literature by studying whether the changes in class distinctiveness in terms of economic preferences can account for the variation in the strength of class-vote linkage over time in a comparative perspective. A measure of class distinctiveness, however, is not readily available in standard survey data. To create such a measure, I adapt the empirical framework of Bertrand and Kamenica (2018) who utilized a simple idea from predictive modeling, which is using the ability to infer a person's true cultural identity from his/her preferences as a measure of the extent to which one can distinguish between cultural identities. I discuss the adaptation of this method to the current setting and its execution in detail in Section 3.

The method I adopt does not only produce a measure that reflects the extent to which the classes are distinct from each other in terms of economic preferences but also allows a meaningful comparison of class distinctiveness across countries and over time. Moreover, the method is scalable and can be implemented with any prediction model. Although I work with a limited number of explanatory variables in this study, this can be increased in other settings. The scalability is a powerful feature because it enables inferences that would otherwise be impossible in settings where a large number of explanatory variables are available and more sophisticated prediction models are required.

In this study, I limit the preference space -in which the class divisions lie- to economic preferences, rather than using all sorts of political preferences mainly for two reasons.[2] First, in the class voting literature, a class is commonly defined based on occupation and the implications of occupation in terms of economic outcomes. It is then natural to expect

that income and occupational status appear to be poor predictors of economic preferences compared to the religious or ethnic characteristics in the Turkish electorate.

[2]The other two sets of political preferences are *grid* and *group* dimensions as described by Kitschelt and Rehm (2014), corresponding to preferences related to, respectively, governance and polity membership. Measuring the class distinctiveness in terms of these two sets of preferences could likewise be interesting and promising in the study of class voting. It might especially shed some light on the evolution of the salience of different political dimensions, which is also closely related to class voting. It is, however, beyond the scope of this paper.

that any clustering of preferences within classes should primarily occur within the realm of economic preferences. The second reason is the common practice in the class voting literature that is to use political parties' economic left-right positioning rather than general left-right positions. This is because the left and right are well-defined on the economic dimension. In contrast, general left-right positions conflate the positions taken in the economic and cultural dimension by political parties.

The rest of the paper proceeds as follows: Section 2 introduces the data sets, the operationalization of the class membership variable and party positions, and finally, the hypotheses to be empirically tested. Section 3 lays out the methodology for the computation of the class distinctiveness measure. Section 4 discusses the main results of the paper. Section 5 concludes.

2 Data & Operationalization

In this study, I combine two different data sets. I first obtain individual level information on occupation, economic policy preferences, party choices, and demographic status from the European Values Study (EVS) data set described in detail in the next section. Second, I obtain the economic left-right positions of political parties from the Manifesto Project, which is described in Section 2.2. I then match these two data sets based on the reported party choice of the respondents in the EVS data set. The hypotheses that are put to empirical tests are explained in Section 2.3.

2.1 Micro-level Survey Data

The European Values Study (EVS) is one of the richest data sources for individual level data on economic preferences, political attitudes, and party choices. It covers several European countries and includes socio-economic and demographic information, such as income, occupation, education, age, and gender. In this study, I use the 1990, 1999, and 2008 waves of the EVS. The countries included in my sample are Austria, Belgium, Bulgaria, Czech Republic, Germany, Denmark, Spain, France, Great Britain, Hungary, Ireland, Italy, Netherlands, Poland, Portugal, Sweden, Slovenia, and Slovakia.

A significant advantage of the EVS data is that it is a harmonized data set in the sense that the same core questions are asked in all countries and waves with the same wording.[3] This enables truthful comparisons between countries and across time, which is not always the case for survey data sets, for example, when combining the national election studies of different countries (Boxell et al., 2020).

In particular, I include the following variables from the EVS to measure class distinctiveness and the class-vote linkage: attitude towards government responsibility, choice over freedom or equality, confidence in labor unions, party choice, occupation, age, gender, and education. The party choice variable is the key variable that enables a connection between the EVS and Manifesto Project data, the latter being the data set that provides political parties' policy stances on a number of issues. For each country in each wave in the sample, I match the individual level EVS data to the parties encoded in the Manifesto Project based on the party choice variable in the EVS.

A second advantage of the EVS data set is that the occupation variable is encoded according to the International Standard Classification of Occupations (ISCO). Hence it is convertible to the Erikson-Goldthorpe class schema (EGP) (Erikson and Goldthorpe, 1992; Ganzeboom and Treiman, 1996). The EGP classification of occupations is the dependent variable of the primary analysis that deals with the measurement of class distinctiveness in the next section.

The Erikson-Goldthorpe classification is constructed based on occupations and the characteristics of these occupations. It classifies occupations into social classes by considering dimensions, such as job security, level of earnings and the way they are earned, promotion prospects, autonomy at work, and working conditions. Evans and Tilley (2017) note that the EGP class schema has been "consistently shown to be related to differences in employment conditions, job autonomy, income, and life-time expected earnings".[4] The EGP has also been adopted in the National Statistics Socio-economic Classification of the U.K. Census (Rose and Pevalin, 2002).

For statistical power concerns, I use the four-class version of the EGP in the same spirit as Jansen et al. (2013) rather than the versions with a larger number of classes.[5] The four classes

[3]The exact wording of the questions and response scales of the EVS variables are given in Appendix B.

[4]A more comprehensive study of the validity of the EGP class schema is provided by Evans (1992).

[5]This classification follows closely the classification in Evans (1992), where he also provides supporting evidence for its validity. The eleven-class full version of EGP and its conversion into a four-class version is provided in Appendix C.

I arrive at are the service class, the routine non-manual working class, the self-employed, and the manual working class (or *working class* as a shorthand). The share of classes within each country at each time point is given in Appendix A.[6] As an example, the typical occupations in these four classes are as follows: Office managers, business professionals, health professionals, legal professionals in the service class. Clerks, salespersons in the routine non-manual working class. Small entrepreneurs, own-account workers in the self-employed class. Machine operators, craft workers in the manual working class.

2.2 Party Positions

I use the Manifesto Project data set to recover the economic left-right positions of political parties. The unit of observation in this data set is a political party belonging to a specific country and election year.

The Manifesto Project uses text analysis techniques to recover how much a political party mentions a particular policy issue in its party manifesto. The project covers 50 countries and spans the period from 1945 up to date. The data set includes a large number of policy issue categories such as freedom and democracy, political system, economy, welfare, and quality of life. The reported numbers for the variables in these categories, however, are not the positions of parties on these issues but the *emphasis* –the share of a particular policy issue in the entire manifesto– that parties put on the respective issue in their manifesto. The Manifesto Project, therefore, does not provide us with a readily available measure of the economic left-right positions of parties.

To create such a measure, I first choose the policy issues that have been identified as useful in the calculation of party positions on the economic dimension by Bakker and Hobolt (2013), who group these policy issues into two categories: the left-wing and right-wing. I then calculate the *total share of right* and *total share of left* emphases in each manifesto by simply summing the emphases of the left-wing and right-wing issues. The difference between the total share of right and left emphases then yields the position of a party on the economic dimension. The share of left-emphases (right-emphases) in a party manifesto can be minimum zero if the party does not mention left-wing (right-wing) economic issues at all in its manifesto and can be maximum one if it only mentions left-wing (right-wing) economic

[6] The heterogeneity of classes in terms of their economic preferences and descriptive statistics of these variables are also reported in Appendix A.

Table 1: Variables from the Manifesto Project

Left Emphases	Right Emphases
regulate capitalism	free enterprise
economic planning	economic incentives
pro-protectionism	anti-protectionism
social services expansion	social services limitation
education expansion	economic orthodoxy: positive
nationalization	labour groups: negative
controlled economy	
labour groups: positive	
corporatism: positive	
Keynesian demand management: positive	
Marxist analysis: positive	
social justice	

Note: The table lists the variables used in Bakker and Hobolt (2013) to calculate the economic left-right positions of political parties. Left emphases refer to left-wing economic policy emphases as a share of total emphases in a party manifesto. Similarly, right emphases refer to right-wing economic policy emphases as a share of total emphases in a party manifesto.

issues. The variables used calculate party positions on the economic dimension are listed in Table 1.

The emphases profiles of party manifestos across countries may vary largely due to country-specific factors such as the party system. For example, it could be that in country A, the economic left-right position of parties (the difference between the total right-emphases and left-emphases) ranges from -1 to +1, in contrast to country B, where it ranges from -0.5 to +0.5. These positionings, however, do not necessarily mean that the party with -1 (+1) economic position in country A is more left-wing (right-wing) than the party with -0.5 (+0.5) economic position in country B. Such a conclusion requires a more in-depth comparison of these two countries in many aspects. To refrain from such hefty arguments and since our unit of observation is at the country-year level, I opt for considering only the ranking of parties on the economic dimension within a country for a given year, but not relative to the positions of parties in other countries. Accordingly, I standardize the party positions within a country-year.

I finally match the Manifesto Project data set to the EVS data set using the preferred political party and year variables. For each political party reported as a response to party choice variable in the EVS data set, the closest available election year for that party in the

Manifesto Project is chosen for matching.

2.3 The Hypotheses

In this study, I test two hypotheses. Both hypotheses have been put forward in previous studies but have not been empirically tested. I first test whether blurring of class divisions (in economic preferences) have taken place by quantifying the relationship between the class membership of respondents and their economic preferences. I then test the statistical association between the class voting and class distinctiveness in economic preferences. A graph of this operationalization is given in Figure 1 below.

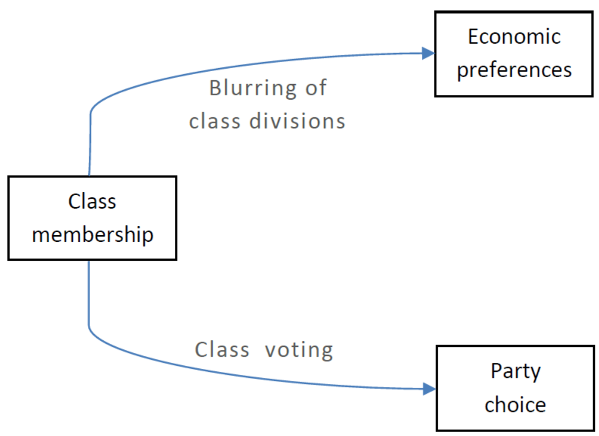

Figure 1: Class voting and blurring of class divisions

Note: The blurring of class divisions is measured by how well the economic preferences predict class membership. The class voting is measured by the strength of the statistical relationship between class membership and the economic left-right position of the preferred political parties by the respondents.

Accordingly, the first hypothesis concerns the development of distinctiveness of classes in terms of their economic preferences over time:

H1: *Classes have become less distinct in terms of their economic preferences over time.*

The first hypothesis exclusively concerns the evolution of class distinctiveness over time, and it is silent on the relationship between class distinctiveness and class voting. In order to shed light on the latter, I also test the following hypothesis:

H2: *The strength of class voting is positively related to class distinctiveness. In other words, the more distinct the classes are, the stronger the class voting is.*

An alternative operating channel in Figure 1 may be preference formation by political parties. This channel could be represented by an arrow from party choice to economic preferences in Figure 1, symbolizing that the partisanship itself forming the economic preferences. Nevertheless, recent works show that the economic preferences –such as those studied in this paper– are more stable than partisanship, and furthermore, their lagged values predict partisanship (Evans and Neundorf, 2018). Moreover, we are interested in the class divisions in economic divisions regardless of its sources, among which there may be partisanship, higher living standards, aggravated economic grievances, etc. The preference formation by parties, therefore, does not pose a threat for quantifying the predictive power of economic preferences for class membership.

3 Measuring Class Distinctiveness

The empirical strategy underlying the computation of the class distinctiveness measure originates from the study of Gentzkow et al. (2018), where they examine the extent of partisanship in the U.S. Congress. To derive a measure of partisanship, they use the texts of congressional speeches from 1873 to 2016. They define the partisanship of a speech in a given session as the predictability of the speaker's party from a single utterance. The underlying reasoning is as follows. If speakers from different parties use more distinct phrases in their speeches, the predictability of party membership from speeches increases and thus indicates a session with higher partisanship as well. A similar study, by Peterson and Spirling (2018), exploits the same idea and derives a polarization index for the British politics using the speeches in the House of Commons.

An application by Bertrand and Kamenica (2018) is perhaps the most similar one to the current study in terms of how prediction accuracy is used as a substantial quantity of interest. In their study, Bertrand and Kamenica (2018) predict an individual's group membership within income categories, education levels, gender, race, or political ideology; from either their media usage, consumption patterns, time usage, or social attitudes. The central statistical apparatus here is that the higher the predictability of group membership from –for example– media usage is, the more distinct the groups in their media usage are.

In this study, I define the *class distinctiveness* as the ease with which one can infer the class membership of a respondent solely from his or her economic preferences. This operationalization immediately lends itself to a classification problem with the dependent variable –*class membership*– being a binary variable for measuring pairwise class distinctiveness and a categorical variable with four levels for measuring overall class distinctiveness. The explanatory variables (predictors) are the economic preferences that are listed in Section 2.1.[7] To obtain pairwise and overall class distinctiveness measures quantitatively, I use logistic and multinomial logistic regressions, respectively, for pairwise class distinctiveness and overall class distinctiveness. The following specification is used in all predictions but run separately for each country in each wave:

$$Class_i = \alpha_1 \cdot Govt\_Resp_i + \alpha_2 \cdot Free\_or\_Eq_i + \alpha_3 \cdot Conf\_Union_i, \quad (1)$$

where $Class_i$ is the class membership of individual i. $Govt\_Resp_i$, $Free\_or\_Eq_i$ and $Conf\_Union_i$ correspond to the economic preferences of individual i, respectively, for government responsibility in economy, the choice of freedom or equality, and confidence in labor unions.

Let me illustrate how the proposed method works with a simple example. Suppose that we have three classes that are class A, B, and C, each with the same number of people in our hypothetical society at a given time. Suppose also, for illustration purposes, that there is only one economic preference, which can take three values that are P_1, P_2, and P_3. Let us first work out the extreme cases. Suppose that each class is homogeneous in its economic preference within itself and different than the other classes. For example, let's assume that everyone in class A favors P_1, everyone in class B favors P_2, and everyone in class C favors P_3. Suppose now that we need to predict the class membership of a person whose economic preference we already know. In this case, we would immediately tell that this person is from class A if she favors P_1, from class B if she favors P_2, and from class C if she favors

[7] One caveat of the proposed way of measuring class distinctiveness is that I use only three economic preferences as the predictors of class membership. Therefore, if I am leaving out an important predictor of class membership, then what I capture with the proposed measure may over- or under-estimate the true underlying class distinctiveness in economic preferences. This limitation in the number of predictors of class membership is due to the lack of data availability in the EVS. The three predictors that I do include in the prediction modeling, on the other hand, are relatively standard and encompassing variables (for the economic policy dimension of politics) that are also asked repeatedly in surveys and used in previous works to measure the positions of voters on the economic left-right dimension. This reassures that these predictors are likely to capture the preferences of voters on the economic left-right dimension.

P_3. Accordingly, our prediction accuracy rate would also be high, as is class distinctiveness. This is the perfect predictability and highest class distinctiveness case.

On the other extreme, suppose that each class has the same composition of economic preferences. For example, let's assume that 20% of class A favors P_1, 30% favors P_2, and 50% favors P_3. Suppose that this is also true for class B and C.[8] Let us try again to predict the class membership of someone whose economic preference we already know. Notice that we do not have much predictive ability in this case since this person may be in any class with probabilities equal to the proportion of the class in our hypothetical society. Therefore, our predictive performance would only be as good as assigning people randomly into classes. This is the case when class distinctiveness is lowest, and prediction accuracy is very close to the performance of a random assignment. For all the other values of the class distinctiveness between its highest and lowest cases, the prediction accuracy is between the perfect predictability and random assignment case.

To implement this idea, for each country in each wave, I first partition the data for a given country-year randomly into training and testing sets, such that the training set constitutes 70% and the testing set constitutes the remaining 30% of the country-year sample. I then estimate Eq.1 for every country-year in the data set with the training set data. The testing set data are reserved for out-of-sample predictions. The class membership of the respondents in the testing set data is predicted with the estimated models and compared to the true class membership of these respondents. This process leaves us with both predicted and true class membership of the respondents in the testing set data. The *prediction accuracy* is then defined as the percentage of correctly classified respondents. This number represents "the ease with which one can infer the class membership of a respondent solely from his or her economic preferences." In short, the obtained prediction accuracy rates serve as the class distinctiveness measure. The higher the predictability is, the higher is class distinctiveness.

I use multinomial logistic regression predictions to obtain the overall class distinctiveness measure –that is how distinct the four classes are in terms of their economic preferences– whereas I use binary logistic regressions to obtain the pair-wise class distinctiveness measure –that is how distinct a given class is from the *working class* in terms of economic preferences.

Note that the predictions could also be made using more sophisticated prediction algo-

[8]This example also works with any other class composition in terms of economic preferences. The only important criterion in this extreme case is that each class has the same percentage of people holding the same economic preference.

rithms such as a regression tree or random forest. Indeed Bertrand and Kamenica (2018) employ three different prediction algorithms and use the ensemble of these to increase the predictive performance. The more sophisticated prediction algorithms and their ensembles result in better predictions in general (Varian, 2014). The level of prediction accuracy, however, is not the quantity of interest *per se* in this study. What is of interest is the variation in prediction accuracy rates across time and space. Consequently, this requires prediction accuracy rates to be comparable across waves and countries.

Regarding the comparability of the class distinctiveness measure across time and space, note that the changes in the class compositions are not a threat for this measure, but they are actually what this measure is supposed to capture. For example, if a certain class becomes more populated by female respondents over time and female respondents are somehow characterized by relatively left-wing economic preferences compared to males, this class will emerge as more left over time. Importantly, however, such a change in the compositions of and (consequently) in the preferences of classes is a part of what I intend to capture. This is because the blurring of class divisions hypothesis encompasses such changes in the composition of classes and consequently in their preferences. Therefore, in the context of this study, the changing composition of classes does not pose a threat to comparability.

A concern that might pose a threat for comparability is the imbalanced nature of the data. The classes in the country-year samples do not constitute equal shares of the sample, and there exist severe imbalances in the number of observations belonging to each class. Moreover, the sample sizes of country-year data also exhibit a variation between countries and across time, posing yet another threat to the comparability of prediction accuracy rates. For example, a country-year sample with more observations than others is more likely to yield different prediction accuracy rates only because it is trained with more data. Such implications of imbalanced samples should, of course, be addressed to enable a truthful comparison.

To illustrate the adverse effects of the imbalance problem, suppose that we try to predict a binary class membership variable with levels A and B with some predictors. Assume that the data set consists of imbalanced classes: Class A has a very low proportion, say 10%, and the remaining of the sample is labeled as Class B. In this case, a classifier such as logistic regression might find it optimal to classify everything into class B since doing so implies a 90% prediction accuracy, and the classifier's job is to maximize the prediction accuracy. As a consequence, the model does a poor job in predicting the class membership, yet it still yields a high predictive accuracy rate. It classifies all observations belonging to class A incorrectly to class B. Therefore, it cannot reflect the true underlying class distinctiveness of the sample. Next, I explain in detail how I mitigate this imbalance problem.

Imbalance Problem. There exist several ways of dealing with the imbalance problem (Kuhn and Johnson, 2013). One of the simplest ways is to tune the model such that it optimizes the prediction accuracy of the minority class, called *sensitivity*. An alternative approach is simply to change cut-off values –for the class membership variable– so that the same prediction model results in different prediction accuracy rates and sensitivity. This method does not alter the prediction model *per se* but post-processes the already predicted values. Neither of them, however, ensures the comparability of prediction accuracy rates across time and space, which is of crucial importance for the purpose of this study. In both approaches, the variation in the sample sizes across countries and years remains a threat to the comparability.

A more convenient way of dealing with the imbalance problem for the purposes of this study might be using case weights. This method places more weight on the observations of minority class(es), for example, by duplicating some observations in the minority class. Although this method solves the problem of class imbalance, the variation in the sample sizes across countries and years remains a problem.

Finally, an approach that solves both the imbalances in class size and sample sizes of different country-year data sets is the resampling method. This method entails either up-sampling the minority class or down-sampling the majority class (or both at the same time). Using up-sampling and down-sampling together, it is possible to balance both class sizes and sample sizes of country-year data, both across time and space. Up-sampling does not increase the information in the data but only puts more weight on the minority class (Chen et al., 2004). On the other hand, since we repeat the resampling process a large number of times, the effect of down-sampling on the representativeness is mitigated too.

Consequently, to balance class sizes, I determine a class size k, which is identical for each class. This amounts to a total sample size of $4 * k$ for the multinomial logistic regressions and $2 * k$ for the logistic regressions with the binary dependent variable. The resampling process entails up-sampling of classes with sample sizes smaller than k and down-sampling of classes with sizes larger than k. I choose the class size $k = 100$ (the average class size in our sample is 300, 234, and 409, respectively, for the years 1990, 1999, and 2008).[9] Note that the sample sizes of country-years are also balanced by fixing the class sizes.

The sampling process of each class is repeated 500 times. Each round of draws from classes constitutes a country-year sample. A logistic or multinomial logistic regression has thus been estimated for each country-year sample for 500 times. The prediction accuracy rates obtained from these regressions are averaged over the draws to obtain a final prediction accuracy for each country-year. This procedure enables a meaningful comparison of prediction accuracy rates across countries and over time since the class sizes and sample sizes no longer idiosyncratically affect the prediction accuracy. A schema illustrating the resampling process for the regressions with four classes is provided in Figure E.1. The resampling process for the regressions with the binary class membership variable is analogous to the one illustrated in Figure E.1.

4 Results

In the following sections, I first provide an empirical validation of the class distinctiveness measure developed in the previous section. I then test the blurring class divisions hypothesis and its relationship with class voting.

4.1 Validation of the Class Distinctiveness Measure

Although similar measures have been developed for studying the evolution of cultural differences between distinct groups (Bertrand and Kamenica, 2018), the class distinctiveness measure used in this study is an innovative measure in the study of electoral politics. I, therefore, find it useful to provide some supporting evidence to demonstrate its validity.

To this end, although it is not the primary goal of this paper to explain cross-country variations of class distinctiveness, I first provide some supporting evidence from the structure of political competition across countries. This evidence aims to ensure that the cross-country pattern of class distinctiveness measure aligns with the established findings in the electoral politics literature. Second, I look at the pairwise class distinctiveness measures and show

[9] I, however, also experiment with $k = 50$ and 200 to show that the results are robust to class size choice. The correlations between the results of these experiments are given in Appendix A. The results with class sizes $k = 50$ and 200 are very similar to those with class size $k = 100$.

that the class distinctiveness measure reflects the *expected* differences between certain classes. Lastly, I argue that the evolution of class distinctiveness measure over time in Great Britain runs parallel to the class differences trends reported by Evans and Tilley (2017). These validation strategies are discussed in detail below.

Before discussing the validation, let me make clarify why the magnitudes of class distinctiveness measure, in other words, prediction accuracy rates, are relatively low. Figure 2 reports that the prediction accuracy rates range between 28% and 35%. The random assignment of the respondents into classes, on the other hand, would yield a prediction accuracy of %25 since the class sizes are equalized beforehand. Therefore, the predictive performance of the trained logistic regressions is slightly better than that of the random assignment case. The main reason for these low prediction accuracy rates, however, is that I include only a few variables –three variables that measure economic preferences– as explanatory variables/predictors in the predictions.

The inclusion of variables, such as income, education, age, etc., is possible and would undoubtedly increase the model's predictive performance as they are expected to differ substantially between distinct classes. Nevertheless, we then would not be able to interpret the prediction accuracy as a measure of class distinctiveness in economic preferences. We want it to reflect only the distinctiveness of classes in terms of economic preferences. We, therefore, restrict the explanatory variables to the ones representing economic preferences only. Finally, let me state once again that we are not interested in the levels of prediction accuracy *per se* but in how it changes across countries and over time.

Political cleavage structures. According to Rokkan, the two revolutions have set the cleavage structure on which political competition takes place. The industrial revolution has led to the emergence of an economic dimension, while the French Revolution has led to nation-state building and the state-church conflict, and hence the emergence of a non-economic cultural dimension. These are the main reasons why today's political competition takes place mostly on these two dimensions –economic and cultural. Manow et al. (2018) elaborate these arguments in detail and link the historically rooted cleavage structures of political competition to different types of welfare regimes by providing supporting empirical evidence.

The four types of welfare regimes described in Manow et al. (2018) are the northern, liberal Anglo-Saxon, continental and southern types. The southern type is added to the well-known Esping-Andersen welfare regime type classification (Esping-Andersen, 2013) by Manow et al. (2018). Among the European countries; Denmark, Finland, and Sweden belong to the northern type; Ireland and the United Kingdom to the liberal Anglo-Saxon type; Austria, Belgium, Germany, France, and the Netherlands to the continental type; finally Greece, Italy, Portugal, and Spain belong to the southern type. Here I provide neither a comprehensive review of the differences between these welfare regimes nor of the underlying reasons (for this, see Polk and Rovny (2016)). I, however, do provide some key features of these welfare regimes that can be linked to the class distinctiveness in economic preferences.

First of all, the northern type of welfare regime has not experienced a church-state conflict. The political competition, therefore, has been taking place mostly on the economic dimension in countries belonging to this group, whereas the cultural dimension is less salient. Second, in the continental type of welfare regime, the state-church conflict is more pronounced compared to the northern type regime, leading both economic and cultural dimensions to be salient in the political competition. Manow et al. (2018), however, notably states that although belonging to the continental type, "France resembles the southern type in many respects."

Third, the southern type of welfare regime is characterized by its relatively more salient cultural dimension than the economic dimension due to the historical state-church conflict. Fourth, the Anglo-Saxon model also has an economic dimension more salient than the cultural dimension.

To sum up, in the light of the evidence provided by Polk and Rovny (2016) and Manow et al. (2018), we expect to see that political competition is oriented towards more economic issues in the countries belonging to the northern type. In contrast, we expect to see that it is oriented towards more cultural issues in the countries belonging to the southern type, and possibly France. These expectations also imply that we should expect more class distinctiveness in northern types -since class distinctiveness is defined in terms of economic preferences- and less class distinctiveness in southern types.

Figure 2 shows the averages (over three points in time) of the class distinctiveness measure for the countries whose welfare regime type we know thanks to Manow et al. (2018). Although the class distinctiveness measure shows only limited variation across countries, it aligns with the expected pattern described above. Northern type countries such as Denmark and Sweden exhibit the highest class distinctiveness, whereas southern type countries such as Portugal, Spain, and Italy exhibit the lowest class distinctiveness. Moreover, although classified as a continental type welfare regime, France is very close to the southern type countries, as noted by Polk and Rovny (2016).

The continental and liberal Anglo-Saxon type countries take positions in between, except Ireland, which is the only country that does not fit the hypothesized pattern. One reason for the observed pattern for Ireland may be the more pronounced state-church conflict since a large part of Ireland's population is Roman Catholic. Considering that the state-church conflict (and consequently the emergence of a salient cultural dimension) was more pronounced in countries where the Catholic Church (instead of the Protestant churches) was more dominant, the case of Ireland may not be shocking (Manow, 2008). All in all, I believe that Figure 2 grants some credibility to the class distinctiveness measure developed in the previous section.

Pairwise class differences. An alternative way to validate the class distinctiveness measure considers binary class predictions reported in Figure 3. The plots in Figure 3 show the prediction accuracy rates between two classes for each country in 1990. The plots for 1999 and 2008 are given in Appendix F.

Informed by the way the EGP classification is constructed, we know that the working class differs from other classes, especially in the way that earnings are obtained, level of earnings, job security, promotion prospects, and work conditions. We expect that these differences in the characteristics of occupations also reveal themselves as differences in the economic preferences of people who hold those occupations. For example, the working class people work in less secure jobs and worse working conditions with lower earnings and low promotion prospects.

The service class, on the other hand, has better terms in all these respects. The self-employed people, on the other hand, have been known traditionally to hold more right-wing economic preferences compared to the working class people (Arzheimer and Carter, 2006; Lubbers et al., 2002; Lubbers and Scheepers, 2001). Finally, the routine non-manual class is expected to be more similar to the working class in terms of economic preferences than the other two classes are. The reason is that the routine non-manual class constitutes the lower tail of the middle class, and it is the closest to the working class in terms of the occupation characteristics. The routine non-manual class is also called "the routine white-collar *workers*" (Evans and Tilley, 2017).

Figure 3 confirms these expectations. Part (a) and (c) indeed report higher prediction accuracy rates on average, in other words higher class distinctiveness, compared to Part (b),

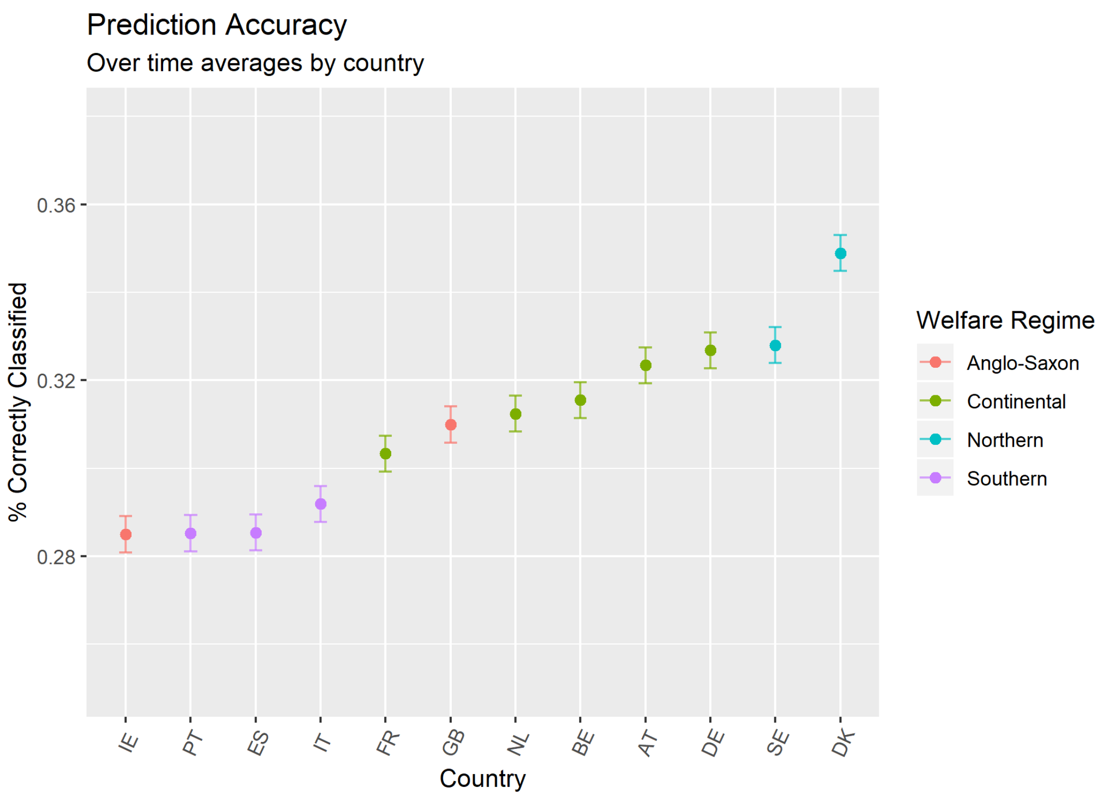

Figure 2: Predictions with multinomial logistic regressions.

Note: In the multinomial logistic regressions, the dependent variable is the categorical class membership variable with four levels. The class membership of the respondents is predicted from their economic preferences only. The reported numbers are the country-level over time averages of the percentages of correctly classified respondents in terms of class membership in the predictions. The bars correspond to the standard errors of the over time prediction accuracy averages of countries. They are computed as the standard deviation of over time prediction accuracy averages of countries over the square root of the number of countries. Countries are labeled as follows: AT = Austria, BE = Belgium, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, IE = Ireland, IT = Italy, NL = Netherlands, PT = Portugal, SE = Sweden.

which shows the relative class distinctiveness between the working class and the routine non-manual class. We observe the same pattern also in 1999 and 2008 (reported in Appendix F). The means of the prediction accuracy rates reported in Part (a), (b), and (c) in Figure 3 are; respectively, 0.59, 0.52 and 0.61. These figures indicate the relative similarity of the routine non-manual class to the working class in terms of economic preferences, as expected.[10]

Great Britain. Although there exists no cross-country evidence for the evolution of class distinctiveness over time, Evans and Tilley (2017) provide some figures for Great Britain.

[10]Similarly, the means of the prediction accuracy rates are 0.57, 0.53 and 0.58 for 1999; and 0.55, 0.51, 0.58 for 2008.

Figure 3: Predictions with logistic regression.

Note: In the logistic regressions, the dependent variable is the binary class membership variable. Its levels are (a) working class and service class, (b) working class and the routine non-manual class, (c) working class and self-employed. The class membership of the respondents is predicted from their economic preferences only. The reported numbers are the percentages of correctly classified observations for the year 1990 in the predictions. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

They use both the British Election Study (the BES) and the British Social Attitudes (the

BSA) data sets and study the period between 1963-2015. According to the figures reported in their study, the differences between classes in terms of economic preferences first decrease between 1990 and 1999, and then slightly increases between 1999 and 2008. This pattern is repeated in Figure 4, where the prediction accuracy rates for Great Britain are reported over time, although the increase from 1999 to 2008 appears to be a minimal one.

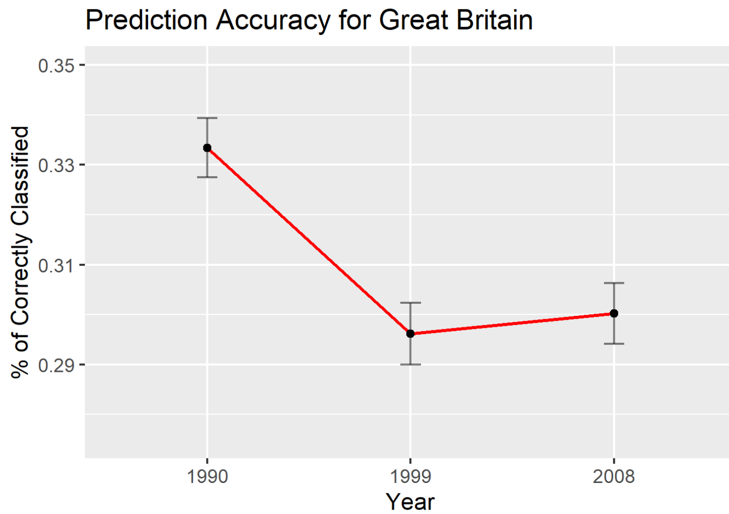

Figure 4: Predictions with multinomial logistic regression for Great Britain.

Note: In the multinomial logistic regression, the dependent variable is the categorical class membership variable with four levels. The class membership of the respondents is predicted from their economic preferences only. The reported numbers are the percentages of correctly classified observations for Great Britain in each wave of the EVS. The bars correspond to the standard errors of the prediction accuracy rates. They are computed as the standard deviation of prediction accuracy rates within a wave over the square root of the number of countries.

4.2 Blurring of Class Divisions?

The blurring of class divisions hypothesis asserts that classes have lost their distinctiveness due to several reasons, including the transition from an industrial society to a post-industrial one, rising quality of life and welfare, more social mobility, etc. Figure 5 shows the evolution of prediction accuracy rates over time for each country in the sample as computed in the previous section. An increase in the prediction accuracy is interpreted as an increase in the class distinctiveness in economic preferences.

Figure 5 reports evidence of decreasing class distinctiveness for countries such as Bulgaria, Hungary, Poland, Portugal, and Slovakia. For the other countries in the sample, however, there is a variation in class distinctiveness rather than a general trend of decline in contrast to the claims of previous studies.

To summarize these results formally, I run a regression of prediction accuracy rates on a time trend with country fixed effects. Table 2 reports the results of this regression, where the time trend turns out to be negative and statistically significant. This statistically negative coefficient of time trend indicates that from one wave of the EVS to the next, the prediction accuracy –or class distinctiveness– decreases by 1.2pp.

In sum, although there is evidence of decreasing class distinctiveness on average and specifically for some countries, this can not be generalized to all countries as done by the previous studies. The findings in Figure 5 point out a variation rather than a general declining trend for class distinctiveness.

Table 2: The time trend of class distinctiveness

	Dependent variable:
	Class Distinctiveness
Time Trend	-0.012^{***} (0.004)
Country FE	✓
Constant	0.347^{***} (0.015)
Observations	54
R^2	0.522

Note: The reported results are from OLS estimations. The dependent variable is the overall class distinctiveness. *p<0.1; **p<0.05; ***p<0.01.

4.3 Class Voting and Class Distinctiveness

In this section, I first estimate the class voting for every country in the sample for three points in time. I then test the statistical association between the class distinctiveness in economic preferences and class voting. To estimate the strength of class voting, I use multi-level/hierarchical modeling. These models estimate a separate regression for each country-year in the sample by assuming a probability distribution for the coefficients of these re-

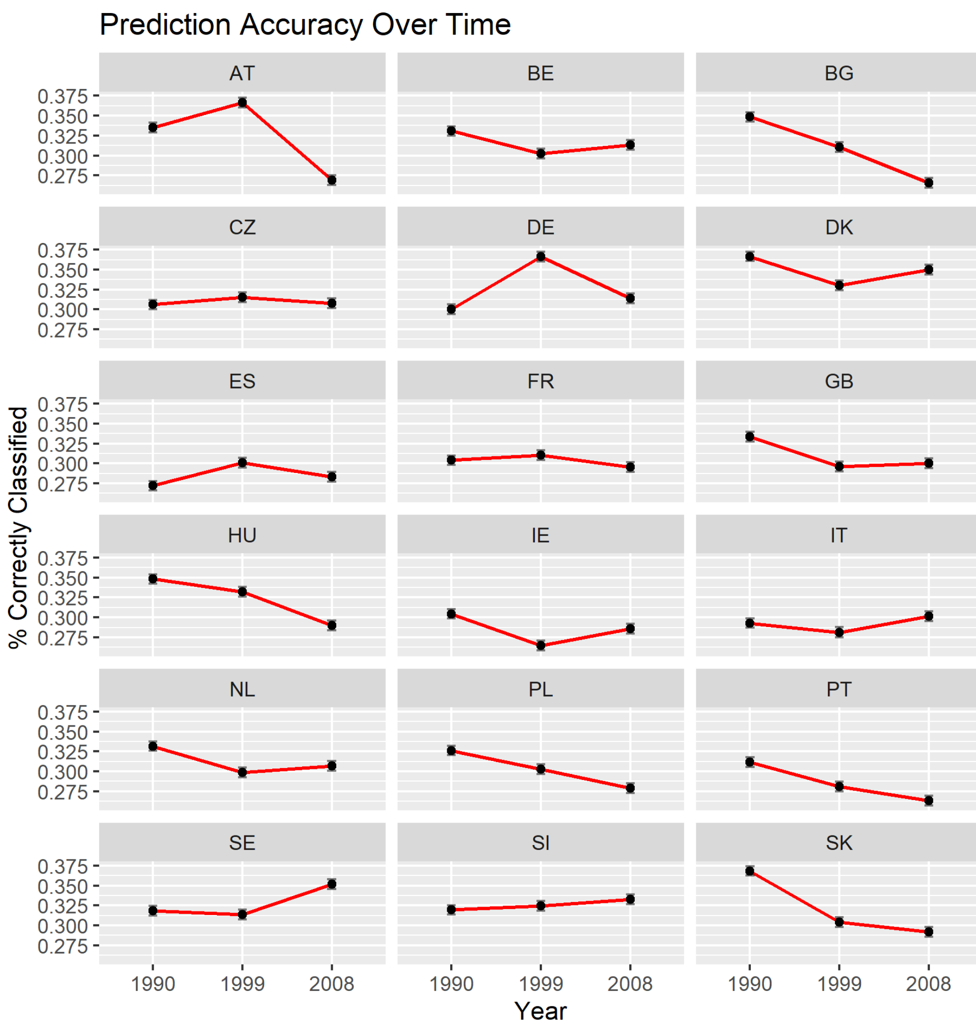

Figure 5: Predictions with multinomial logistic regression.

Note: In multinomial logistic regressions, the dependent variable is the categorical class membership variable with four levels. The class membership of the respondents is predicted from their economic preferences only. The reported numbers are the percentages of correctly classified respondents in terms of class membership in each wave of the EVS. The bars correspond to the standard errors of the prediction accuracy rates. They are computed as the standard deviation of prediction accuracy rates within a wave over the square root of the number of countries. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

gressions. A multilevel model can be considered as a generalization of the linear regression model with intercepts and/or slopes being allowed to vary by group –in the current setting by country-year (Gelman and Hill, 2006).

The reason why I opt for multilevel modeling is the compromise that these models make between *complete pooling* and *no pooling* cases. Complete pooling corresponds to estimating a single intercept and slope for all country-years. No pooling, on the other hand, fits a separate regression (thus separate intercept and slope terms) for each country-year. While complete pooling ignores the variation among country-years, no pooling makes country-year level estimates seem more different than they actually are due to overstating the variation within country-years. A multilevel model estimate for a country-year is a weighted average of the no pooling and complete pooling estimates. These models give more weight to groups with larger sample sizes since those groups are likely to carry more information. The estimates for countries with smaller numbers of observations, therefore, are pulled towards the complete pooling estimates, whereas the estimates for countries with larger numbers of observations are drawn towards the no pooling estimate of the country.

In order to obtain the strength of class-vote linkage, the following specification is estimated by a multilevel model:

$$
\begin{aligned}
LR_{ijt} \quad = \quad & \beta_{1jt} \cdot Class_{ijt} + \beta_2 \cdot Age_{ijt} + \beta_3 \cdot Gender_{ijt} + \beta_4 \cdot Educ_{ijt} + \\
& Country\_FE + Time\_FE + u_{ijt},
\end{aligned}
$$

where LR_{ijt} is the left-right economic position of the preferred political party by respondent i, in country j, at time t. $Class_{ijt}$ corresponds to the class membership of respondent i. $Country\_FE$ and $Time\_FE$ correspond to the country and time fixed effects, respectively. The coefficients of $Class_{ijt}$ variable (β_{1jt}'s that vary at the country-year level) provide us with the strength of class voting.

The multilevel model estimates the coefficient of the class membership variable for every country-year in the sample. The class membership variable is a categorical one with four levels, and I choose the reference category as the working class. The coefficient of the class membership variable, therefore, estimates the conditional differences in preferred party positions between i) the working class and the service class, ii) the working class and the routine non-manual class, and iii) the working class and the self-employed. Consequently, this estimation yields three sets of coefficients (β_{1jt}'s) for the class membership variable for each country-year.

Figure 6 reports the differences between the economic left-right positions of the preferred parties by the working class and service class people.[11] Note that, before the estimation, the party positions are standardized such that they have a mean zero and standard deviation of one within each country-year. In Figure 6, if we take Austria (the first country plot) for example, the coefficient of the class membership variable is very close to 0.4 in 1990. This number corresponds to the difference between the economic left-right positions of the preferred political parties by the service class and working class. It indicates that service class people prefer parties with 0.4 standard deviations more on the economic right compared to the working class people.

In the last part of the analysis, I test whether class voting is associated with class distinctiveness in a panel regression. The coefficients of the class membership variable that we obtained in the previous multilevel model estimation are the dependent variable of the following specification. The main variable of interest is the pairwise class distinctiveness measure developed in Section 3. The panel regressions include class-pair, country, and time fixed effects. The reason why I also include a class-pair fixed effect is because the three sets of pairwise class coefficients –which are reported in Figure 6, G.1, and G.2– are used as the dependent variable in this panel regression. Finally, I estimate the following panel regression:

$$\begin{aligned} Class\_Voting_{kjt} &= \gamma \cdot Class\_Dist_{kjt} + Class\_pair\_FE + Country\_FE + \\ &\quad Time\_FE + \nu_{kjt}, \end{aligned}$$

where k represents the class-pair: working class vs. service class, working class vs. routine non-manual class, working class vs. self-employed. $Class\_Voting_{kjt}$ is the coefficient of the class membership variable obtained in the previous multilevel model, for class pair k, in country j, at time t. $Class\_Dist_{kjt}$ corresponds to the class distinctiveness measure for the class pair k, in country j, at time t. $Class\_pair\_FE$, $Country\_FE$ and $Time\_FE$ correspond, respectively; to class pair, country, and time fixed effects.

I estimate this specification with a panel regression. Table 3 summarizes the results of this regression. The first model in Table 3 is a panel regression of class voting on class distinctiveness with only class-pair fixed effects. It yields a positive and statistically significant coefficient for class distinctiveness. This result supports the second hypothesis that

[11] The results for the other two class pairs are given in Figure G.1 and G.2.

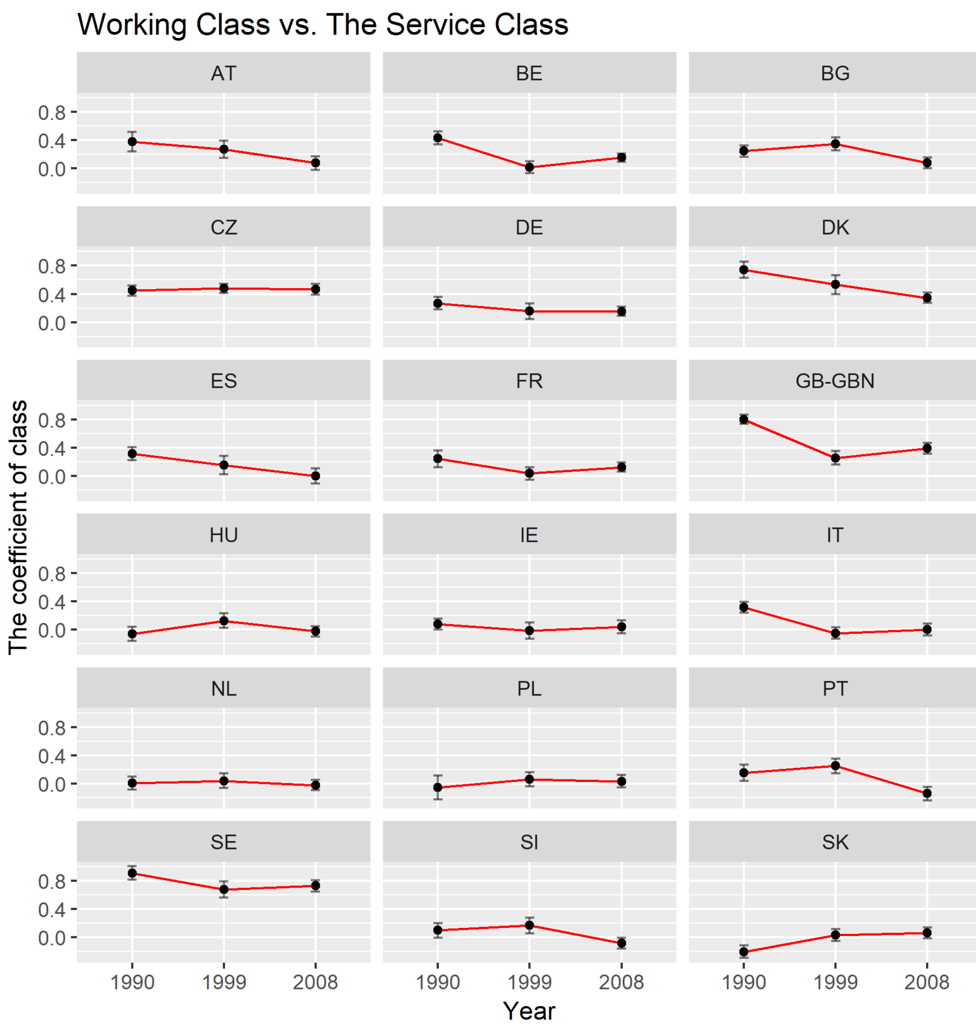

Figure 6: Multilevel estimation.

Note: The dependent variable is the left-right economic position of the preferred political party by the respondent. The coefficients of the class membership variable are reported for each country and time point in the sample. The reported coefficients correspond to the differences between the left-right economic positions of the preferred political parties by the working class and the service class. The bars around the point estimates correspond to the standard errors of the coefficient estimates. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

the higher the class distinctiveness is, the stronger the class voting is. The second model adds the country fixed effects to the first model. The third model additionally includes the time fixed effects. The coefficient of class distinctiveness remains positive and statistically significant in both specifications.

Furthermore, in addition to being statistically significant, the class distinctiveness has a substantially significant effect on the strength of class voting. The coefficient of 1.417 on the class distinctiveness variable implies that a 10% increase in the class distinctiveness variable is associated with a 0.14 increase in the strength of class voting. This 0.14 increase in class voting, in turn, corresponds to 0.14 difference between the economic left-right positions of political parties preferred by the working class and any of the other classes (considering that the party positions are standardized beforehand such that they have a mean zero and standard deviation of one).

For robustness, I run the same panel regressions as in Table 3 with observation weights. These weights correspond to the inverse of the standard errors of the coefficient on the class membership variable in the multilevel estimation. Doing so ensures that we also take into account the uncertainty of the estimated class coefficients from the previous multilevel model. The results, which are reported in Table H.1, are very similar to the ones in Table 3 in terms of the magnitude, sign, and statistical significance. Overall, these findings provide strong evidence for our second hypothesis of a positive relationship between class voting and class distinctiveness.

Table 3: Panel regressions of class voting on class distinctiveness

	Dependent variable:		
	Strength of class voting		
	(1)	(2)	(3)
Class distinctiveness	1.924***	1.671***	1.417***
	(0.462)	(0.357)	(0.366)
Constant	−0.890***	−0.770***	−0.579**
	(0.264)	(0.213)	(0.222)
Class-pair FE	✓	✓	✓
Country FE		✓	✓
Time FE			✓
Observations	162	162	162
R^2	0.225	0.696	0.711

Note: The reported results are from panel regressions. The dependent variable is the strength of class voting, which is obtained from the multilevel model estimation. *p<0.1; **p<0.05; ***p<0.01.

5 Conclusion

The recent literature on class voting has centered around the mechanisms that drive class voting. This paper concerns one of the potential mechanisms: the blurring of class divisions. The previous scholarship in political science, and also in sociology, has either announced the death of class as a useful concept in electoral politics or has claimed that its explanatory power in electoral politics halted. Although this hypothesis has been pronounced in several studies, there has been no empirical evidence neither on the blurring of class divisions nor on its relationship with class voting so far.

In this paper, I address this specific question using the European Values Survey and Manifesto Project data sets and by operationalizing the class divisions as the differences in economic preferences between distinct classes. I transform differences in economic preferences between classes into a single measure called *class distinctiveness* by predictive modeling. This transformation is based on the simple idea that the predictability of class membership from economic preferences is itself an indicator of class distinctiveness.

In order to validate the newly-developed class distinctiveness measure, I provide supporting evidence; first from the structure of political competition within countries, second from expectations about pairwise class distinctiveness in consideration of how the EGP classification is constructed, and third from the trend of class differences in Great Britain. Based on the newly developed class distinctiveness measure, I subsequently present some evidence of declining class divisions over time. This subsiding trend is nonetheless pronounced for some countries, whereas others exhibit a variation rather than a generally declining trend. This result cast serious doubt on the commonly postulated hypothesis of blurring class divisions over time and indicates that this hypothesis may only be a part of the story.

I then test the relationship between class distinctiveness and class voting. To obtain a quantitative measure of the latter, I run regressions of the left-right economi positions of preferred political parties by survey respondents on their class memberships. A multilevel model is estimated for this purpose since it deals with both within and between country variation. Having obtained both the class distinctiveness and class voting measures, I then proceed to estimate panel regressions to reveal the nature of the statistical relationship between class-voting and class distinctiveness. These regressions provide strong evidence in favor of a positive relationship between class voting and class distinctiveness. Overall, these findings point out that the class concept is still relevant and of interest to the studying electoral politics.

A major caveat of this study relates to the number of explanatory variables that are used for representing economic preferences. We know that the larger the number of variables available on the same issue, the more representative they are (Ansolabehere et al., 2008). Due to missing variables in some countries, however, the number of variables related to economic preferences are limited to three in this study. Another limitation is the fixed nature of classes due to the EGP class schema. This limitation implies that I do not allow the occupations to switch class membership over time even if their working conditions, related earnings, etc., change over time. I, however, believe that this is not a severe issue for the majority of occupations considered in the EGP. As a final note, I believe that replicating this type of an analysis for other political dimensions, such as the cultural dimension, would be an exciting and promising direction for future research.

References

Ansolabehere, S., Rodden, J., and Snyder Jr, J. M. (2008). The strength of issues: Using multiple measures to gauge preference stability, ideological constraint, and issue voting. *American Political Science Review*, pages 215–232.

Arzheimer, K. and Carter, E. (2006). Political opportunity structures and right-wing extremist party success. *European Journal of Political Research*, 45(3):419–443.

Arzheimer, K., Evans, J., and Lewis-Beck, M. S. (2016). *The Sage handbook of electoral behaviour*. Sage.

Bakker, R. and Hobolt, S. (2013). Measuring party positions. *Political Choice Matters: Explaining the Strenght of Class and Religious Cleavages in Cross-National Perspective*, pages 27–45.

Bertrand, M. and Kamenica, E. (2018). Coming apart? cultural distances in the united states over time. Technical report, National Bureau of Economic Research.

Boxell, L., Gentzkow, M., and Shapiro, J. M. (2020). Cross-country trends in affective polarization. Technical report, National Bureau of Economic Research.

Chen, C., Liaw, A., and Breiman, L. (2004). Using random forest to learn imbalanced data. *University of California, Berkeley*, 110:1–12.

Clark, T. N. and Lipset, S. M. (1991). Are social classes dying? *International sociology*, 6(4):397–410.

Connelly, R., Gayle, V., and Lambert, P. S. (2016). A review of occupation-based social classifications for social survey research. *Methodological Innovations*, 9:2059799116638003.

Dalton, R. J. (2013). *Citizen politics: Public opinion and political parties in advanced industrial democracies*. Cq Press.

Erikson, R. and Goldthorpe, J. H. (1992). *The constant flux: A study of class mobility in industrial societies*. Oxford University Press, USA.

Esping-Andersen, G. (2013). *The three worlds of welfare capitalism*. John Wiley & Sons.

Evans, G. (1992). Testing the validity of the goldthorpe class schema. *European Sociological Review*, 8(3):211–232.

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

Evans, G. (2000). The continued significance of class voting. *Annual review of political science*, 3(1):401–417.

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

Evans, G. and Neundorf, A. (2018). Core political values and the long-term shaping of partisanship. *British Journal of Political Science*, pages 1–19.

Evans, G. and Tilley, J. (2012a). The depoliticization of inequality and redistribution: Explaining the decline of class voting. *The Journal of Politics*, 74(4):963–976.

Evans, G. and Tilley, J. (2012b). How parties shape class politics: Explaining the decline of the class basis of party support. *British journal of political science*, 42(1):137–161.

Evans, G. and Tilley, J. (2017). *The new politics of class: The political exclusion of the British working class*. Oxford University Press.

Ganzeboom, H. B. and Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social science research*, 25(3):201–239.

Gelman, A. and Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

Gentzkow, M., Shapiro, J. M., and Taddy, M. (2018). Measuring group differences in high-dimensional choices: Method and application to congressional speech. Technical report, Technical report Stanford University. URL: http://web. stanford. edu ....

Guiso, L., Herrera, H., and Morelli, M. (2017). Demand and supply of populism.

Inglehart, R. and Rabier, J.-R. (1986). Political realignment in advanced industrial society: from class-based politics to quality-of-life politics. *Government and Opposition*, 21(4):456–479.

Jansen, G., Evans, G., and De Graaf, N. D. (2013). Class voting and left–right party positions: A comparative study of 15 western democracies, 1960–2005. *Social science research*, 42(2):376–400.

Kitschelt, H. and Rehm, P. (2014). Occupations as a site of political preference formation. *Comparative Political Studies*, 47(12):1670–1706.

Kuhn, M. and Johnson, K. (2013). *Applied predictive modeling*, volume 26. Springer.

Lubbers, M., Gijsberts, M., and Scheepers, P. (2002). Extreme right-wing voting in western europe. *European Journal of Political Research*, 41(3):345–378.

Lubbers, M. and Scheepers, P. (2001). Explaining the trend in extreme right-wing voting: Germany 1989–1998. *European Sociological Review*, 17(4):431–449.

Manow, P. (2008). Electoral rules, class coalitions and welfare state regimes, or how to explain esping-andersen with stein rokkan. *Socio-Economic Review*, 7(1):101–121.

Manow, P., Palier, B., and Schwander, H. (2018). *Welfare Democracies and Party Politics: Explaining Electoral Dynamics in Times of Changing Welfare Capitalism*. Oxford University Press.

Oesch, D. and Rennwald, L. (2018). Electoral competition in europe's new tripolar political space: Class voting for the left, centre-right and radical right. *European Journal of Political Research*.

Peterson, A. and Spirling, A. (2018). Classification accuracy as a substantive quantity of interest: Measuring polarization in westminster systems. *Political Analysis*, 26(1):120–128.

Polk, J. and Rovny, J. (2016). Worlds of welfare and multidimensional party competition in europe. *Unpublished manuscript*.

Rose, D. and Pevalin, D. J. (2002). *A researcher's guide to the national statistics socio-economic classification*. Sage.

Spies, D. (2013). Explaining working-class support for extreme right parties: A party competition approach. *Acta Politica*, 48(3):296–325.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2):3–28.

Yagci, A. H., Harma, M., and Tekgüç, H. (2020). Who wants left-wing policies? economic preferences and political cleavages in turkey. *Mediterranean Politics*, pages 1–26.

APPENDIX

A Country-year Samples

Table A.1: The sample sizes of country-year data sets and classes

Country	Wave	# of Obs.	Service	Routine	Self-emp.	Working
AT	1990-1993	1061	22	534	140	365
AT	1999-2001	1219	53	588	138	440
AT	2008-2010	1191	368	366	111	346
BE	1990-1993	1856	96	612	274	874
BE	1999-2001	1415	198	705	83	429
BE	2008-2010	1296	558	212	81	445
BG	1990-1993	830	156	152	22	500
BG	1999-2001	596	180	56	37	323
BG	2008-2010	1151	360	207	55	529
CZ	1990-1993	1909	197	617	66	1029
CZ	1999-2001	1397	376	298	65	658
CZ	2008-2010	1347	372	322	64	589
DK	1990-1993	881	51	389	73	368
DK	1999-2001	748	25	362	61	300
DK	2008-2010	1329	625	272	79	353
FR	1990-1993	690	81	365	81	163
FR	1999-2001	1156	135	584	113	324
FR	2008-2010	1391	569	293	89	440
DE	1990-1993	2864	59	1291	172	1342
DE	1999-2001	1287	43	656	127	461
DE	2008-2010	1672	484	406	71	711
HU	1990-1993	783	107	110	36	530
HU	1999-2001	803	91	159	60	493
HU	2008-2010	1268	403	192	61	612
IE	1990-1993	880	147	198	100	435
IE	1999-2001	537	45	203	106	183
IE	2008-2010	691	204	226	50	211
IT	1990-1993	1173	227	446	86	414
IT	1999-2001	1195	298	315	142	440
IT	2008-2010	1067	360	211	172	324
NL	1990-1993	814	187	265	47	315
NL	1999-2001	620	148	185	36	251
NL	2008-2010	1364	691	295	100	278
PL	1990-1993	651	22	153	137	339
PL	1999-2001	778	103	161	102	412
PL	2008-2010	1045	273	262	95	415

continued

Country	Wave	# of Obs.	Service	Routine	Self-emp.	Working
PT	1990-1993	819	38	231	87	463
PT	1999-2001	625	83	97	76	369
PT	2008-2010	1071	195	258	96	522
SK	1990-1993	932	106	247	13	566
SK	1999-2001	836	123	222	29	462
SK	2008-2010	1062	282	226	50	504
SI	1990-1993	726	97	175	44	410
SI	1999-2001	708	109	188	41	370
SI	2008-2010	1074	423	166	58	427
ES	1990-1993	1457	131	275	217	834
ES	1999-2001	564	70	75	75	344
ES	2008-2010	1086	215	274	105	492
SE	1990-1993	893	149	446	40	258
SE	1999-2001	629	61	277	75	216
SE	2008-2010	887	428	216	49	194
GB	1990-1993	1309	216	305	93	695
GB	1999-2001	578	118	135	31	294
GB	2008-2010	1271	590	265	90	326

Note: The reported numbers are the sample sizes of country-year data sets and classes across the EVS waves. "Service" corresponds to the service class, "Routine" to the routine non-manual class, "Self-emp." to the self-employed class, and "Working" to the working class. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

Table A.2: The share of classes within countries - 1990

Country	Service	Routine	Self-emp.	Working
AT	0.02	0.50	0.13	0.34
BE	0.05	0.33	0.15	0.47
BG	0.19	0.18	0.03	0.60
CZ	0.10	0.32	0.03	0.54
DK	0.06	0.44	0.08	0.42
FR	0.12	0.53	0.12	0.24
DE	0.02	0.45	0.06	0.47
HU	0.14	0.14	0.05	0.68
IE	0.17	0.22	0.11	0.49
IT	0.19	0.38	0.07	0.35
NL	0.23	0.33	0.06	0.39
PL	0.03	0.24	0.21	0.52
PT	0.05	0.28	0.11	0.57
SK	0.11	0.27	0.01	0.61
SI	0.13	0.24	0.06	0.56
ES	0.09	0.19	0.15	0.57
SE	0.17	0.50	0.04	0.29
GB	0.17	0.23	0.07	0.53

Note: The reported numbers are the shares of classes within countries. "Service" correspond to the service class, "Routine" to the the routine non-manual class, "Self-emp." to the self-employed class, and "Working" to the working class. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

Table A.3: The share of classes within countries - 1999

Country	Service	Routine	Self-emp.	Working
AT	0.04	0.48	0.11	0.36
BE	0.14	0.50	0.06	0.30
BG	0.30	0.09	0.06	0.54
CZ	0.27	0.21	0.05	0.47
DK	0.03	0.48	0.08	0.40
FR	0.12	0.51	0.10	0.28
DE	0.03	0.51	0.10	0.36
HU	0.11	0.20	0.07	0.61
IE	0.08	0.38	0.20	0.34
IT	0.25	0.26	0.12	0.37
NL	0.24	0.30	0.06	0.40
PL	0.13	0.21	0.13	0.53
PT	0.13	0.16	0.12	0.59
SK	0.15	0.27	0.03	0.55
SI	0.15	0.27	0.06	0.52
ES	0.12	0.13	0.13	0.61
SE	0.10	0.44	0.12	0.34
GB	0.20	0.23	0.05	0.51

Note: The reported numbers are the shares of classes within countries. "Service" correspond to the service class, "Routine" to the the routine non-manual class, "Self-emp." to the self-employed class, and "Working" to the working class. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

Table A.4: The share of classes within countries - 2008

Country	Service	Routine	Self-emp.	Working
AT	0.31	0.31	0.09	0.29
BE	0.43	0.16	0.06	0.34
BG	0.31	0.18	0.05	0.46
CZ	0.28	0.24	0.05	0.44
DK	0.47	0.20	0.06	0.27
FR	0.41	0.21	0.06	0.32
DE	0.29	0.24	0.04	0.43
HU	0.32	0.15	0.05	0.48
IE	0.30	0.33	0.07	0.31
IT	0.34	0.20	0.16	0.30
NL	0.51	0.22	0.07	0.20
PL	0.26	0.25	0.09	0.40
PT	0.18	0.24	0.09	0.49
SK	0.27	0.21	0.05	0.47
SI	0.39	0.15	0.05	0.40
ES	0.20	0.25	0.10	0.45
SE	0.48	0.24	0.06	0.22
GB	0.46	0.21	0.07	0.26

Note: The reported numbers are the shares of classes within countries. "Service" correspond to the service class, "Routine" to the the routine non-manual class, "Self-emp." to the self-employed class, and "Working" to the working class. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

B Question Wordings of the Variables from the EVS

- **e037 - Government Responsibility:** On this card you see a number of opposite views on various issues. How would you place your views on this scale?

  1: Individuals should take more responsibility for providing for themselves,

  10: The state should take more responsibility to ensure that everybody is provided for.

- **e032 - Freedom or Equality:** Which of these two statements comes closest to your own opinion?

  A) I find that both freedom and equality are important. But if I were to make up my mind for/to choose one or the other, I would consider personal freedom more important, that is, everyone can live in freedom and develop without hindrance.

  B) Certainly both freedom and equality are important. But if I were to make up my mind for/to choose one or the other, I would consider equality more important, that is that nobody is underprivileged and that the social class differences are not so strong.

  1: Agreement with Statement A, 2: Agreement with Statement B, 3: Neither.

- **e69_05 - Confidence Labour Unions:** Please look at this card and tell me, for each item listed, how much confidence you have in them, is it a great deal, quite a lot, not very much or non at all?

  1: A great deal, 2: Quite a lot, 3: Not very much, 4: None at all.

- **e179 - Which political party would you vote for?: First Choice:** Response scales includes the political parties and change from country to country.

C EGP and Its Versions

The original eleven-class version of EGP class schema is coded into four-version according to Connelly et al. (2016) and Jansen et al. (2013).

Classes in the original EGP:

- **I: Higher Controllers:** higher grade professionals, administrators, officials; managers of large industrial establishments.
- **II: Lower Controllers:** lower grade professionals; higher grade technicians; managers in small industrial establishments; supervisors of non-manual employees.
- **IIIa: Routine Non-manual:** higher grade employees (administration and commerce).
- **IIIb: Routine Lower Sales-Service:** lower grade employees (sales and services).
- **IVa+IVb: Self-employed:** small proprietors, artisans with and without employees.
- **IVc: Self-employed Farmer**
- **V: Manual Work Supervisors:** foremen, supervisors of manual workers.
- **VI: Skilled Worker**
- **VIIa: Unskilled Worker**
- **VIIb: Farm Worker**

The class schema that is used in this study is:

- **1. The service class:** I + II.
- **2. The routine non-manual class:** IIIa + IIIb.
- **3. The self-employed:** IVa + IVb + IVc.
- **4. The manual working class:** V + VI + VIIa + VIIb.

D Robustness to Class Size Choice k

Table D.1: The correlations of the prediction accuracy rates between different class size choices

	Class pair:		
	Service vs. Work.	Routine vs. Work.	Self-emp vs. Work.
$k = 50$ vs. 100	0.98	0.95	0.94
$k = 50$ vs. 200	0.87	0.91	0.88
$k = 100$ vs. 200	1.00	0.96	0.96

Note: The reported numbers are the correlations between the prediction accuracy rates under different class size choices (k) for different class pairs. "Service" stands for the service class, whereas "Routine" for the non-manual routine class, "Self-emp" for the self-employed class, and "Work." for the working class.

E Resampling: Austria 1990 data set as an example

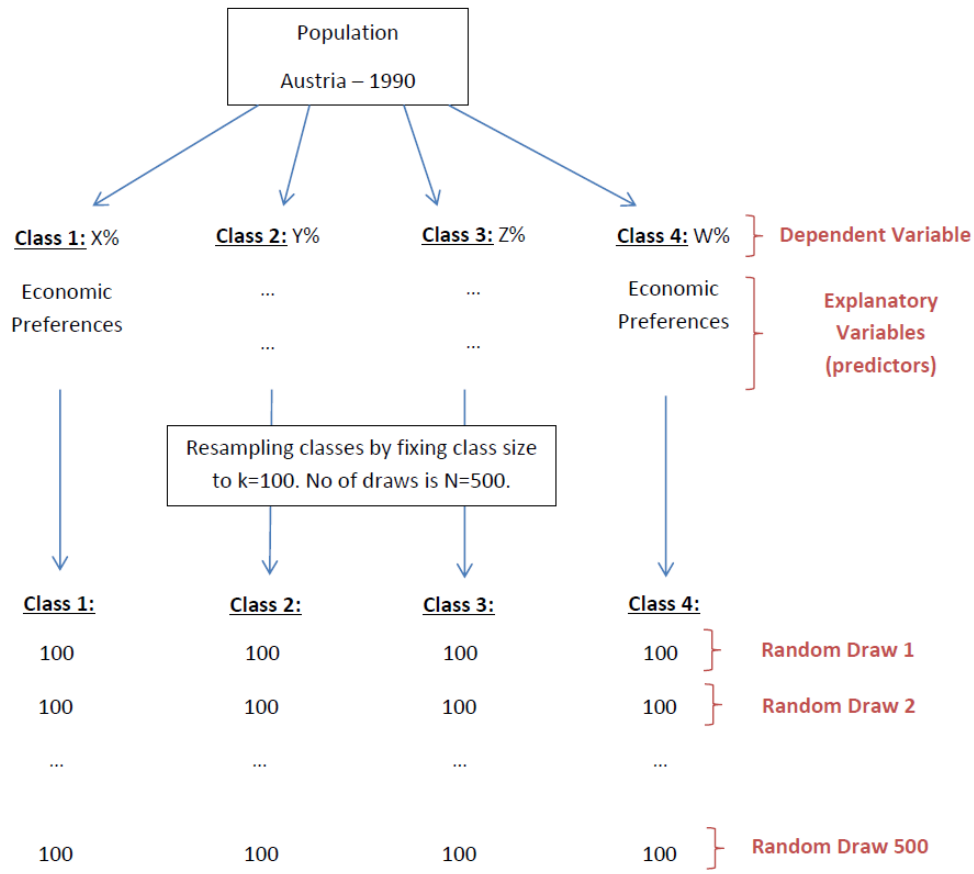

Figure E.1: The figure illustrates the resampling process used to balance class size and sample sizes of the country-year data sets.

F The Prediction Accuracy Rates for 1999 and 2008

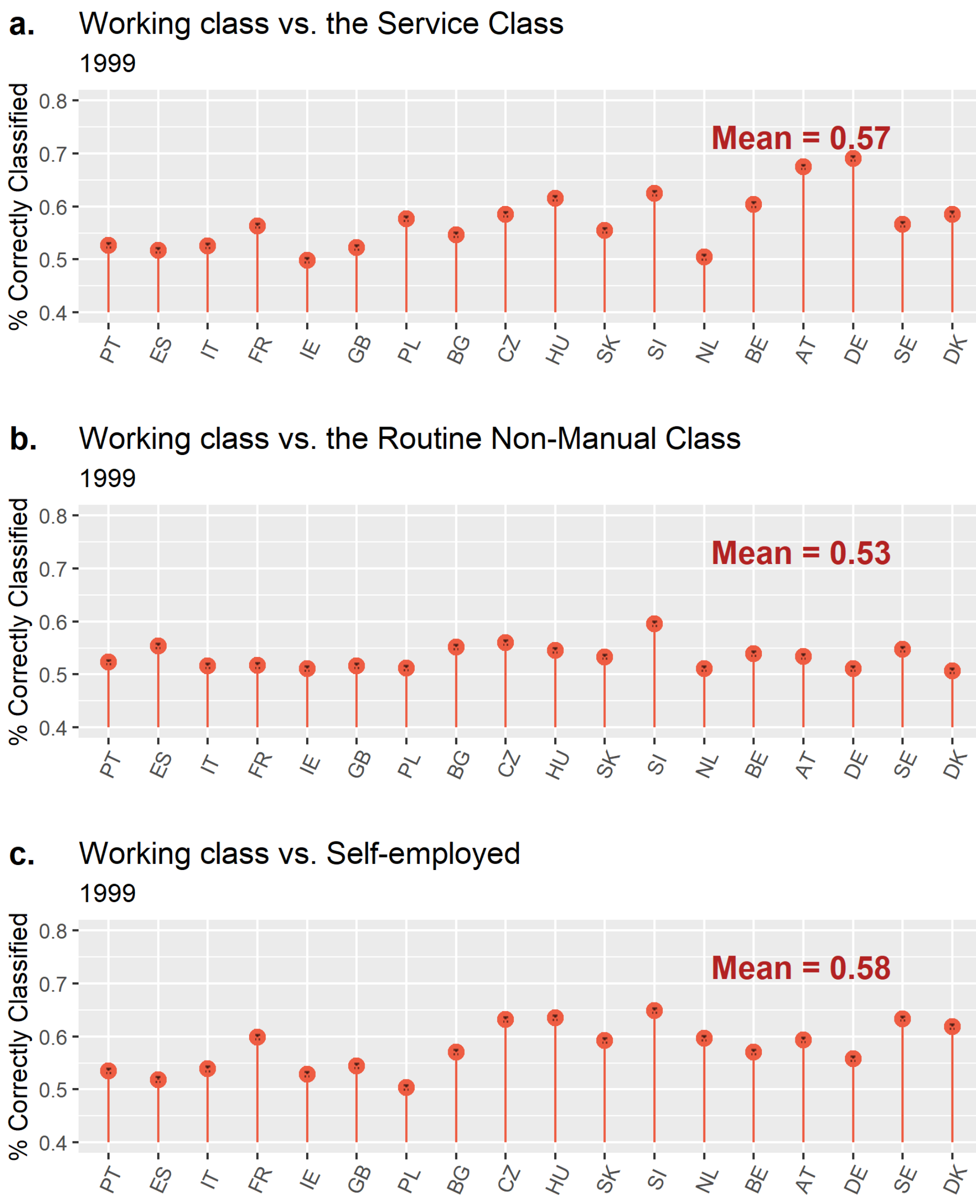

Figure F.1: Predictions with logistic regression.

Note: In the logistic regressions, the dependent variable is the binary class membership variable. Its levels are (a) working class and service class, (b) working class and the routine non-manual class, (c) working class and self-employed. The class membership of the respondents is predicted from their economic preferences only. The reported numbers are the percentages of correctly classified observations for the year 1999 in the predictions. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

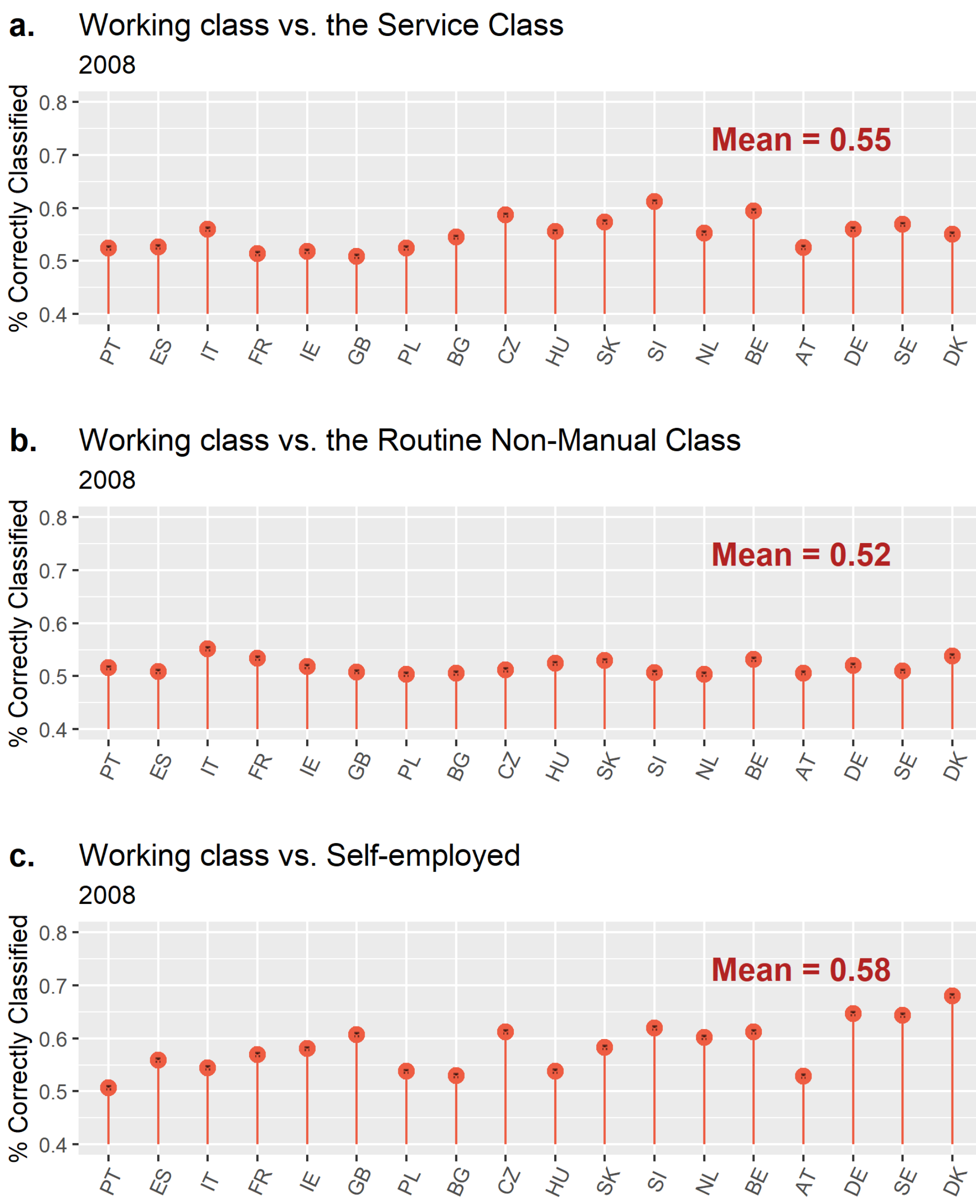

Figure F.2: Predictions with logistic regression.

Note: In the logistic regressions, the dependent variable is the binary class membership variable. Its levels are (a) working class and service class, (b) working class and the routine non-manual class, (c) working class and self-employed. The class membership of the respondents is predicted from their economic preferences only. The reported numbers are the percentages of correctly classified observations for the year 2008 in the predictions. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

G Additional Figures

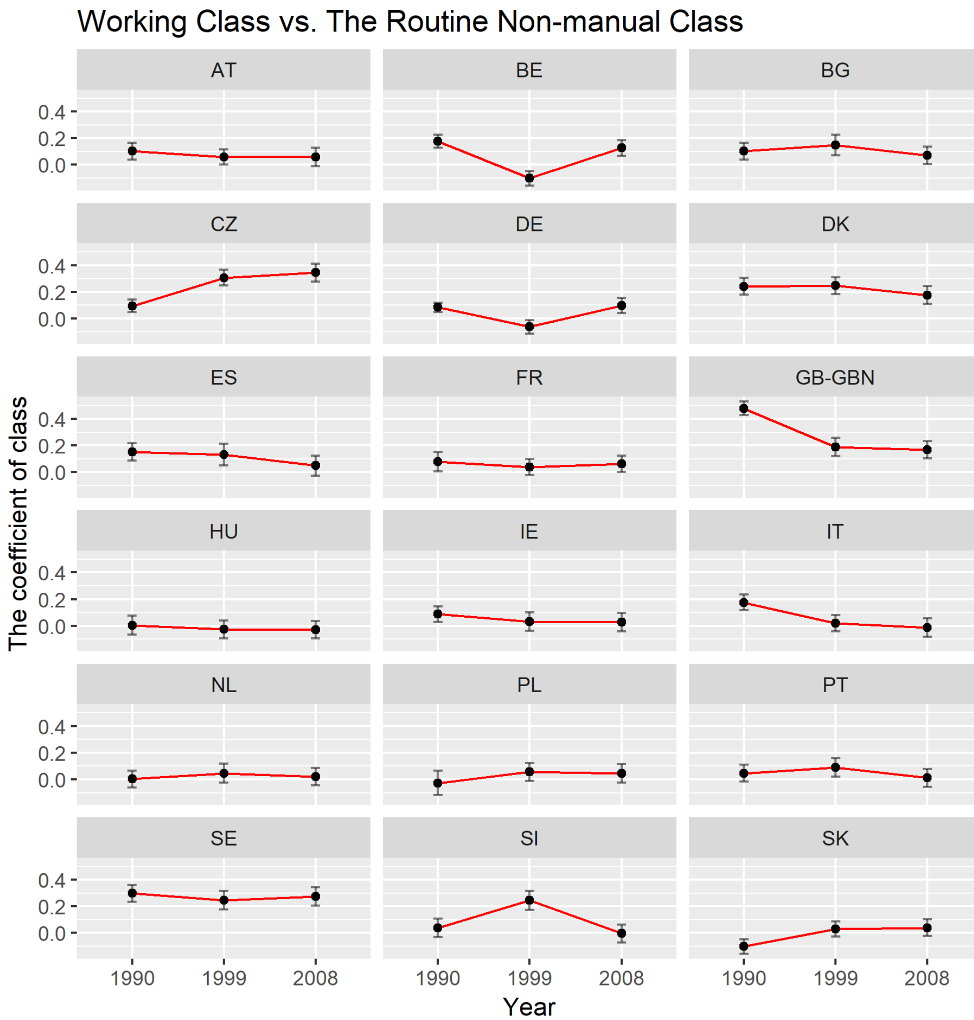

Figure G.1: Multilevel estimation.

Note: The dependent variable is the economic left-right position of the preferred political party by the respondent. The coefficients of the class membership variable are reported for each country and time point in the sample. The reported coefficients represent the difference between the economic left-right positions of the preferred parties by the working class and the routine non-manual class. The bars around the point estimates correspond to the standard errors of the coefficient estimates. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

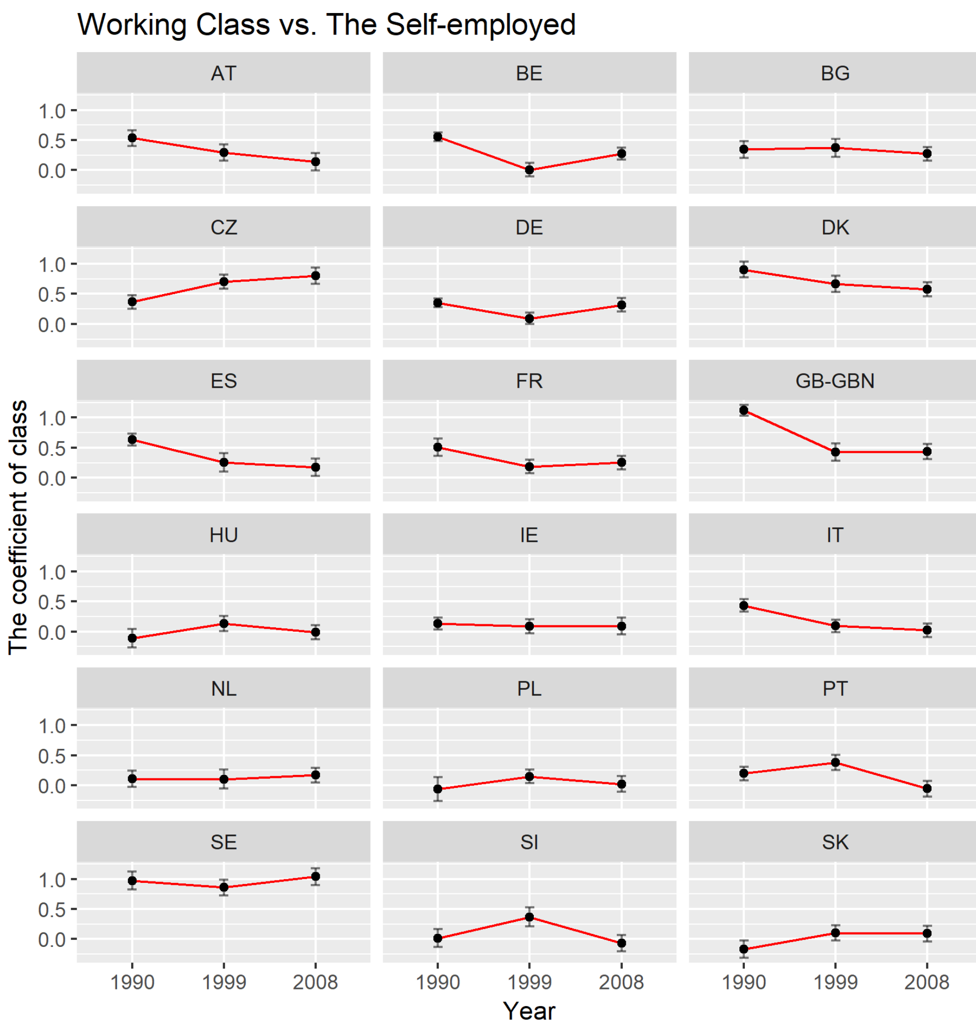

Figure G.2: Multilevel estimation.

Note: The dependent variable is the economic left-right position of the preferred political party by the respondent. The coefficients of the class membership variable are reported for each country and time point in the sample. The reported coefficients represent the difference between the economic left-right positions of the preferred parties by the working class and the routine non-manual class. The bars around the point estimates correspond to the standard errors of the coefficient estimates. Countries are labeled as follows: AT = Austria, BE = Belgium, BG = Bulgaria, CZ = Czech Republic, DE = Germany, DK = Denmark, ES = Spain, FR = France, GB = Great Britain, HU = Hungary, IE = Ireland, IT = Italy, NL = Netherlands, PL = Poland, PT = Portugal, SE = Sweden, SI = Slovenia, SK = Slovakia.

H Additional Tables

Table H.1: Panel regressions of class voting on class distinctiveness with case weights

	Dependent variable:		
	Strength of class voting		
	(1)	(2)	(3)
Class Distinctiveness	1.918^{***}	1.673^{***}	1.419^{***}
	(0.462)	(0.357)	(0.366)
Constant	-0.890^{***}	-0.774^{***}	-0.584^{***}
	(0.264)	(0.213)	(0.222)
Class-pair FE	✓	✓	✓
Country FE		✓	✓
Time FE			✓
Observations	162	162	162
R^2	0.225	0.696	0.711

Note: The reported results are from panel regressions. The dependent variable is the strength of class voting, which is obtained from the multilevel model estimation. The observations are weighted by their case weights. The case weights correspond to the inverse of the standard errors of the class membership variable in the multilevel estimation. *p<0.1; **p<0.05; ***p<0.01.