Essays in Applied Microeconomics

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

durch die

Rechts- und Staatswissenschaftliche Fakultät
der Rheinischen Friedrich-Wilhelms-Universität
Bonn

vorgelegt von

Laurenz Richard Kuno Günther

aus Bonn

2023

Acknowledgements

I am happy to thank my supervisors, Florian, Sebastian, and Moritz, for their excellent supervision. Florian and Sebastian spent countless hours discussing my crazy research ideas and always gave me good advice. They helped me out on numerous occasions, sometimes on very short notice or in the middle of the night. Even though the third supervisor usually only has ceremonial tasks, Moritz too supported me with great advice.

I also want to thank my BGSE cohort. We went a long way together, which to me looks very short in hindsight but felt like an eternity during the first year, which I would not have survived without them. Moreover, I am indebted to the BGSE and the University of Bonn in general. Bonn University has always been my dream place to study ever since my interest in economics was awakened during the great recession. Likewise, there was no graduate school I would have preferred over the BGSE. Both places gave me the opportunity to grow and supported my research financially without demanding anything in return except proctoring.

Among the many great people I met there, I want to single out Holger and Ilona. If you agree with me that this dissertation looks nice, both of us thank Holger for his TEXnical genius. In addition, I thank him for counting goals during our football matches as well as for his swift and genius programming support. Ilona was my main contact person for organizational stuff during my Ph.D. I must thank her for her patience and for doing all the things I hate to do.

Moving up to larger social entities, I wish to thank the German people. The Germans have created a magnificent society which I benefited from a lot. It gave me safety, freedom, and the opportunity to spend 30 years of my life just studying and doing science without demanding anything in return.

Speaking of kinship, I want to thank my family in general, particularly my brother, mother, and father. They have always been and still are the most important persons in my life and coined me more than anyone else for the better. In addition, I want to thank my ancestors. Many of them lived harsh and often miserable lives in pursuit of a better life for their descendants. Their suffering and tenacity built a stepping stone for me.

Of course, I could not finish these acknowledgments without talking about my friends. I divide them into three groups. First, there are those who helped me a lot

during my Ph.D. by giving constructive feedback on my papers or by raising stupid points that might be raised in seminars too. Most important among these are my great co-author, Timo, and Luca. Many others are listed in the acknowledgments of my papers. Still, I have to mention that many of these friends were suspiciously often absent from parties. I hope this will change in the future. A second group of friends did not support me in completing my thesis, but at least they did not hinder my progress. There are many of this kind, too many to be named. Come to Milan soon! Finally, there is a very special subset of my friends, members who did everything they could to hinder me in my scientific ambitions and the timely completion of my thesis. Their methods included social pressure and an annoying tenacity in permanently requesting excessive alcohol consumption during the week. There is no better example of this behavior than the infamous Bene. Had it not been for this group of friends, I would have finished my Ph.D. at least two years earlier and published at least two papers in the top five journals by now. I could not be more thankful to them. So much of life is about whether to consume one marshmallow today or two tomorrow. Of course, choosing "two tomorrow" will make you more successful regarding social status and wealth, but it also means that you never enjoy the fruits of your work. Without consuming marshmallows with these hedonistic friends, my life would not be worth living.

I will now move on the save the continent. I have waited for this opportunity since the great recession. Finally, my formal education is complete, and I will move to the best place imaginable to execute my plan. The only thing that makes me feel some regret is that it means having less time to spend with my family and friends.

Contents

List of Figures

List of Tables

Introduction

Many economic studies analyze the effectiveness of policies in dealing with perceived problems. For instance, a study might investigate the effectiveness of an increase in the inheritance tax in reducing economic inequality. However, multiple policies are often available to deal with a perceived problem, and all policies lead to costs. For instance, increasing the income tax is an alternative to an increase in the inheritance tax, and both increases could lead to an exodus of labor or capital, thereby hampering economic growth. It is often difficult to establish in general whether a policy is worth the costs and even harder to prove that it is generally superior to an alternative policy because preferences about the costs and benefits of policies differ. To stay with the example, one person might favor a tax increase because he or she prioritizes equality over economic growth, while another person might consider growth more important, therefore opposing a tax hike. Hence, thinking about the welfare consequences and implementation of policies requires examining the preferences of those affected.

Compared to the rich literature on potential policies, economists have devoted little attention to examining policy preferences and whether actual policymaking is in the interest of relevant stakeholders. This dissertation contributes to filling this gap in the literature by estimating policy preferences on various policy issues in multiple developed economies and comparing these preferences to actual policymaking.

Chapter 1 of this dissertation examines to what extent parliaments in Europe represent the policy attitudes of their voters and citizens. Most social scientists agree that the extent of representation is an important welfare criterion because it is a key measure of how democratic a country is (Golder and Ferland, 2017). Yet, there has been much less research on estimating its extent than on other welfare-relevant variables, such as unemployment or inequality. Just as it matters whether the unemployment rate equals 2% or 50%, it also makes a big difference whether the extent of representation is high or low. Similarly, it matters whether deviations of policymaking from people's preferences —representation gaps— arise on issues that are important or unimportant to people, whether representation gaps on different issues go in the same direction, and whether all major parties are biased in the same direction relative to voters. Chapter 1 provides an extensive analysis of representation gaps in Europe, addressing these and related questions.

I document large, significant, and systematic representation gaps. European citizens and voters prefer more redistribution and less state intervention than is desired or implemented by parliaments. These representation gaps on economic topics are large. For instance, the difference between policymaking and the redistribution preferences of the average voter is about as large as the average difference in redistribution preferences between poor and rich Europeans. Representation gaps on non-economic issues are even larger. Policymaking is more left-wing than voters, or citizens prefer regarding nearly all non-economic issues. Even most christian democratic and conservative parliamentarians are more left-wing on these issues than the average voter. Representation gaps are most pronounced regarding immigration, punishment for criminals, and gender relations. This pattern holds for almost all countries considered and when comparing European voters in general to the European Parliament.

These findings cannot be explained by standard models of electoral competition and raise the question of why voters do not eliminate representation gaps by adjusting their voting decisions. To answer this question, I build a stylized formal model. The intuition it formalizes is that voters do not only care about the distance of their policy preferences and the policy positions of a party but also about the competence of a party in implementing its positions effectively. Hence, voters might vote for a party even though its policy positions are relatively different from their own attitudes if they believe that this party is more competent than its rivals. I show formally that this model can explain the existence of representation gaps and derive empirical predictions, which, as I show, are born out by the data.

Finally, Chapter 1 relates representation gaps to two key recent developments in Europe — a loss of trust in democratic institutions and the rise of populism. It shows that citizens whose policy attitudes are less well represented by their parliaments are less satisfied with the way democracy works in their country, less trusting in democratic institutions, and less likely to vote, even after controlling for many demographic characteristics. Moreover, it documents that representation gaps are tightly linked to the rise of populism. Populist politicians are more likely than non-populist politicians to identify representation gaps, more motivated to fill them, and the policies they propose also contribute to filling representation gaps.

Are representation gaps intrinsically problematic, independent of potential negative effects on trust in democratic institutions? I address this question in Chapter 2 of this dissertation. I start from the idea that whether representation gaps are a problem or not depends on the reasons for their existence. If representation gaps reflect an information advantage of politicians, they might not only be acceptable but even desirable. In contrast, if representation gaps result from differences in deep preferences (like risk attitudes or time preferences) between voters and politicians, they are more likely problematic from a welfare perspective. Accordingly, it is important to find out whether information asymmetries between voters and politicians drive representation gaps.

To do so, Chapter 2 develops a new estimation technique based on vignette studies. I focus on the topic of immigration because representation gaps are particularly large there, and recent research has shown that Europeans hold strongly biased beliefs about immigrants (Alesina, Miano, and Stantcheva, 2023). The key idea behind the technique is to combine data on citizen attitudes toward detailed described hypothetical asylum seeker vignettes and data on the characteristics of actual asylum seekers. By matching these two data sources, I estimate the asylum-seeker attitudes of citizens, who potentially hold biased beliefs about asylum seeker characteristics, if they were well-informed about their actual characteristics.

I find that most Europeans would have opposed immigration by most asylum seekers that really immigrated if they had been well-informed about their characteristics. This "informed opposition" is common across many demographic dimensions. For instance, a majority of respondents in all 15 European countries examined reject immigration by the average asylum seeker. The same is true for most Europeans on the left and the right of the political spectrum, both sexes, those with and without an immigration background, and all age groups. These findings indicate that the desire of Europeans to reduce the immigration of asylum seekers is not driven by misinformation. Hence, the representation gap regarding immigration is unlikely to be driven by biased beliefs of ordinary citizens either.

Chapter 3 of my dissertation analyzes preferences toward inherited inequality. Economic inequality has grown in recent decades in most developed countries (Roine and Waldenström, 2015). Much of this increase in inequality is driven by people profiting to differential extents from their parents or other related persons — inherited inequality (Bowles and Gintis, 2002). Still, surprisingly many citizens oppose redistributive policies, like inheritance taxation, even if they would profit financially from them (Bartels, 2008; Stantcheva, 2021).

How can we explain this opposition to redistributive policies? Chapter 3 argues that such opposition can be explained through fairness preferences. Previous research has shown that many people adhere to a meritocratic fairness ideal according to which inequality that results from differential efforts is deemed fair while inequality resulting from differential luck is judged as unfair (Cappelen, Falch, and Tungodden, 2020). However, meritocrats might judge inherited inequality that results from the efforts of those who bequest fair and unfair at the same time. On the one hand, according to the meritocratic fairness ideal, those who bequest earned their resources and should be able to spend them freely, which includes transferring them to a person of their choice. On the other hand, those who inherit these resources did not earn them and are therefore equally deserving, which renders inequality between them unfair by meritocratic standards. Hence, it is unclear whether meritocrats accept or oppose inherited inequality resulting from the efforts of those who bequest. If they accept such inequality because they prioritize fairness toward those who bequest, this might explain why people oppose redistributive preferences regarding inherited inequality.

Chapter 3 reports the results of a survey experiment that was designed to test this explanation. In the experiment, impartial spectators, who constitute a representative sample of adult US citizens, redistribute payments between pairs of individuals. The experiment varies a) whether initial payments are based on a random draw or effort and b) whether spectators redistribute between individuals who have worked themselves or who merely benefit from the work of real-life friends. Redistribution levels are substantially higher if inequality is based on luck instead of effort, suggesting that many US citizens hold a meritocratic fairness ideal. However, whether individuals worked themselves or merely inherited their initial payoffs does not matter notably for spectators' redistribution decisions. This suggests that many US citizens accept inherited inequality as long as it is merited at some stage, which may explain why many people oppose redistributive policies like, for instance, inheritance taxation.

References

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva. 2023. "Immigration and Redistribution." *Review of Economic Studies* 90 (1): 1–39. DOI: 10.1093/restud/rdac011. (Visited on 03/27/2023). [3]

Bartels, Larry M. 2008. *Unequal Democracy: The Political Economy of the New Gilded Age*. STU - Student edition. Princeton University Press. URL: https://www.jstor.org/stable/j.ctt7t9ks (visited on 04/26/2023). [3]

Bowles, Samuel, and Herbert Gintis. 2002. "The Inheritance of Inequality." en. *Journal of Economic Perspectives* 16 (3): 3–30. DOI: 10.1257/089533002760278686. (Visited on 04/09/2023). [3]

Cappelen, Alexander W., Ranveig Falch, and Bertil Tungodden. 2020. "Fair and Unfair Income Inequality." en. In *Handbook of Labor, Human Resources and Population Economics*. Edited by Klaus F. Zimmermann. Cham: Springer International Publishing, 1–25. DOI: 10.1007/978-3-319-57365-6_130-1. (Visited on 03/19/2023). [3]

Golder, Matt, and Benjamin Ferland. 2017. "Electoral Systems and Citizen-Elite Ideological Congruence." In *The Oxford Handbook of Electoral Systems*. Edited by Erik S. Herron, Robert J. Pekkanen, and Matthew S. Shugart. Oxford University Press, 0. DOI: 10.1093/oxfordhb/9780190258658.013.11. (Visited on 04/12/2023). [1]

Roine, Jesper, and Daniel Waldenström. 2015. "Long-Run Trends in the Distribution of Income and Wealth." en. In *Handbook of Income Distribution*. Vol. 2, Elsevier, 469–592. DOI: 10.1016/B978-0-444-59428-0.00008-4. (Visited on 03/19/2023). [3]

Stantcheva, Stefanie. 2021. "Why Do We Not Support More Redistribution?: New Explanations from Economics Research." en. In *Combating Inequality*. Edited by Olivier Blanchard and Dani Rodrik. The MIT Press, 263–70. DOI: 10.7551/mitpress/13469.003.0040. (Visited on 03/19/2023). [3]

Chapter 1

Political Representation Gaps in Europe[*]

1.1 Introduction

Representative democracies build on the idea that voters elect members of parliament (MPs), who in turn represent the policy attitudes of voters (Golder and Stramski, 2010). The extent to which MPs' policy *decisions* align with voters' attitudes is called substantive representation, while the lack thereof is referred to as a representation gap (Pitkin, 1967; Andeweg, 2012). This paper estimates the extent of substantive representation in Europe and documents large, significant, and systematic representation gaps. It also examines how representation gaps can arise despite political competition, and assesses their relationship with the rise of populism and distrust in democratic institutions.

The main analysis builds on a survey dataset I compiled from several existing anonymized MP and voter surveys. This dataset contains the policy attitudes regarding a wide range of political issues of 31,461 citizens (27,852 of which are voters) and 2,074 parliamentarians. The samples are representative of the underlying universes of citizens and voters of 27 European countries as well as 15 national European parliaments and the European parliament, respectively. A key advantage of

this dataset is that responses of citizens and parliamentarians were elicited through identical items, enabling comparability.

I find that the policy attitudes of MPs and their voters differ strongly, significantly, and systematically. MPs are much more liberal/left-wing on nearly all cultural issues and more in favor of EU integration in nearly all 27 countries. Attitude differences regarding immigration and punishment for criminals are particularly large. On a cultural issue index, the distance between the mean voter and the mean MP is more than half of the distance between the mean social democratic and the mean conservative MP. Alternatively, the average MP is about three-quarters of a standard deviation of citizen attitudes more left-wing than the average voter. Moreover, the average MP of any established party family is culturally further to the left than the average voter. In the economic policy dimension, MPs tend to be more market-oriented and anti-redistribution than voters. The largest economic attitude differences regard redistribution, where differences between voters and MPs resemble those of the poor and the rich. However, attitude differences in the economic dimension tend to be smaller than in the cultural and EU dimensions. Furthermore, they are strongly heterogeneous across countries. Moreover, social democratic MPs are more state-oriented on economic issues than the average voter, while liberal and Christian democratic MPs are more market-oriented. Thus, the average voter has established voting options to their left and right on economic issues but only to their left on cultural topics.

Attitude differences matter because, as I show, most MPs base their decisions on their own attitudes rather than on the attitudes of their voters. When asked whether a parliamentarian should follow the opinion of their voters or their own opinion, 69% of national MPs and 84% of members of the European Parliament state that the parliamentarian should follow their own opinion.

To compare *decisions* rather than attitudes of parliaments to the attitudes of voters, I develop an estimation framework. It formalizes the idea that attitude differences result in representation gaps to the extent that MPs decide according to their own attitudes rather than according to the attitudes of their voters. Consequently, the framework takes both attitude differences between voters and MPs and the extent to which MPs aim to implement voters' attitudes as arguments. Estimates of representation gaps based on this framework are very similar to attitude differences between MPs and voters. To validate these survey-based estimates, I use a second, hand-collected dataset on politicians' and ordinary citizens' voting and initiation behavior in 126 Swiss referendums. Differences in voting behavior in referendums between Swiss MPs and voters closely resemble representation gaps as estimated from survey data. Moreover, this dataset reveals that representation gaps can be stable over many decades.

The existence of representation gaps is robust to many alternative specifications. For example, representation gaps are larger when comparing citizens instead of voters to MPs, and they are not driven by MPs or voters with extreme attitudes. They

are also larger on issues that both voters and parliamentarians find more important. Demographic characteristics can explain 30–40% of representation gaps, and the remaining gaps significantly differ from zero.

Standard probabilistic voting models of electoral competition cannot explain the existence of representation gaps (Adams, 1999; Lin, Enelow, and Dorussen, 1999; Schofield, 2004, 2007; Schofield and Zakharov, 2010). These models extend the logic of Downs (1957) to the competition of multiple parties in multiple policy dimensions. As in Downs (1957), parties are assumed to maximize their vote share and competition forces parties to converge at the electoral center.

To explain how representation gaps can be stable despite electoral competition, I build a stylized model of electoral competition between two parties akin to that used in Downs (1957). Motivated by my empirical findings, I alter the utility function of the parties relative to Downs (1957) and standard probabilistic voting models such that they are not motivated by winning the election but only by the implemented policy. My model also assumes that voters see parties as differentially competent, as I verify empirically. I prove that every equilibrium of the model features a representation gap in the direction of the bliss point of the party that is seen as more competent. Intuitively, the party that is seen as more competent can shift its policy position away from the median voter and still win the election due to its perceived competence advantage. It also profits from this shift because it is policy-motivated. The model makes testable predictions about the distribution of European parties' perceived competence. For instance, it predicts that culturally left-wing parties are considered more competent than culturally right-wing parties, which enables the former to win elections despite positions far from the electoral center. I find strong empirical support for all of these predictions.

In contrast, I find no empirical evidence in favor of several alternative explanations, including lobbyism, a desire of politicians to protect minorities from a "tyranny of the majority," or the possibility that MPs' policy attitudes change when they get elected.

Finally, I examine the relationship between representation gaps and two crucial political developments: distrust in democratic institutions and the rise of populism. Even after controlling for demographic characteristics, I find that citizens whose policy attitudes differ more from their representatives have less trust in democratic institutions and are less likely to vote. In the public discourse, populists claim to be a reaction to representation gaps (Mudde and Kaltwasser, 2017). According to their narrative, policymaking deviates from the people's policy preferences. Populists claim to be aware of these representation gaps and to have entered politics to fill them. I test the claims contained in this "populist narrative." I find that populist MPs are more likely to correctly identify representation gaps and to state that an MP should follow the opinion of their voters than non-populist MPs. Moreover, populist parties provide a mixture of left-wing economic and right-wing cultural policy positions that many citizens demand but few parties have supplied so far. Their policy

positions differ more strongly from those of non-populist parties on issues where representation gaps are larger. Finally, the attitudes of populist MPs are, in contrast to those of non-populist MPs, close to voters' attitudes on cultural issues. These findings are consistent with the populist narrative and the claim that populists are a reaction to representation gaps.

Consequently, my findings help to clarify the rise and characteristics of populist parties.[1] Recent studies have shown empirically that factors like trade exposure (Colantone and Stanig, 2018a,b; Autor et al., 2020), economic insecurity Funke, Schularick, and Trebesch (2016), Algan et al. (2017), Fetzer (2019), and Gabriel, Klein, and Pessoa (2022) and immigration (Halla, Wagner, and Zweimüller, 2017; Harmon, 2018; Dustmann, Vasiljeva, and Piil Damm, 2019; Hangartner et al., 2019; Tabellini, 2020) increase the vote shares of populist parties. It is less well understood why these events increase the populist vote share (Guriev and Papaioannou, 2022). There is no conclusive evidence that crises generally lead to populist voting. For instance, popular government support usually increases temporarily during international crises (Mueller, 1970) and the Covid pandemic increased government support and trust in political institutions (Esaiasson et al., 2021; Kritzinger et al., 2021).

Why do only certain events strengthen challenger parties, and not other established parties? Why do these challenger parties see the "pure people" in a struggle with the "corrupt elite," hold right-wing cultural positions, and have anti-media and anti-expert attitudes (Mudde and Kaltwasser, 2017)? Representation gaps offer an explanation. Suppose that voters vote for parties close to them in policy space on issues that are relevant at the moment (Bakker, Jolly, and Polk, 2018). If representation gaps exist, challenger parties can fill empty policy space and will ultimately rise. Because established parties are more market-oriented and culturally left-wing than voters, challenger parties are more state-oriented and culturally right-wing than established parties. More specifically, this line of reasoning suggests that restrictive immigration policies, stricter sentencing, and increased redistribution decrease the vote share of populist parties. Indeed, recent studies have found that increased redistribution weakens populists (Albanese, Barone, and Blasio, 2022), (large) increases in immigration strengthen populists (Barone et al., 2016; Becker and Fetzer, 2016; Halla, Wagner, and Zweimüller, 2017; Dinas et al., 2019; Dustmann, Vasiljeva, and Piil Damm, 2019; Edo et al., 2019; Hangartner et al., 2019; Ajzenman, Aksoy, and Guriev, 2022; Guriev and Papaioannou, 2022), and being soft on crime leads to more right-wing voting (Drago, Galbiati, and Sobbrio, 2020).[2]

---

1. Surveys on populism are provided by Berman (2021) and Guriev and Papaioannou (2022). Several recent papers have also argued that the rise of populism is partly due to mainstream parties failing to represent voters' policy attitudes (Berger, 2017; Grzymala-Busse, 2019; Lindner et al., 2020; Berman and Kundnani, 2021; Gethin, Martínez-Toledano, and Piketty, 2021; Bó et al., 2023). However, these papers have not estimated representation gaps.

2. My findings suggest that closing other representation gaps might reduce the vote share of populist parties too. I am not aware of studies that have examined whether this is the case.

Challenger parties likely rise especially quickly when issues become relevant where representation gaps are large, but not when issues become relevant where no representation gap exists. Representation gaps are particularly large when it comes to redistribution and immigration, which explains why the populist vote share rose especially quickly during the global financial crisis and the refugee crisis that highlighted inequality and immigration. This explanation complements recent evidence that increased relevance of cultural topics makes cultural attitudes stronger drivers of policy views (Bonomi, Gennaioli, and Tabellini, 2021) and that changes in voters' issue priorities are the main driver behind the rise of the populist right (Danieli et al., 2022).

To reduce trust in their opponents, challenger parties can argue that established parties do not represent the people and instead make policy according to their own attitudes. Given that representation gaps exist, this statement contains a kernel of truth, and it is likely to damage the reputations of their rivals. This explains why challenger parties propagate the populist narrative.

Moreover, representation gaps help to explain why many citizens vote for populist parties, even though it has been shown that having populists in power reduces economic growth (Funke, Schularick, and Trebesch, 2016) and damages democratic institutions (Funke, Schularick, and Trebesch, 2016; Bellodi, Morelli, and Vannoni, 2021; Morelli, Nicolò, and Roberti, 2021; Docquier, Peluso, and Morelli, 2022). My results indicate that many voters are faced with two similarly unattractive options because they must choose between established parties that do not implement the policy attitudes of the voters and populists who threaten democratic institutions. Some voters, particularly those more distant from the culturally left-wing established parties, might consider populists the lesser evil.

The fact that policymaking is generally incongruent with mass attitudes also helps to explain why populist parties can be strong in countries not affected by crises. Hence, just sitting out or preventing crises might not eliminate populism. The current paper highlights another policy alternative: established parties could close representation gaps. They could do so either by convincing the public that their policies were in the public's interest or by changing policymaking. Which alternative is preferable depends on whether representation gaps result from information- or value differences between voters and MPs. Because, as I show, representation gaps persist even when comparing MPs to the educated or those very interested in politics, value differences seem to be at least partly responsible for representation gaps.

The present paper also adds to the theoretical political economy literature. I show empirically that voters consider some parties much more competent than others, and that most politicians are policy-motivated. While it is well known that adding these insights individually to the stylized model of Downs (1957) does not produce representation gaps (Persson and Tabellini, 2002), I show that adding both at the same time does. Predictions of a model that incorporates both insights, are well supported by an empirical analysis. At the same time, these modifications do not

notably increase the model's complexity. Therefore, it might prove useful to modify other models of electoral competition in the same way.

Finally, this paper contributes to the literature on representation. Economists have focused on the numerical representation of sociodemographic groups in positions of power, called descriptive representation (Golder and Ferland, 2017).[3] In contrast, I focus on substantive representation. My empirical analysis reveals that the association between descriptive and substantive representation is weak. Whether descriptively overrepresented or underrepresented groups are better represented substantively depends on the policy dimension and the demographic variable. For instance, immigrants are better represented substantively than natives on cultural topics, but less well on economic issues. The only numerically underrepresented group that is systematically disadvantaged substantively is the poor. This raises the question of how effective affirmative action policies are. First, it suggests that such policies might not change substantive representation as desired. Second, it indicates that improving the substantive representation of descriptively underrepresented groups might not be desirable because they might not be underrepresented substantively.

Outside of economics, most research on substantive representation stems from political science.[4] The present paper makes a methodological contribution to this literature by presenting a stylized formal framework to define and estimate substantive representation from survey data. A formal framework might be useful because existing studies use identical terms to refer to similar but different concepts (Kertzer, 2022). The framework formalizes the insight that attitude differences between voters and MPs matter little if MPs base their decisions on voters' attitudes. Therefore, the framework uses information on attitude differences and the intention of politicians to implement their voters' attitudes. In contrast, most previous studies only assess attitude differences (Costello, Thomassen, and Rosema, 2012; Ferland, 2016; Dalton, 2017; Kübler and Schäfer, 2022).

Moreover, I investigate several potential causes of representation gaps that have not been examined to date. These include the cause primarily investigated in the present paper, namely an interaction of heterogeneous perceived competence and

---

3. Studies have primarily focused on women (Chattopadhyay and Duflo, 2004; Beaman et al., 2009; Duflo, 2012; Besley et al., 2017) and ethnic minorities (Pande, 2003; Banerjee and Pande, 2007; Munshi and Rosenzweig, 2015), often in developing countries. Recent papers have examined descriptive representation regarding socioeconomic background in Europe (Bó et al., 2017, 2023).

4. See Kertzer (2022) for a recent discussion and meta-analysis. Most studies focus on single countries (Bühlmann, Widmer, and Schädel, 2010; Andeweg, 2012; Holmberg, 2012; Andreadis and Stavrakakis, 2017; Schakel and Hakhverdian, 2018; Hakhverdian and Schakel, 2022; Jaime-Castillo and Coller, 2022; Lesschaeve, 2022) and/or estimate substantive representation at one point in time (Costello, Toshkov, et al., 2021; Hakhverdian and Schakel, 2022; Lesschaeve, 2022). Coverage of European countries is particularly low (Shapiro, 2011). Notable exceptions are Costello, Thomassen, and Rosema (2012) and Dalton (2017), who compare policy *attitudes* of candidates to the European Parliament and European voters, and Evans and Hall (2019), who analyze whether the positions of parties and voters *change* in the same direction over time.

policy-motivated politicians. Finally, this study demonstrates the relevance of representation gaps for the debates on political trust and populism.

The paper proceeds as follows. Section 1.2 presents the estimation framework of substantive representation. After that, Section 1.3 derives theoretical predictions. Section 1.4 provides an overview of the data used in the empirical part of the paper. Within this empirical part, Section 1.5 documents the pattern of substantive representation in Europe, Section 1.6 examines potential causes of this pattern, and Section 1.7 relates representation gaps to political trust and populism. Section 1.8 concludes the paper.

1.2   A Formal Framework for Estimating Representation Gaps

By substantive representation, I refer to the degree of congruence between the *decisions* of parliamentarians and the policy attitudes of voters. By a representation gap, I refer to a lack of substantive representation, a difference between the policy decision from the policy decision preferred by voters.

Formally, consider a particular single policy dimension d. Different values on d refer to different ordered policy alternatives. Voters and MPs have preferences over policy alternatives. The set of voter preferences is denoted by X_V, the set of MP preferences is denoted by X_{MP}, and mp is the index for MPs. MPs implement a policy. When doing so, they need not support a policy that coincides with their preferences. I abstract from the decision-making process of the parliament. There are two types of MPs: MPs of type p(olicy motivated) only care about their own preferences on d. MPs of type r(epresentation motivated) only care about representing voters. Let the representation intention of the parliament (RI) be the share of MPs that are representation motivated: $RI = \frac{\#r}{\#p + \#r} = 1 - \frac{\#p}{\#p + \#r}$. I define a representation gap (RG) for a given dimension d and metric $m(\cdot)$ as

$$RG_m^d \; := \; m(\text{policy alternatives supported by MPs}) - m(X_V).$$

Because standard probabilistic voting models make predictions about representation gaps in terms of means, I assume that m is the ordinary mean with equal weights in the main text, but results can be generalized to any metric. For simplicity, I assume that MPs of type r support the policy decision preferred by the overall mean voter. Then

$$
\begin{aligned}
RG_{mean}^d &= \text{mean}(\text{policy alternatives supported by MPs}) - \text{mean}(X_V) \\
&= \text{mean}(X_V) \cdot \frac{\#r}{\#p + \#r} + \text{mean}(X_{MP}|mp \in p) \cdot \frac{\#p}{\#p + \#r} - \text{mean}(X_V) \\
&= [\text{mean}(X_{MP}|mp \in p) - \text{mean}(X_V)] \cdot (1 - \text{RI}) \\
&:= ARG \cdot (1 - \text{RI}).
\end{aligned}
\tag{1.1}
$$

Equation 1.1 proposes a way to measure representation gaps in terms of the behavior of MPs and preferred behavior by voters from data on attitudes of MPs and voters. It states that representation gaps in means can be measured as the difference between the mean attitude of voters and the mean attitude of policy-motivated MPs (which I label attitude representation gap or ARG for short) multiplied by the representation intention. This suggests that just comparing the attitudes of all MPs and all voters might lead to biased estimates for representation gaps. Intuitively, attitude differences don't lead to representation gaps if all MPs are representation motivated.

The RG on an issue is zero if the means of voters' and MP's attitudes are identical. It can be positive or negative, making it possible to identify which direction MPs differ from voters. Because I want to compare voters to the MPs that actually sit in parliament, RGs are measured without accounting for demographic differences between voters and MPs. Moreover, in contrast to most political scientists, I do not claim that differences in attitudes arise causally due to the election of politicians (Kertzer, 2022).[5][6]

Representation gaps might be relevant because they might lead to important reactions by voters and citizens. For instance, voters might reconsider their voting decision and enable new parties and social movements to rise. Moreover, representation gaps might also reduce voters' trust in democracy, leading to protest voting and vote abstention. Large representation gaps on topics that are important to many voters might even lead to attacks on politicians or revolutions. These effects might manifest even if representation gaps result from wise policymaking of politicians who might be, for example, more informed than voters. Additionally, representation gaps might be the symptom of corruption where public resources are used for the benefit of MPs to the detriment of the overall social welfare if they result from differences between MPs and voters regarding values or deep preferences like risk attitudes. Section 1.B discusses in more detail under what circumstances representation gaps might be a problem.

1.3 Theoretical Predictions of Standard Models

Economic theory on electoral competition and voting has largely been in the tradition of the spatial approach of Downs (1957). In this approach, different political positions are related to each other by how similar they are. More similar policy positions are thought of as being closer to each other. Voters have ideal points in this policy space, and political candidates announce to implement certain points if they

---

5. However, I do examine to what extent such differences are due to a direct effect of being elected and how much of them can be accounted for by demographic differences between MPs and voters to examine their origins Section 1.F.4 and Section 1.F.5 in the appendix.

6. Section 1.A discusses how this measure differs from other measures used in the political science literature.

get elected. Voters are typically assumed to base their vote solely on the distance between their ideal points and the points announced by the candidates and vote with higher probability for candidates closer to them. Parties take ideal points of the electorate as given and are usually assumed to strategically position themselves in the policy space to maximize their chances of election (Persson and Tabellini, 2002).

In European countries more than two parties compete in a policy space that has been found to be multidimensional by empirical studies (Bakker, Jolly, and Polk, 2012). Probabilistic voting models are the standard way to model electoral competition in such situations (Hinich, Ledyard, and Ordeshook, 1972; Coughlin and Nitzan, 1981). They often feature a (sometimes unique) convergent equilibrium. In such a convergent equilibrium all parties position themselves at the electoral *mean* on each policy dimension (Caplin and Nalebuff, 1991; Adams, 1999; Lin, Enelow, and Dorussen, 1999; McKelvey and Patty, 2006; Schofield, 2007). This implies that the mean positions of parliamentarians and voters are identical. Hence, RGs in terms of means are predicted to be small. Convergence of all parties to the mean is sufficient but not necessary for small RGs. Candidates could position themselves symmetrically around the electoral mean, which could result in low RGs in terms of means. While representation gaps have not been examined directly as far as I know in these models, simulation exercises suggest that if equilibria are not convergent, candidates are indeed positioned symmetrically around the electoral mean (Adams, 1999; Schofield, 2007; Schofield and Zakharov, 2010).[7] These results are robust to the lifting of several assumptions, like strategic instead of sincere voting (McKelvey and Patty, 2006), different candidate motivations (Adams, 1999), the distance measure voters use to compare their own positions to those of candidates (Lin, Enelow, and Dorussen, 1999) and heterogeneity in perceived competence of the candidates (Schofield, 2007). Moreover, Adams (1999) combines Monte Carlo simulations with parameter estimates from empirical studies to argue that the conditions for a convergent equilibrium are often fulfilled in the real world. These theoretical results suggest that representation gaps regarding the mean should be nonexistent or at least small and not systematic.

7. There are deterministic and probabilistic spatial voting models. Under deterministic voting, the probability that a voter votes for any candidate is one if he is closest to the voter and zero otherwise. Models of probabilistic voting add a random error term to the voting behavior, such that the probability of voting for a candidate increases continuously as the distance decreases. Most models of deterministic voting do not feature pure strategy Nash equilibria unless the distribution of ideal voter points fulfills strong symmetry conditions (Plott, 1967; Eaton and Lipsey, 1975; McKelvey and Wendell, 1976; Schofield, 1978; McKelvey, 1979; Cohen and Matthews, 1980; Schofield, 1983; McKelvey and Schofield, 1987; Caplin and Nalebuff, 1991; Banks, 1995; Saari, 1997). In particular, the median voter theorem by Black (1948) and Downs (1957) does not hold. Probabilistic voting models usually have equilibria. In addition, they are more realistic than deterministic voting because the latter requires much political knowledge from voters.

1.4 Data

The main analysis builds on a survey-based dataset that I generated by harmonizing comparable surveys among parliamentarians and citizens. Survey data on MPs come from the European Candidate Study 2009 (Weßels, 2013) and wave one of the Comparative Candidate Survey (CCS) (CCS, 2016). Both surveys were conducted in a combined effort of local institutions with much experience in collecting MP data. The European Candidate Study 2009 was fielded to nearly all candidates for the 2009 European Parliament, while the CCS was fielded to nearly all candidates to national parliaments for all elections between 2005 and 2013 in the following European countries: Austria, Belgium, the Czech Republic, Denmark, Estonia, Finland, Germany, Greece, Hungary, Iceland, Ireland, Italy, the Netherlands, Norway, Portugal, Romania, Sweden, Switzerland, and the United Kingdom. All responses were elicited several months after the election and include information on whether the candidate was elected, which enables me to identify elected members of parliament. The European Candidate Study includes information on 6,558 candidates from 260 parties which implies a response rate of 20.5% overall, and information on 169 out of the 736 elected MEPs that made up the 2009-2014 European Parliament which is equal to a response rate of MEPs of 23%. The CCS data contains data on 14,392 candidates at a response rate of 40.81% and 1,905 elected national MPs.

Data on voter attitudes come from the European Voter Study 2009 (Egmond et al., 2017) and the Swiss Electoral Studies 2007 (Selects, 2009). The European Voter Study 2009 was conducted alongside the European Candidate Study 2009 and designed to match it as closely as possible. The timing, structure, and wording of all policy-attitude questions are identical. The Voter Study was fielded to a representative probability sample of roughly 1,000 citizens in each 27 EU country in 2009. I add data on a representative sample of 4,392 Swiss voters from the Swiss Electoral Studies 2007 because the European Voter Study 2009 does not contain data on Swiss voters.

The combined dataset includes policy attitudes and demographic data about 2,074 elected European Parliamentarians and 31,461 citizens from 27 European countries. Citizens are representative of the respective adult population and, as shown in Section 1.C, sample parliamentarians are representative of the universe of parliamentarians.

The European Candidate Study 2009 and the European Voter Study 2009 include 14 identical policy attitude items. Subjects were provided with a statement like "Immigration to [Country] should be reduced significantly" and could then indicate how much they agreed or disagreed with the statement or refuse to answer. Out of the 14 attitude items, 10 were elicited on a 5-point Likert scale, one had three answer opportunities, and one had 11 answer opportunities. Table 1.H.1 provides wordings and more information for all questions.

Voters and MEPs were asked about their opinions about immigration, assimilation of immigrants, the importance of private enterprise, same-sex marriage, state ownership, state intervention, abortion, punishment for criminals, redistribution, teaching authority in schools, direct democracy, gender relations, EU unification, and EU membership. All question wording and response categories were identical for MEPs and voters. Of these 14 items, seven overlap precisely with items given to national MPs. These include assimilation, same-sex marriage, abortion, state intervention, redistribution, EU unification, and EU membership. The items on punishment for criminals that MEPs and voters saw differ slightly from the one that national MPs responded to. Voters and MEPs were asked how much they agreed or disagreed with the following statement:

> *People who break the law should be given much harsher sentences than they are these days.*

National MPs were asked how much they agreed or disagreed with a slightly different assertion:

> *People who break the law should be given stiffer sentences.*

Nonetheless, I compare the responses of national MPs and voters on that item because, as I show, the difference in formulation likely creates a bias that works against my finding.

I also use the 2014 and 2019 Chapel Hill Expert Surveys (CHES) (Polk et al., 2017; Jolly et al., 2022). The CHES contains estimates of established policy experts of positions of various European parties on several policy issues, such as redistribution and immigration, many of which match policy issues from the survey dataset. Section 1.C shows that measures for party positions based on the candidate survey data are highly correlated with measures for party positions based on the CHES, which suggests that candidate surveys provide valid data.

To estimate representation gaps from behavioral data, I make use of an original referendum dataset. In principle, referendums are ideal for comparing the decisions of voters, MPs, and parties because voters and MPs are confronted with the same well-defined issue, which enables comparability. Take the people's initiative "Against the construction of Minarets" in Switzerland as an example. The initiative wanted to ban the construction of minarets by constitutional article. 57.5% of Swiss citizens who voted in 2009 on that initiative voted in favor of it. In contrast, over 72% of members of the Swiss lower house voted against the initiative. Opposition was even larger in the upper house, where 39 out of 42 delegates voted against it. Similarly, the vote share of parties that officially positioned themselves against the initiative amounted to over 64%, and the government too openly positioned itself against it. The common interpretation of these results has been that the Swiss voters were much more right-wing on issues related to assimilation and Islam than the political elite.

However, one should be cautious in drawing such inferences from a single referendum. Comparing the political elite with the population requires a larger dataset of voting behavior on referendums. Not every referendum is useful for estimating RGs, however. First, the initiative behind the referendum needs to be on one narrowly defined topic. For instance, initiatives that advocate for a complex package of laws do not fulfill this criterion. If one found more opposition among the population than among MPs, it would be hard to assess which law was responsible for the disagreement. Second, to enable comparability with survey data, these topics should match the topics of the survey data. Third, it must be clear whether a yes or a no vote indicates a right-wing or left-wing position. To the best of my knowledge, no available dataset on referendums fulfills these conditions. Hence, I created it by building on the database Swissvotes (Swissvotes, 2021). Swissvotes is the primary data source for referendums in Switzerland. It contains information on all referendums in Switzerland since 1884. Among this information are the numbers of yes and no votes of voters and MPs. It also contains the share of votes jointly received in the last national election for the parties that officially declared themselves in favor of the initiative and the same for its opponents and the government's official position.

To generate the referendum dataset, I read up on all referendums contained in Swissvotes and classified each referendum into topic categories used in the EES to enable comparability. Referendums that cannot be classified in that way are labeled as belonging to the category "other." Some initiatives might be classified into several topics. Therefore, I recorded up to three topics for each initiative. This procedure resulted in 126 classified referendums over 130 years. I also coded the political direction of a referendum. This variable refers to whether the referendum would shift policy-making to the political left or the political right on that particular issue. For instance, I coded the Minaret Referendum as mainly belonging to the "assimilation" topic, matching a category from the survey data, and as a right-wing initiative. Classifying referendums as right-wing or left-wing is simple in most cases. However, I also included a variable that captures how clear the classification was. Finally, I coded whether the referendum was initiated by the parliament, "ordinary" citizens, or whether it resulted from an interaction of the two groups. A discussion on descriptive statistics of the referendum dataset can be found in Section 1.D.1 in the appendix.

1.5   Documenting Representation

1.5.1   Documenting the Representation Intention of MPs

The survey data contains items in which MEPs and MPs were directly asked what they would do if their policy attitudes differed from that of their voters. MEPs were asked the following question:

How should, in your opinion, a member of European Parliament vote if his/her own opinion does not correspond with the opinion of her/his voters?

Possible answers included "Should vote according to her/his party's opinion" and "Should vote according to her/his voters' opinion." National MPs were asked:

An MP in a conflict between own opinion and the constituency voters should follow:

Possible answers included "own opinion" and "voter opinion." All answers were anonymous. Both items contrast the candidate with those who voted for him. Parliamentarians are likely particularly inclined to represent their own voters compared to all voters or all citizens. Hence, one might expect that parliamentarians would be less likely to signal a high representation intention if these questions had been asked in terms of voters or citizens in general. Similarly, one might fear that social desirability bias biases the estimate for the representation intention upward.

Despite that, of the 72% of MEPs who answered the question, about 84% stated that the MEP should follow their own opinion. Similarly, of the approximately 77% of national MPs that responded, about 69% stated that the national MP should follow their own opinion rather than the opinion of their voters. Hence, neither MEPs nor MPs seem to have a high representation intention. This suggests that differences in attitudes translate into actual policymaking.

Is there important heterogeneity among MPs regarding their representation intention? Figure 1.G.1 in the appendix depicts the average representation intention of various demographic subgroups of national MPs. Male, less educated and younger MPs have higher representation intention. However, even in these subgroups majorities state that an MP should follow his own opinion rather than the opinion of his voters. There is more variation between countries. The lowest average representation intention exist in Germany, the Netherlands, Island, the UK, and Switzerland. The highest are observed in Hungary, Romania, Austria, and Italy. Hungary and Romania are striking outliers as they are the only countries in which only a majority states that an MP should follow the opinion of his voters. In sum, this evidence suggests that the low average representation intention is not driven by specific demographic groups but a majority position in most demographic groups and in most countries.

1.5.2 Documenting Attitudes Differences between Voters and MPs

1.5.2.1 Comparing Attitudes of Voters and MEPs by Policy Issue

To make the magnitude of attitude differences easy to interpret I divide all policy attitude variables by the standard deviation of citizens of that variable. I then estimate regressions of the following form by OLS:

$$std(a)_{i,s} = \alpha + \beta \cdot \mathbb{1}[\text{MEP}]_{i,s} + \varepsilon_{i,s}. \tag{1.2}$$

$std(a)_{i,s}$ is the standardized attitude of subject s on issue i and $\mathbb{1}[\text{MEP}]_{i,s}$ equals one if s is an elected MEP and zero if s is not an MEP and voted in the European Parliament election. Regressions are weighted to adjust for differences in population sizes. As a result, I compare representative samples from the voters for the 2009 European Parliament election with a representative sample of MEPs. I am interested in β as a descriptive measure for differences in policy attitudes between the mean MEP and the mean European Parliament election voter.
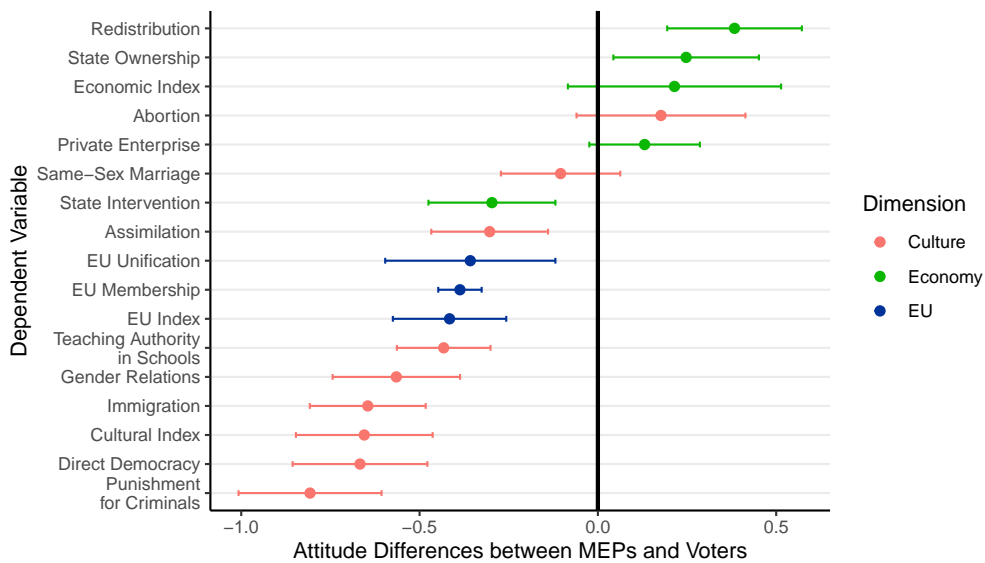
Figure 1.1. Attitude Differences between Voters and MEPs by Issue

Note: The vertical axis shows different policy issues or policy dimensions. The horizontal axis shows OLS estimates for β's from Equation 1.2. Regressions compare MEPs with a representative sample of those who voted in the European Parliament election. All variables were scaled to have a standard deviation (in terms of citizen attitudes) of one. Higher values indicate that MEPs are more right-wing on an issue than voters. 95% confidence intervals are based on standard errors clustered at the country level.

Figure 1.1 shows β's and 95% confidence intervals based on standard errors clustered at the country level for all 14 policy issues and indexes for three policy dimensions, to be discussed below. Higher values indicate that the mean MEP is more right-wing on an issue compared to the mean voter. For instance, a preference for a restrictive immigration policy is intuitively right-wing, while a wish for more redistribution is intuitively left-wing. Many economists and political scientists argue that the policy space in Europe consists of at least two dimensions: one cultural (sometimes also labeled social or GAL-TAN) and one economic dimension (Bakker, Jolly, and Polk, 2012; Norris and Inglehart, 2019). Recent studies have argued that topics related to EU integration should be considered a third separate dimension (Bakker, Jolly, and Polk, 2012). I follow the literature and categorize all issues into three categories Economy, Culture, and EU, as indicated by colors in Figure 1.1. I

also calculate the attitude of each subject on a dimension as the mean of this subject's positions on all topics belonging to this dimension with equal weights.

Regarding cultural issues, MEPs do not differ significantly from voters on abortion and same-sex marriage. On all other cultural issues, MEPs are significantly more left-wing than voters. Differences amount to 40% to 80% of a standard deviation of citizen attitudes. They are largest for punishment for criminals, direct democracy (voters are more in favor of it), immigration and gender relations. On the cultural dimension index, the mean MEP is about two-thirds of a standard deviation more left-wing in the cultural dimension than the mean voter. Similarly, the mean MEP is about 42% of a standard deviation more in favor of EU integration than the mean voter. Differences for economic variables are less systematic. Voters are significantly more opposed to state intervention but more in favor of redistribution and state ownership than MEPs while attitudes do not differ significantly regarding the role that private enterprise should play in the economy. On the economic index, the mean MEP is estimated to be about 0.2 standard deviations more right-wing than the mean voter, but the difference is not significant ($p \approx 0.16$). Hence, it is important to distinguish between economic topics and other topics when examining representation.

I compare voters to MEPs in this section to make use of all attitude items I have data on. Differences in attitudes are robust to many alternative specifications. First, Figure 1.G.2 in the appendix compares national MPs to voters. Results are qualitatively very similar but quantitatively larger. For instance, national MPs are over 75% of a standard deviation of citizen attitudes more left-wing in the cultural dimension than their national voters. Similarly, Figure 1.G.4 in the appendix shows the mean positions of national MPs together with those of MEPs, voters, and citizens. MEPs and national MPs tend to have similar attitudes but MPs are a bit further away from voters and citizens on most issues. Moreover, attitudes of MPs and MEPs differ more from those of citizens than from those of voters. Finally, Section 1.F.1 in the appendix shows that MEPs and voters consider the same topics important and find cultural and economic topics similarly important while they find topics related to the EU much less important. If anything, differences in attitudes are larger on topics that are considered to be more important by MEPs or voters.

1.5.2.2 Cross Party Heterogeneity in Policy Attitudes

Figure 1.2 visualizes the European policy space by policy dimension. I normalize the space such that the most right-wing attitudes equal one and the most left-wing attitudes equal zero. I depict the mean attitudes of voters and MEPs of the main European party groups surrounded by 95% confidence intervals. Again, I compare representative samples of voters for the 2009 EP election with a representative sample of members of the European Parliament.

Party groups line up as expected in the economic dimension. The European Parliament is polarized with liberals and christian democrats/conservatives on the right and social democrats on the left. The mean voter's position in the economic dimension lies between the mean positions of the major right-wing and left-wing party groups. Weighted t-tests and median tests confirm that christian democrats/conservatives and liberals are significantly more right-wing than the average voter, while social democrats are significantly more left-wing. Hence, the average voter has established options to his left and right. Consequently, coalitions between economically right-wing and left-wing parties are likely to feature economic policy attitudes close to the mean voter. Indeed, the mean attitudes of MEPs and voters are similar, and their difference is only marginally significant according to a t-test despite the large sample size ($p \approx 0.07$). However, Mood's median test finds a significant difference ($p \approx 0.003$).
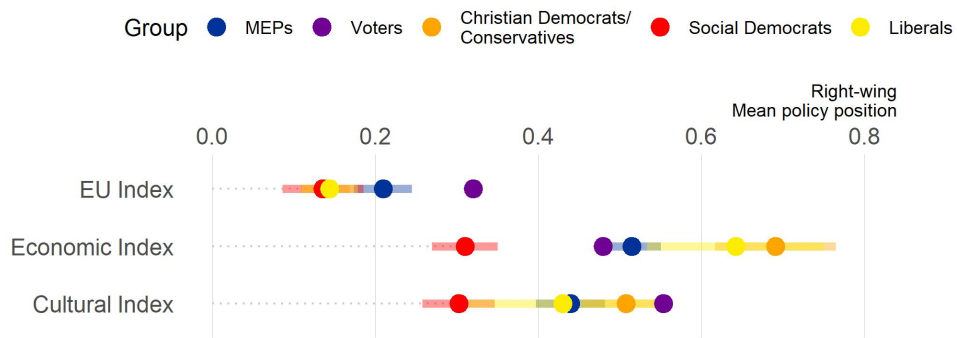
Figure 1.2. Mean Attitudes of Voters and MEPs by Policy Dimension
Note: The vertical axis shows different policy dimensions. The horizontal axis shows the positions of voters and parliamentarians. All variables were scaled to range from zero to one and such that higher values indicate a position that is more right-wing. Dots indicate the mean-attitudes of voters, christian democrats/conservatives, social democrats, liberals and greens respectively. Data is pooled across Europe.

The pattern looks qualitatively different for the other two dimensions, even though party groups again line up as expected. On cultural issues, christian democrats/conservatives are most right-wing, social democrats are most left-wing, and liberals are in between. However, voters hold attitudes to the right of all major party families. Weighted t-tests and median tests confirm that social democrats and liberals are more left-wing than voters. For instance, social democratic parties are more than 25 percentage points to the left on cultural issues than voters, which amounts to more than one-quarter of the range of the cultural index. Mood's median test finds that the median position of christian democrats/conservatives is significantly more left-wing than the median position of voters, while a weighted t-test just fails to find a significant difference ($p \approx 0.1003$). MEPs as a whole are more than

12 percentage points of the range of the outcome variable more left-wing than voters, and differences are highly significant according to mean and median tests. Another way to interpret this result is that the mean voter differs more from the mean MEP than the mean liberal MEP differs from the mean conservative MEP or by about as much as the mean liberal MEP differs from the mean social democratic MEP. In the EU dimension, the mean attitudes of all party groups are close to one another, while voters are about 10 percentage points of the range of the outcome variable more opposed to EU integration. The mean voter differs much more strongly from the mean MEP than the mean MEPs of either party group differ from one another. Weighted t-tests and median tests confirm that means and medians of all party groups and MEPs as a whole are significantly more pro-EU than those of voters.

Hence, when focusing on the EU dimension and to a lesser degree in the cultural dimension no established party group represents the attitudes of the mean voter well, and all party groups are biased relative to him in the same direction. This implies that average attitudes of coalitions between established parties differ from the attitude of the mean voter too. Thus, voters cannot achieve high attitude congruence with parliamentarians on the culture and EU dimension by voting for the main established party families. To do so, they have to increase the vote shares of parties positioned far to the right of established parties. Figure 1.G.5 in the appendix shows that MEPs of the major European party groups are not only to the left of the mean voter regarding the cultural and EU dimension in general but also on most cultural and EU-related issues individually.

1.5.2.3 Distributions of Policy Attitudes

Figure 1.3 depicts the two-dimensional culture-economy policy space. Both dimensions range from zero to one and higher values indicate attitudes that are more right-wing. I focus on the economy and culture dimension because Section 1.F.1 in the appendix shows that voters and MEPs consider them much more important than the EU dimension. Figure 1.3 depicts 2D densities of voter and MEP preferences in the policy space. Figure 1.G.3 in the appendix shows the same distribution of voter attitudes in the same policy space together with MEP attitudes distinguished by party family.

The distribution of voters is unimodal and quite symmetric. In particular, there is no strong correlation between the economic and cultural dimensions. This simple structure makes it easy for parties to determine the vote-maximizing policy positions in the center of the voter distribution and makes a convergent equilibrium more likely in standard models (Plott, 1967; McKelvey and Wendell, 1976; McKelvey, 1979; McKelvey and Schofield, 1987; Schofield, 2007). Moreover, most voters are not extremists but have center-right cultural attitudes.

The distribution of MEP attitudes has a more complex form. It has two modes. The first is in the lower-left corner where many social democrats and green MPs
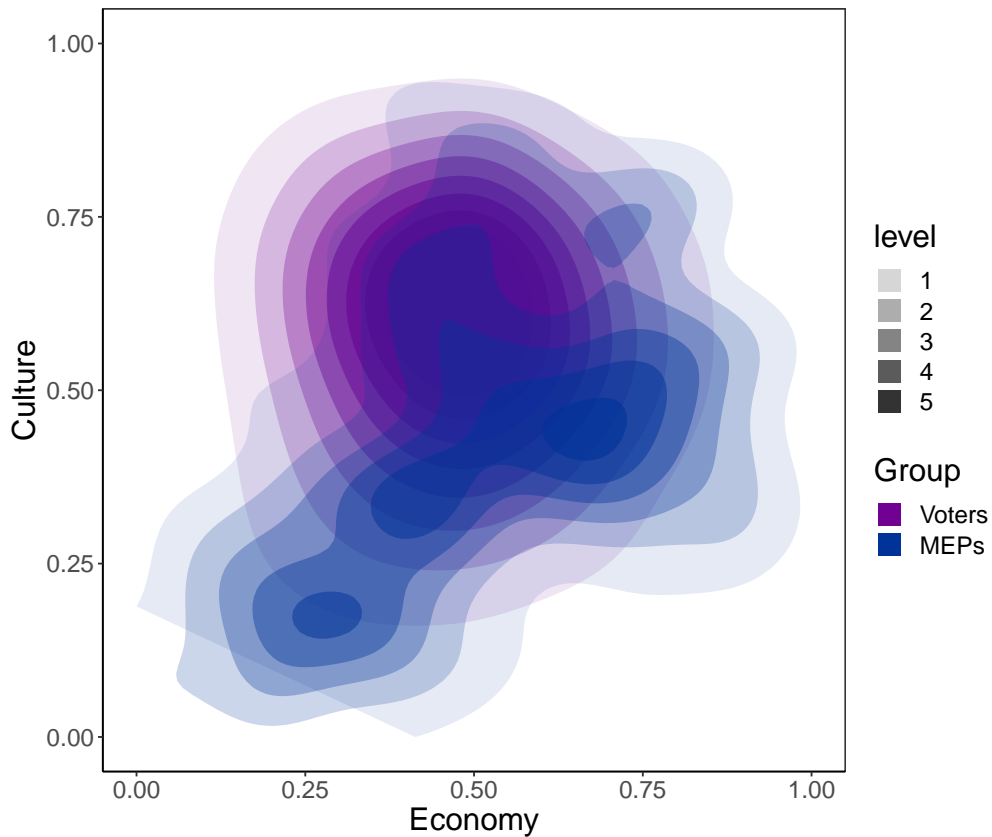
Figure 1.3. Distribution of Voters and MEPs in the Culture–Economy Policy Space
Note: The level of transparency indicates the density. The density is higher in less transparent areas. For both policy dimensions higher values indicate a stance that is more right-wing. Data is pooled across Europe.

are located as Section 1.F.1 reveals. The second is closer to the middle of the graph but still below the center of the voter-attitude distribution. Section 1.F.1 shows that many christian democrats and conservatives are located there. Figure 1.3 reveals a strong positive correlation between attitudes in the cultural and economic dimensions. Variation in attitudes is higher among MEPs than among voters. This is particularly pronounced regarding the economic dimension. The MEP-attitude distribution does not only have a different shape than the voter-attitude distribution; it is also shifted. Most MEPs, even those in the "upper" right part of the MEP attitude distribution, are below the center of the voter-attitude distribution. As a result, few MEPs have policy attitudes close to the electoral center. Most are more culturally left-wing. Voters with conservative cultural and left-wing economic attitudes are particularly badly represented. Hence, differences in mean attitudes are not driven by voters or MEPs with extreme policy attitudes. Rather, most MEPs are culturally more left-leaning than most voters.

1.5.3   Estimates of Representation Gaps

1.5.3.1   Representation Gaps by Country

Figure 1.4 combines information on representation intention and attitudes differences between MPs and voters through the estimation framework of representation gaps (Equation 1.1) to estimate representation gaps in means for 27 European countries and the European Union taken as a whole. Regions are displayed on the horizontal axis. The heights of the bars indicate the magnitude of the representation gaps, which are measured in standard deviations of citizen attitudes. Positive values indicate that parliamentarians of the respective country tend to be more right-wing/conservative/anti-EU regarding the dimension than voters, while a negative representation gap shows the opposite result. A value of zero indicates that there is no representation gap. I pool national MPs and MEPs to increase the sample size and because MEPs and national MPs have very similar policy attitudes, as shown in Figure 1.G.4 in the appendix. In the case of the EU as a whole, I exclude national MPs. There, I also weigh to adjust for population sizes in order to compare representative samples of MEPs with a representative sample of voters from the EU.
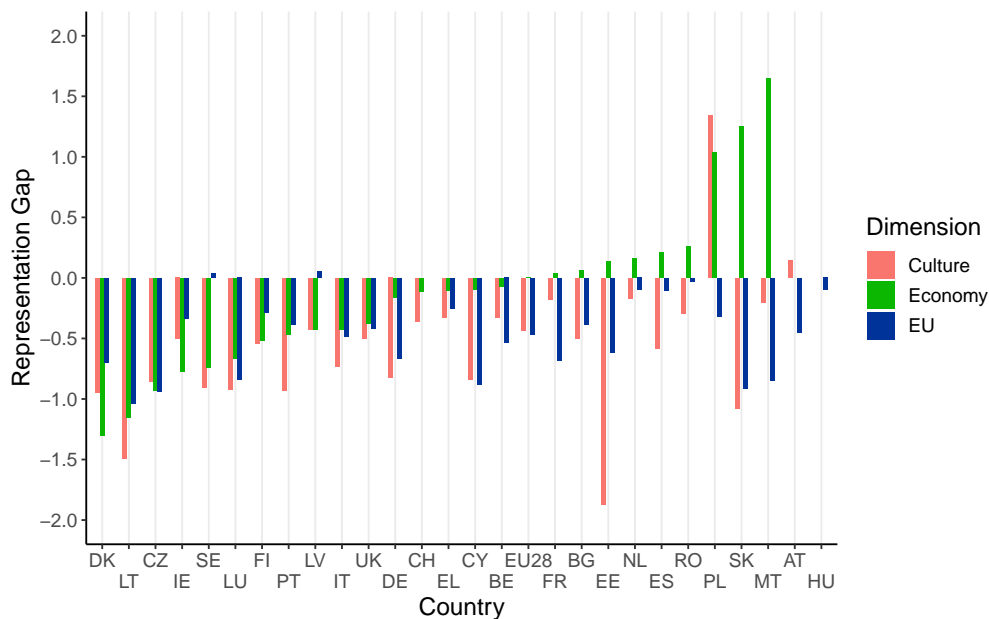
Figure 1.4. Representation Gaps by Country and Policy Dimension
Note: Bars show representation gaps by country and dimension between voters and parliamentarians (MPs and MEPs) from the same country. The larger the value, the more right-wing parliamentarians are compared to voters from their country. Slovenia, the economic and cultural dimension in Hungary, and the economic dimension in Austria are missing due to missing data.

Cultural representation gaps are negative in all regions except Poland and Austria, indicating that policymaking is more left-wing than voters prefer in nearly all European countries and the EU as a whole. Magnitudes are substantial. Policy-

making is more than one standard deviation of citizen attitudes more left-wing than the mean voter in several countries. The largest cultural representation gaps emerge in Estonia, Lithuania, Denmark, Sweden, Slovakia, Portugal and Poland. They are smallest in Switzerland, Greece, France, the Netherlands and Austria. The patter is similar regarding the EU dimension although representation gaps tend to be smaller. Voters are slightly more in favor of European integration than the policymaking in only two countries, Latvia and Sweden. Policymaking by the European Parliament is about half a standard deviation more left-wing regarding cultural and EU-related topics than those who voted in the 2009 European Parliament election.

Representation gaps in the economic dimension follow a different pattern. They are negative in 16 countries, positive in nine countries, and virtually zero regarding the EU. Moreover, economic representation gaps are smaller but still sizeable than cultural ones in absolute value for most countries. They are the smallest in the EU, Belgium, France, Bulgaria, and Cyprus. In many other countries, they are large. In some countries, voters are much more market-oriented than parliamentarians, like in Poland and Slovakia, but also in Malta. The opposite is true in, for instance, Denmark, Lithuania, and the Czech Republic. Consistent with the finding that most parliamentarians are policy-motivated, Figure 1.G.6 in the appendix shows that attitude differences between voters and parliamentarians look very similar to representation gaps in all countries.

1.5.3.2   Unequal Representation of Demographic Groups

The European Parliament, like most parliaments, numerically over-represent those with a high living standard, men, natives, the old, and the educated. Are the these groups also better represented substantively by the European Parliament? Figure 1.5 helps to answer this question. It depicts representation gaps and 95% confidence intervals by demographic group and policy dimension. Data is pooled for all EU member states from 2009 and weighted to account for cross-country differences in population size. The vertical axis displays various groups of voters. The three panels refer to representation gaps on the three political dimensions. Higher absolute values indicate a larger representation gap in means between the voter group and the European Parliament. Positive values indicate that the mean attitude of the voter group is more left-wing on the dimension than policymaking of the European Parliament and negative values have an analogous interpretation. I distinguish between subgroups regarding the self-assessed standard of living, sex, whether the voter lives in an urban or rural area, whether the voter has an immigration background, religious denomination, the degree of religiosity, above versus below the median age, and education.

Figure 1.5 reveals little heterogeneity between various demographic subgroups regarding the economic dimension. Men and women, those living in rural or urban areas, immigrants and natives, and the educated and uneducated are all sim-
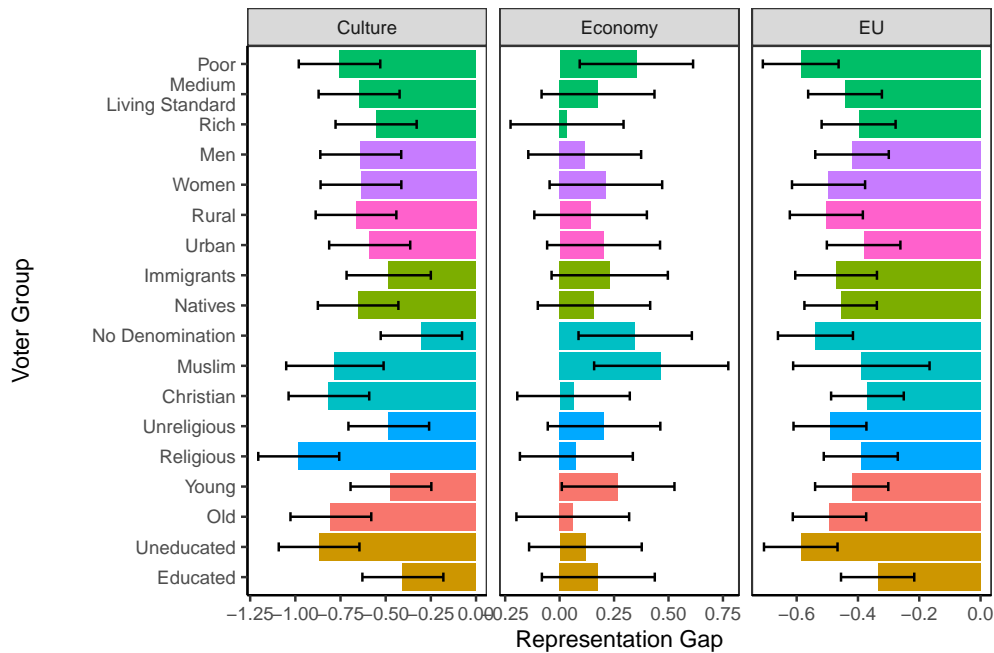
Figure 1.5. Representation Gaps by Voter Group
Note: The horizontal axis shows representation gaps in means between demographic subgroups of those who voted in the 2009 European Parliament elections and members of the European Parliament. The vertical axis displays the demographic subgroups. Data is pooled for all 27 EU member states from 2009 and weighted to account for cross-country differences in population size. I also depict 95 % confidence intervals.

ilarly well represented substantively by the policymaking of the European Parliament. However, the rich are much better represented substantively than the poor, Christians are better represented than those without a denomination or Muslims, and the old are better represented than the young. In general, the numerically over-represented groups, the rich, men, natives, the old, and the educated, are also substantively better represented in four out of five cases. The only exception is the uneducated, who are better represented substantively than the educated.

In the EU dimension, the rich, men, and the educated have lower RG than the poor, women, and the uneducated, respectively. However, the largest differences between subgroups arise regarding education, not income. Moreover, there is no notable difference in representation gaps between immigrants and natives, and the young have lower representation gaps than the old. Hence, the numerically over-represented group is better represented substantively in three out of five cases. On the cultural dimension, the rich have a somewhat lower RG than the poor, and the educated have a much lower representation gap than the uneducated, but representation gaps for men and women are very similar. Immigrants have a lower representation gap than natives, and the young have a much lower RG than the old. Hence, the numerically over-represented have lower representation gaps in only two out of five

cases, while the opposite is true in two cases. Taken together, these findings show that numerical over-representation does not necessarily lead to substantive over-representation. In particular, the relationship depends on the policy dimension and the demographic variable. While numerical and substantive representation seems to be positively related in the economic dimension, this is less the case in the cultural dimension.

Figure 1.5 also helps to examine the origins of representation gaps. On the one hand, representation gaps could result from competition between large social groups. If power is unequally distributed between these groups, the more powerful group could force the parliament to bias policymaking in its favor, away from the mean voter. Such models would predict that representation gaps go in opposite directions for demographic subgroups of some variable. For instance, the rich could be more right-wing than the parliament, and the poor more left-wing. If the parliament is closer to the position of the rich, this could explain an overall representation gap. In such a model, parliamentarians balance group interests in a biased way. On the other hand, representation gaps could result from differences in attitudes between a small elite of parliamentarians and most people of the society, including majorities in all demographic subgroups. In such a model, parliamentarians do not balance group interests but hold attitudes that differ from all major groups in the same direction. As can be seen in Figure 1.5, all subgroups are more left-wing in the economic dimension than MEPs. At the same time, all groups are significantly more right-wing in the cultural and EU dimension than MEPs. This is evidence against the formerly mentioned "balancing" models and is consistent with the low representation intentions of parliamentarians.

1.5.4 Validating Survey-Based Estimates with Referendum Data

Estimates based on survey data might lead to biased results for several reasons (Heckman, Jagelka, and Kautz, 2021). For instance, politicians might interpret questions differently than voters or have stronger incentives than citizens to give "politically correct" responses. Hence, it is essential to validate survey-based data with behavioral data, which are less vulnerable to biases. When voting on referendums, MPs and voters make decisions with real political consequences, mitigating potential biases. Consequently, this section calculates representation gaps between voters and parliamentarians by comparing their voting behavior on referendums and uses these estimates to examine where survey-based estimates are reliable.

To this end, I restrict the analysis to Switzerland because it is the only European country with a sufficiently large number of referendums. Hence, I add a dataset that includes information for 126 Swiss referendums on whether the referendum initiative would push the country to the right or the left and the shares of voters and national Swiss MPs who voted in favor of the initiative. I calculate the representation gap for a referendum r as

$$RG_r = \begin{cases} \text{share of "yes"-voting voters} - \text{share of "yes"-voting MPs}, & \text{if } r \text{ is left-wing} \\ \text{share of "yes"-voting MPs} - \text{share of "yes"-voting voters}, & \text{if } r \text{ is right-wing.} \end{cases}$$

To calculate the representation gap for a policy issue i, I take the mean with equal weights of all RG_r belonging to issue i, matching the categories in the survey data. To do so, I pool referendum data from a 20-year time window around the date surveys were administered.

Referendum data only contains information on yes-no decisions. To make the Likert-scale data from the surveys comparable to it, I use the share of those holding a right-wing stance on an issue as a measure for the position of a group. For instance, I calculate attitude differences regarding the punishment of criminals as the share of Swiss voters who agree or strongly agree that punishment for criminals should be more severe minus the share of Swiss MPs who agree or strongly agree with that statement.

Figure 1.6 shows a scatter-plot of survey-based estimates of representation gaps ($\hat{RG}$) and estimates based on referendum voting (RG) by issue. The horizontal axis shows the estimate for representation gaps on an issue based on survey data, while the vertical axis shows the corresponding estimate as estimated from referendum voting. The graph also displays a 45° line on which all data points would lie if the two measures were exactly the same. The better this line describes the relationship between the two measures, the more reliable the survey data. Reassuringly, the two measures are positively associated, and the 45° line describes their relationship well. This suggests that survey-based estimates are a reliable indicator of representation gaps.

In Figure 1.6, I aggregate referendums from 1997 until 2017 because the number of topics for which I can calculate the representation gap decreases when smaller time intervals are chosen. Figure 1.G.7 and Figure 1.G.8 in the appendix show that the results change little if one chooses a 10-year or a 30-year time window around 2007. Moreover, Section 1.D.2 in the appendix shows estimates for Swiss representation gaps on all issues that could be identified from the referendum data, including those that could not be matched to issues from the survey data. Consistent with my other results, voters are consistently more right-wing on cultural and EU-related issues than MPs, while the pattern is more mixed on economic issues. Section 1.D.4 in the appendix shows that these results are similar if voters' actual votes are compared to the vote recommendations of parties.[8]

---

8. Appendix Section 1.D.3 shows how representation gaps varied over time by policy dimension. During the last 90 years, MPs have been constantly more left-wing on cultural topics than voters, and they have been more pro-EU than voters since the first referendums on EU-related topics in the 1990s. On economic topics, MPs have been more state-oriented than voters until the 1980s. Since then, they have been more market-oriented than voters.
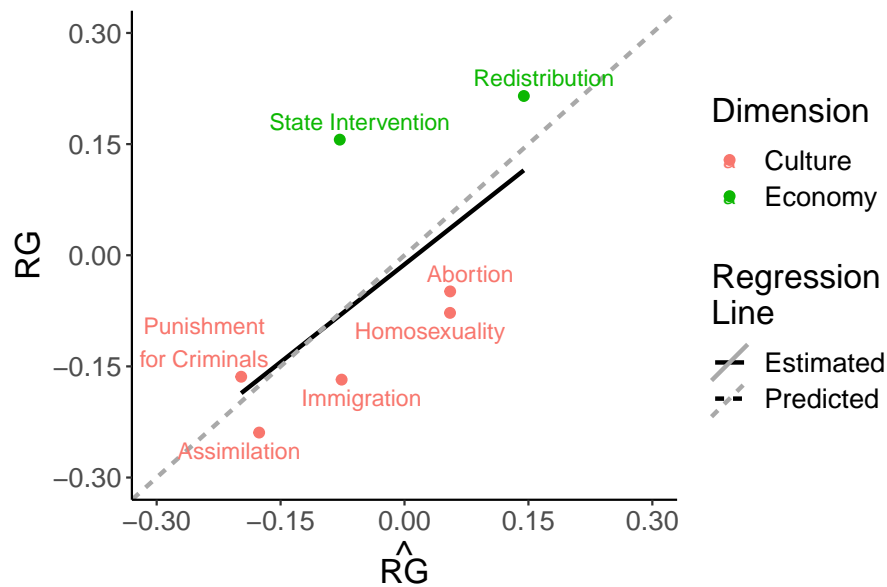
Figure 1.6. Estimates for Representation Gaps Based on Survey and Referendum Data (1997–2017)

Note: The horizontal axis shows RGs calculated from 2007 Swiss survey data. The vertical axis depicts RGs calculated from referendum votes of MPs and ordinary citizens. I take the unweighted means of RGs in all referendums belonging to a topic between 1997 and 2017. The figure also shows a dashed 45° line and a fitted line from a binary regression.

A potential problem of using referendum voting as a measure for representation gaps is that the idea behind referendums is to let voters decide. Hence, MPs might vote based on their personal policy attitudes in referendums but follow voters' attitudes in other decisions. That would imply that the estimates provided by this section are estimates for the attitude differences, not for the representation gaps between voters and MPs. Under this interpretation, the results should be interpreted as a revealed preference approach to the policy attitudes of MPs. It mitigates biases specific to surveys like lying or politically correct responses and, therefore, still illustrates the robustness of representation gaps. However, it would not include the representation intention.

A measure that mitigates this concern is initiation behavior. Referendums can be initiated by different actors. In some cases, the government or the parliament can call for a referendum. Examples include changes in the constitutions or accession to supranational organizations in Switzerland, in which case a referendum is obligatory. In other cases, referendums are initiated by ordinary citizens. For instance, referendums have to be held when an initiative for a referendum has collected a certain number of signatures. There are also mix-versions. For example, citizens may call for a referendum after the parliament makes a decision with which they disagree. Similarly, the parliament may offer counter-proposals to referendum ini-

tiatives put forward by the people. If an actor initiates a referendum on an initiative that would push policymaking to the right, this is evidence of a right-wing move of this actor. Moreover, deciding on which initiatives to hold a referendum on is not purely left to voters. Hence, MPs are more likely to incorporate the representation intention when deciding whether to propose left-wing or right-wing initiatives.

Section 1.D.5 in the appendix shows that "ordinary" citizens and MPs initiate different types of referendums. Citizens mostly initiate referendums that enable a cultural shift to the right, like stricter immigration measures, while most proposals put forward by MPs would push the country culturally to the left. This difference is highly significant. In contrast, there is no significant difference in initiation behavior on economic issues. Hence, initiation differences are similar to attitude and voting differences, suggesting that survey-based estimates for representation gaps are reliable.

1.6 Potential Causes of Representation Gaps

In standard models, electoral competition forces candidates to adopt positions close to the electoral center (Downs, 1957; Persson and Tabellini, 2002; Schofield, 2007). Why doesn't that happen? And why do voters vote in such a way that representation gaps emerge?

There are many potential reasons why electoral competition does not force candidates to implement the policy attitudes of voters. Section 1.F shows that many of these factors are unlikely to explain the existence of the representation gaps documented by this paper. These factors include that voters or MPs find some policy dimensions unimportant (Section 1.F.1), the intention of parliamentarians to protect minority groups from a "tyranny of the majority" (Section 1.F.2), lobbyism (Section 1.F.3), a causal effect of being elected as a parliamentarian on policy attitudes (Section 1.F.4), and the intention of parliamentarians to represent citizens at large or only those who are very interested in politics (Section 1.F.4). The main text focuses on the explanation that seems most likely due to a plausible theoretical argument and support from an empirical analysis.

1.6.1 A Model That Can Explain Representation Gaps

The stylized model presented in this section abstracts from many features of electoral competition in Europe to focus on a particular mechanism. Consider as a reference point a model resembling Downs (1957). There is a continuum of voters. Each voter i has a policy attitude a_i. Policy attitudes are distributed on a single continuous policy dimension according to CDF $A(\cdot)$. Two candidates, L, and R, announce policy positions x_L and x_R before an election takes place. Each voter casts his vote either for L or for R. The candidate who gets more votes is elected. If they get the same number of votes, the election is decided by the toss of a fair coin. Finally, the elected

candidate implements the policy he has announced before the election. This final policy is denoted by x. The utility of voter i is given by $w_i(x) = -(a_i - x)^2$. The utility of the candidates equals one if they are elected and zero if they are not elected. This implies that they are not policy-motivated, and hence their RI equals one. By the median voter theorem (Black, 1948; Downs, 1957) this model has a unique equilibrium in dominant strategies in which both candidates set $x_L = x_R = A^{-1}(0.5)$; they both announce the attitude of the median voter. Hence, the baseline model without valence heterogeneity and a representation intention of one does not feature a representation gap. Intuitively, if one candidate announces a policy position that does not equal $A^{-1}(0.5)$, then the other candidate wins for sure by announcing $A^{-1}(0.5)$ because the half of the electorate that is on the opposite side from $A^{-1}(0.5)$ as the other candidate and the median voter will vote for the candidate at the median position.

The evidence discussed in Section 1.5.1 suggests that candidates are primarily motivated not by winning but by implementing their own policy attitudes.[9] The only thing that changes relative to the baseline model is the utility functions of the candidates. Now $u_L = -(a_L - x)^2$ where a_L is the policy attitude of candidate L. The utility for R is defined similarly. Assume $a_L < A^{-1}(0.5) < a_R$. Again, there exists a unique Nash Equilibrium in dominant strategies in which both candidates announce the attitude of the median voter $A^{-1}(0.5)$.

To see why consider the case in which a_L is closer to $A^{-1}(0.5)$ than a_R is to $A^{-1}(0.5)$. If both candidates announce their preferred policy attitude as their policy positions, L wins, and $x = a_L$ is chosen as the final policy. But this strategy profile cannot be an equilibrium because R can increase his utility by announcing a policy position that is larger than $A^{-1}(0.5)$ and closer to $A^{-1}(0.5)$ than a_L. But if R announces such a policy position, L can increase his utility in the same manner by announcing a position that is even closer to $A^{-1}(0.5)$. In this manner, both candidates can be thought of as outbidding each other until they both announce $A^{-1}(0.5)$ as their policy positions which is implemented with certainty.[10] Hence, policy-motivated candidates alone are not sufficient to generate a representation gap.

Lets modify the model in a third and final step by adding heterogeneity in valence. Valence can loosely be translated as perceived competence of a political actor. Such competence might matter for voting decisions above and beyond ideological distance (Stokes, 1992). Let the valence of candidate L be denoted by λ_L and the valence of candidate R by $\lambda_R < \lambda_L$. I follow Schofield (2007) by capturing valence directly in the utility function of the voters. Now the utility of a voter depends not only on x but also on which candidate is elected. Formally, i's utility if L is elected is given by $w_{i,L}(x) = \lambda_L - (a_i - x)^2$ and his utility if R is elected equals

---

9. Models in which candidates are solely policy-motivated, i.e., RI=0, have been introduced by Wittman (1977) and Calvert (1985).

10. A formal proof can be found in Persson and Tabellini (2002).

$w_{i,R}(x) = \lambda_R - (a_i - x)^2$. The addition of heterogeneous valence changes the results on representation qualitatively. Section 1.E proves that in any equilibrium of this game the final policy is biased away from the median voter toward L's bliss point.

Intuitively, the best that R can do to maximize his vote share is to choose the position of the median voter. But because the policy space is continuous and $\lambda_L > \lambda_R$ there are $a_L \leq x_L < A^{-1}(0.5)$ such that the median voter and all voters with $a_i < A^{-1}(0.5)$ vote for L who therefore wins the election. So L can ensure the final policy is biased away from the median voter toward his bliss point independent of what R does. Moreover, L has an incentive to do so because he is policy-motivated. This simple model illustrates that heterogeneous valence and low RI are both needed to lead to an RG. Low representation intention makes candidates willing to bias policy positions away from the electoral center. Heterogeneous valence allows the higher valence party to get elected despite competition from candidates who are closer to the electoral center.

In the model, policy outcomes are always biased toward the bliss point of the party that has the higher valence. Hence, the systematic representation gaps documented by the paper at hand can only be explained by systematic valence differences between parties. Viewed differently, the model predicts that left-wing parties are seen as more competent than right-wing parties by voters on country-dimension pairs where policy making is more left-wing than voters prefer and vice versa. Due to the complex country-dimension structure of representation gaps shown in Figure 1.4, valence differences between left-wing and right-wing parties are predicted to follow a very specific pattern. For instance, cultural left-wing parties are predicted to have a valence advantage compared to cultural right-wing parties in most countries while economic left-wing parties are predicted to have a valence advantage compared to economic right-wing parties in about half of the countries.

1.6.2 Testing the Valence Hypothesis

It is possible to test these prediction using the survey data. Voters were asked to name the problem they considered most important for their country. Afterward, they were asked which party in their country they considered "best at dealing" with that problem. I calculate the valence for each party based on this variable. Let $\mathbb{1}[\text{p is best}]_{i,p}$ equal one if respondent i states that party p is best at dealing with the most important issue and zero otherwise. Let R_c be the set of those in country c who responded to the item. Let P_c be the set of parties in the country. I define the valence of party p as

$$valence_p^c = \frac{\sum_{i \in R_c} \mathbb{1}[\text{p is best}]_{i,p}}{\sum_{p \in P_c} \sum_{i \in R_c} \mathbb{1}[\text{p is best}]_{i,p}}.$$

In words, $valence_p^c$ is the share of people in the country of party p that believe party p is most competent.

Let c be a country and d a dimension $\in$ {Culture, Economy, EU}. Let P_c be the set of parties in country c. Let $PL_{c,d}$ be the set of parties in country c that are more left-wing than the mean voter on dimension d. Similarly, Let $PR_{c,d}$ be the set of parties in country c that are more right-wing than the mean voter on dimension d. I define the valence advantage of right-wing parties in country c on dimension d as

$$VR_{c,d} = \left| \sum_{p \in PR_{c,d}} valence^c_{p,d} \right| - \left| \sum_{p \in PL_{c,d}} valence^c_{p,d} \right|. \quad (1.3)$$

Hence, I calculate the valence share in country c that all right-wing parties have combined and subtract the shares of valence that all left-wing parties have combined. $VR_{c,d}$ might take on different values for different dimensions, although there is only one valence measure because parties might be culturally right-wing but economically left-wing. The measure is distributed between -1 and 1, and higher values indicate that right-wing parties are seen as more competent compared to left-wing parties.

Figure 1.7 compares representation gaps with $VR_{c,d}$ by country and policy dimension. The horizontal axis shows country abbreviations. The vertical axis relates to the valence measure and the Attitude representation gap. Consistent with the model,
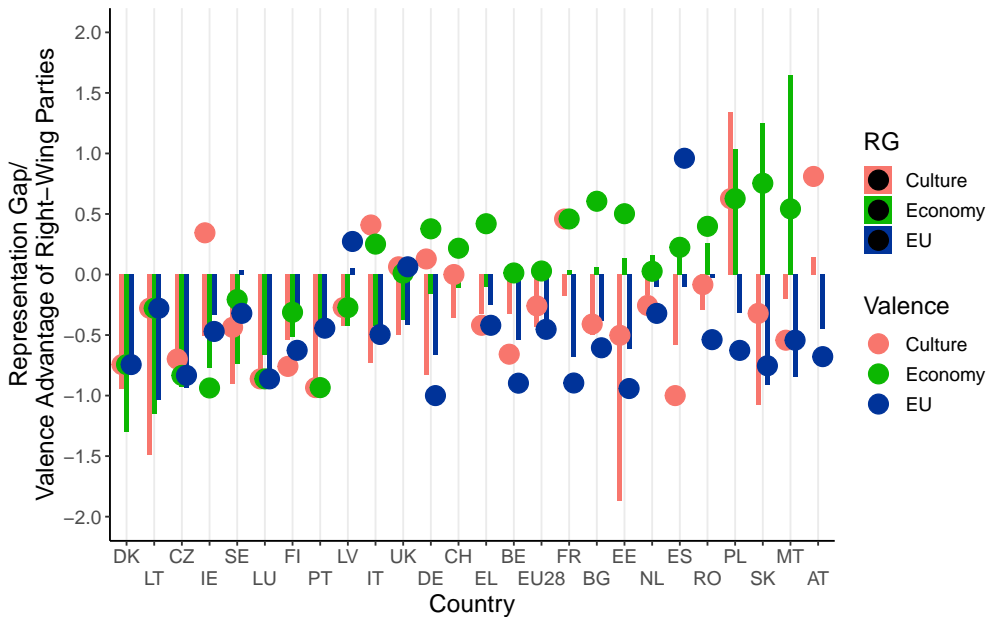
Figure 1.7. Comparing the Valence Measure with Representation Gaps
Note: The horizontal axis shows country abbreviations. The vertical axis shows two variables: the representation gap and $VR_{c,d}$ from Equation 1.3. Bars show values for the representation gap. Points show values for $VR_{c,d}$. Colors indicate the dimension for both measures in the same way. Slovenia, Hungary and the economic dimension in Austria are missing due to missing data.

the valence measure is negative for the cultural dimension 23 out of 27 countries.

Thus, parties that are culturally left-wing relative to voters combine more valence than culturally right-wing parties. Similarly, all but three of the valence measures relating to the EU dimension are negative. This means that in most countries, parties who are in favor of EU integration are seen as more competent as a collective than parties who are rather opposed to it. As predicted, the pattern is more even regarding the economic dimension. The measure is positive in 17 and negative in eight countries. Moreover, parties that are economically more right-wing than voters tend to unite more valence than those on the left of voters.

To examine whether the valence advantages of right-wing parties correlate with the RG between countries *within* a dimension, one can compare the points to the bars in Figure 1.7. Points and bars seem to be positively correlated. This can be seen most easily in the economic dimension, where points track bars closely. In particular, when the valence measure is negative, the RG is also mostly negative, while the RG is usually positive when the valence measure is positive. To assess this relationship formally, I estimate the following equation by OLS:

$$RG_{c,d} = \alpha + \beta \cdot VR_{c,d} + \theta \cdot \mathbf{X}_{c,d} + \varepsilon_{c,d}.$$

$\mathbf{X}_{c,d}$ is a vector of control variables. I divide all variables by their standard deviation to make results easier to interpret. Standard errors are clustered on the country levels. Table 1.H.2 shows the results. The highly significant estimate for β in column (1) shows that an increase in the valence advantage of the political right by one standard deviation is associated with policymaking that is about 0.61 standard deviations more to the right of the mean voter. The R^2 from this binary regression is about 0.38, which shows that valence alone can account for much of the variation in representation gaps. The association is robust to the inclusion of several control variables which might influence representation gaps, like a democracy index and a measure for corruption of which data is taken from the World Bank. Moreover, most other variables are not significant and the coefficient on valence is the largest of all coefficients in all specifications. In column (5), I additionally include indicators for the political dimension. This means that β is estimated using only variation within each of the political dimensions. As a result, β changes little and stays highly significant. Overall, these results are consistent with the hypothesis that differences in valence are an important driver of representation gaps.

A key concern with the valence measure is that valence cannot be completely disentangled from proximity in policy space. Subjects might believe that a party is competent because it expresses views that are similar to those of the subject. This would imply that the valence measure mixes the perceived valence of a party and proximity in policy space between voter and party. While I cannot rule out that my measure does mix up these two factors, this bias strengthens my conclusions instead of invalidating them. Assume that valence and position in policy space are identical measures. Given that voters find left-wing parties more competent than right-wing

parties, their attitudes should align with those of left-wing parties. In contrast, I find that most voters find left-wing parties more competent but are at the same time more right-wing than parliaments. Hence, a positive correlation between valence and proximity in policy space cannot explain the pattern described in Table 1.H.2 and would lead to a downward bias of β.

1.7 Potential Consequences of Representation Gaps

1.7.1 Political Trust

Representation gaps might make citizens believe that the parliament does not consider their concerns which could in turn lead to reduced trust in democratic institutions. The previous analysis has also shown that in the cultural and EU dimensions, all established party families are more left-wing than the average voter. Citizens who are more conservative than established parties on these dimensions might struggle to find a party they can vote for because non-established parties tend to have low valence. Such citizens might therefore abstain from voting. Because there exist established parties to the right and the left of voters on the economic dimension, a representation gap there is less likely to cause vote abstention.

To test these predictions, I compare the extent to which citizens, whose attitudes are differently well represented, trust political institutions, are satisfied with them, and are willing to vote. Let

$$|\Delta_{i,d}| = |\text{attitude index}_{i,d} - \overline{\text{EP attitude index}_d}|,$$

where $\overline{\text{EP attitude index}_d}$ is the mean value for the index on policy dimension d of MEPs. $|\Delta_{i,d}|$ measures the absolute distance between voter i on policy dimension d from the mean MEP. Higher values indicate that i's policy attitudes differ more from the attitudes of the mean MEP. I regress several outcome variables of interest y_i on this variable using OLS:

$$y_i = \alpha + \beta_d \cdot |\Delta_{i,d}| + \theta \cdot \mathbf{X}_i + \varepsilon_i. \tag{1.4}$$

$\mathbf{X}_i$ includes the following control variables: age, gender, degree of religiosity, marital status, city size, living standard, social class, occupation, education, immigration background, and religious denomination. To simplify the interpretation, I standardize all dependent variables except the voting indicator and divide all $|\Delta_d|$ by the standard deviation of citizen attitudes regarding the corresponding dimension index. Consequently, β_d, which measures by how many standard deviations y_i differs between a citizen and another citizen with similar demographics but policy attitudes that are a standard deviation of attitudes further away from those of the mean MEP. Standard errors are clustered at the country level.

In principle, I am interested in the causal effect of $|\Delta_{i,d}|$ on y_i. However, Equation 1.4 estimates a correlation. Including $\mathbf{X}_i$ can help mitigate omitted variable bias but cannot account for reversed causality. Reversed causality is likely to be most strongly pronounced in the EU dimension. Citizens might be more opposed to EU integration than the EP because they believe that the EP doesn't represent citizens rather than the other way around. Reversed causality is likely to bias estimates for β_d upward, while measurement error likely leads to a bias in the opposite direction. Hence, the causal effect might be larger or smaller than β_d.

Figure 1.8 shows the β_d coefficients and 95% confidence intervals by democracy attitude variable y and policy dimension d. Results are consistent with the expectations. Most coefficients are positive and statistically significant. Voters with an additional bias from the mean MEP on EU-related issues of one standard deviation are about 0.41 standard deviations more likely to believe that the European Parliament does not consider the concern of the citizens. Associations are the largest in the EU dimension and weakest in the economic dimension.
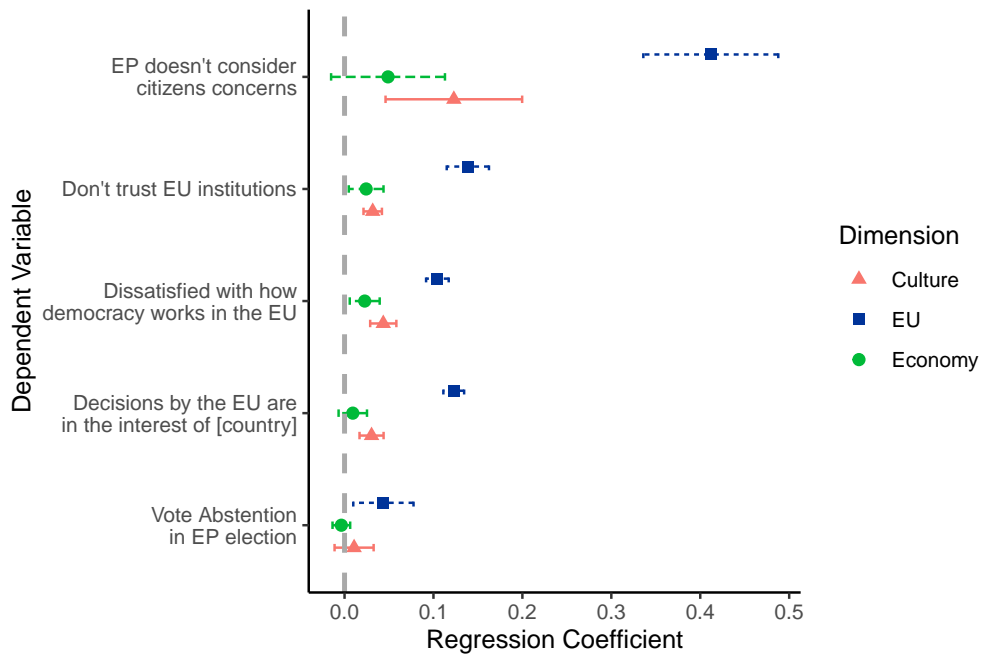
Figure 1.8. Association between Representation and Political Trust

Note: This figure shows a coefficient plot. It depicts estimates for β_d from different versions of regression Equation 1.4. The vertical axis shows different dependent variables. Colors indicate the policy dimension d on which the difference between voters and the mean MEP is calculated. the horizontal axis shows the magnitude of coefficients. All regressions condition on a large set of demographic characteristics. Data is pooled for all EU countries. Standard errors are clustered at the country level.

The strongest associations emerge regarding whether the European Parliament considers the opinions of its citizens. Those with different policy attitudes than the mean MEP are much more likely to think that it doesn't consider the concerns of

its citizens. Moreover, citizens with different attitudes than the mean MEP are less likely to trust EU institutions, more likely to be dissatisfied with how democracy works in the EU, and to believe that decisions made by the EU are not in the interest of their country.

The association with voting in the 2009 European Parliament election is quantitatively smaller because it is measured on a different scale. The dependent variable is an indicator of whether the citizen voted or not. Citizens with the same opinion as the mean MEP regarding the EU dimension were 15 percentage points more likely to vote than citizens with completely different attitudes toward the EU than the mean MEP. The association regarding the cultural dimension is also estimated to be positive but weaker and insignificant. The association regarding the economic dimension is insignificant and estimated to be negative.[11]

Figure 1.G.11 in the appendix shows that similar results obtain at the national level. Citizens who have more distant policy attitudes from their national MPs are more likely to believe that the national parliament does not consider the concerns of its citizens, are dissatisfied with how democracy works in their country, tend to disapprove of the country's government, and are less likely to vote at national elections. Overall this evidence is consistent with the idea that representation gaps decrease trust in democracy and the EU as a whole and that the lack of established parties that are right-wing on cultural and EU-related topics reduces voting.

1.7.2 Populism as a Reaction to Representation Gaps

The inclusion of valence and politically motivated candidates in the Downs (1957) framework explains why neither existing parties nor voters deviate from an equilibrium that has representation gaps. But why do new parties not fill the representation gaps and consequently rise? Indeed, in recent decades, a new group of challenger parties, the populists, have achieved remarkable electoral success. Could the rise of populism be a consequence of representation gaps?

1.7.2.1 Defining Populism

The rise of populism refers to the increase in the vote share and power of populist parties. However, whether a party is populist is often hard to define because "populist" is usually not used by parties to describe themselves but rather ascribed to them by observers. Nonetheless, a consensus on the key properties of populists has emerged. Populism is a thin ideology. It should be distinguished from fully fleshed out ideologies like liberalism or conservatism, which include a consistent worldview and derive policy objectives from it. Rather, the most important feature of populists

---

11. Bakker, Jolly, and Polk (2020) find, using a different methodology, whether ideological incongruence between parties and voters is associated with lower political trust and increased vote-shares for anti-establishment parties.

is to draw a clear distinction between the "corrupt elite" and the "pure and homo-geneous people". This distinction is linked to a narrative that populists base their legitimacy on and that I will label the populist narrative (Mudde and Kaltwasser, 2017; Guriev and Papaioannou, 2022).

According to the populist narrative the people and the elite are caught in a struggle. The elite governs in a way that is not congruent with the policy preferences of the people. Seeing this, populists enter politics to replace the elite and better represent the policy preferences of the people (Mudde and Kaltwasser, 2017; Guriev and Papaioannou, 2022).

Neither the distinction between the people and the elite nor the populist nar-rative make any claims on political orientation. Hence, populism is consistent with right-wing or left-wing policy stances. Because most scholars agree that the elite-people distinction and anti-elite sentiment are the most basic feature of populism, most past attempts to define populism have been based on it (Guriev and Pa-paioannou, 2022). I follow this literature. Data for classification is taken from the CHES. In one CHES item, political experts are asked to assess the salience of anti-establishment and anti-elite rhetoric on a scale from 0 = "Not important at all" to 10 = "Extremely important." This item was included in the 2014 and the 2019 CHES, which is roughly the time period I focus on and makes it possible to examine tem-poral stability.

I label all parties as populists who score more than one standard deviation above the mean on this item. This procedure results in a set of 50 parties which are classi-fied as populist in 2014. While one could use other thresholds, this approach results in a set of populist parties that is large and plausible. Populist parties are listed in Table 1.H.11. Parties that are typically described as populist, like the National Rally in France, Lega Nord in Italy, or SYRIZA in Greece, are all found to be populist using this approach. The most notable exceptions are the PiS party in Poland and the SVP in Switzerland, which are not classified as populist. The dataset also includes many minor parties to which researchers have paid less attention. Many of them already have the key property of populism in their official party name, like the "ANO 2011, Action of Dissatisfied Citizens" in the Czech Republic or the "Ordinary People and Independent Personalities" in Slovakia.

The correlation of the anti-elite salience item in 2014 with the anti-elite salience item in 2019 is about 0.84, which shows that the populism measure is robust over time. An alternative item for identifying populists is the "people vs. elite" item, which measures whether "the people" or elected representatives should have the final say on the most important issues. However, it was only included in the 2019 CHES. Reassuringly, the correlation between the two measures is about 0.78.

1.7.2.2 Testing the Populist Narrative

The populist narrative is simplistic and, taken literally, wrong. The "people" and the "elite" are not well defined, and no group is completely homogeneous. However, to take the narrative literally misses the point that it does not claim to be scientific. Rather it is used by politicians to get attention and votes. Therefore, one should expect the populist narrative to be an exaggeration. Let's interpret it more leniently. "Elite" and "the people" are vague terms. Populists usually refer to politicians when talking about the elite, and most populist movements are political. Hence, it makes sense to use national and EU parliament members as a substitute for the elite. "The people" likely refers to citizens of the country who do not belong to the elite. Sometimes it seems like populists exclude ethnic minorities from the people. However, this would not alter the considerations presented below, as ethnic minorities are small in most European countries.

The populist narrative makes three main claims. First, if parliamentarians and citizens are caught in a struggle for the country's political direction, a representation gap must exist. I will label this the "representation gap claim." Second, it makes a "homogeneity claim" when it argues that parliamentarians and citizens are homogeneous groups. The most important implication of the homogeneity claim and the representation gap claim is that all parliamentarians have attitudes that differ from those of citizens in the same direction. Hence, closing representation gaps requires a rise of a counter-elite representing citizens. Finally, populists make a "representation claim" when they argue to be this counter-elite. Hence, representation gaps are central to the populist narrative and to their justification for entering politics. In particular, populists argue themselves to be a reaction to representation gaps (Mudde and Kaltwasser, 2017).

This paper shows that the representation gap claim is true. Large representation gaps indeed exist in Europe. The homogeneity claim is wrong, taken literally, but it contains a kernel of truth. There is important within-group heterogeneity in the parties' positions and voters, and there are parliamentarians who are more right-wing than the mean voter in most parliaments. Moreover, there are parties to the right and the left in the economic dimension in most European Parliaments. Hence, the homogeneity claim is an exaggeration. However, it is true that in most European countries, all large established parties are more left-wing than most voters in the same direction in the cultural and the EU dimension, and many populists focus on cultural topics like immigration, policy toward minorities, or punishment of criminals. Moreover, while differences between party positions exist, they are often small compared to the difference between the mean citizen and the mean parliamentarian. This is particularly strongly pronounced in the EU dimension. Hence, the homogeneity claim could be seen as a simplification that allows one to zoom in on a central feature of the political situation in Europe.

Populists have identified a large gap in the policy space. Filling this gap is a reasonable and plausible motive for a politician, independent of his long-term goals. Some people might have a taste for representing the citizens' attitudes at large and aim to fill policy space for idealistic reasons. Similarly, citizens who do not feel represented might become politicians to represent their own attitudes and like-minded people (Bó et al., 2023). On the other hand, an opportunistic politician who wants to acquire power might fill empty policy space to maximize his vote share. But do populists indeed fill representation gaps?

1.7.2.3 Testing the Representation Claim of the Populist Narrative

If populists react to representation gaps, they should be relatively likely to believe in low substantive representation. To test this, I compare the responses of national MPs belonging to populist and non-populist parties to the following item:

> *Thinking about how elections in [country] work in practice, how well do you think they ensure that the views of MPs accurately reflect the views of voters?*

Possible answers included "very well," "fairly well," "not very well," and "not well at all."
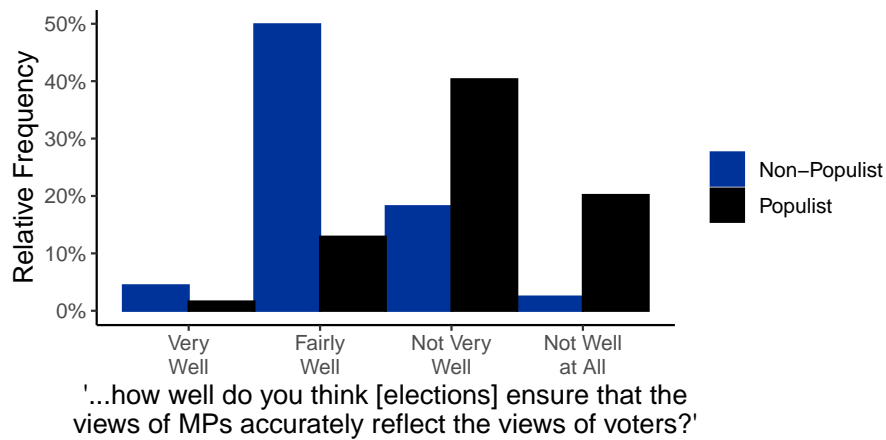
Figure 1.9. Perceptions of Attitude Representation Gaps by Populist and Non-populist MPs
Note: This figure depicts two histograms that visualize answer frequencies among parliamentarians from populist and non-populist parties to the following statement: "Thinking about how elections in [country] work in practice, how well do you think they ensure that the views of MPs accurately reflect the views of voters?" Results are based on national MPs.

Figure 1.9 visualizes the response distributions of populist and non-populist MPs. Populists are much more likely than non-populists to believe that attitude differences are large. A majority of non-populists believe that the views of MPs reflect the views of voters "fairly well," while most populist MPs think that the views of MPs reflect the views of voters "not very well" or "not well at all." Given the large attitude differences found by this study, populist MPs have a more realistic perception of representation

in Europe than non-populist MPs. Overall, this suggests that populists truly perceive attitude differences to be large which is consistent with their narrative. Figure 1.G.9 in the appendix shows similar results for MEPs.

The populist narrative further asserts that populists do represent citizens. To examine this, I first analyze their representation intention. Of the 1,366 national MPs belonging to non-populist parties, 978, or about 72%, said that an MP should vote according to his own opinion. In contrast, out of the 101 populist national MPs, only 31, about 31% stated that an MP should vote according to his own opinion. The difference is highly significant according to Fisher's exact test (p < 0.0001). Of the 114 MEPs I identified as belonging to non-populist parties, 98 or about 86% stated that an MEP should vote according to his own opinion if it differed from the opinion of his voters, and 14% stated that an MEP should vote according to the opinion of his voters. Of the eight MEPs belonging to populist parties, four stated that an MEP should vote according to his own opinion, and four stated that an MEP should vote according to the opinion of his voters. Fisher's exact test reveals that this difference is significant, too (p ≈ 0.024). Hence, parliamentarians from populist parties have a much higher Representation Intention than those from non-populist parties.

Are attitude representation gaps between populist MPs and voters also smaller than between mainstream MPs and voters? Populism is not necessarily correlated with substantive political positions. To examine whether and how populist parties differ from mainstream parties, I estimate the following regression equation for several political issues individually by OLS:

$$Y_{p,t,i} = \alpha + \beta_i \cdot \mathbb{1}[\text{Populist}]_{p,t} + \delta_t + \theta_c + \varepsilon_{p,t,i}. \tag{1.5}$$

$Y_{p,t,c,i}$ is the political position of party p (in country c) at time t on issue i. $\mathbb{1}[\text{Populist}]_{p,t}$ is an indicator for whether the party is populist, which can be time-varying. I also include time- and country-fixed effects to absorb constant differences between countries and parallel issue-position-trends. I pool data from the 2014 and 2019 CHES to estimate this model and use all policy issues included in either of the two surveys while most items are included in both of them. To make items comparable and easier to interpret, I scale them such that higher values indicate a stance that is intuitively more right-wing on the issue and such that all variables have a standard deviation of one.

I then estimate Equation 1.5 for each i individually. As a result, β_i measures descriptively how many standard deviations of party positions the average populist party is more right-wing on issue i than the average mainstream party, controlling for time and country effects.

Figure 1.10 depicts β_i for several issues i together with 95% confidence intervals based on standard errors clustered at the country level. The horizontal axis refers to different political issues. The vertical axis measures the corresponding estimated coefficient. Nearly all estimates are significant and quantitatively large. Populist
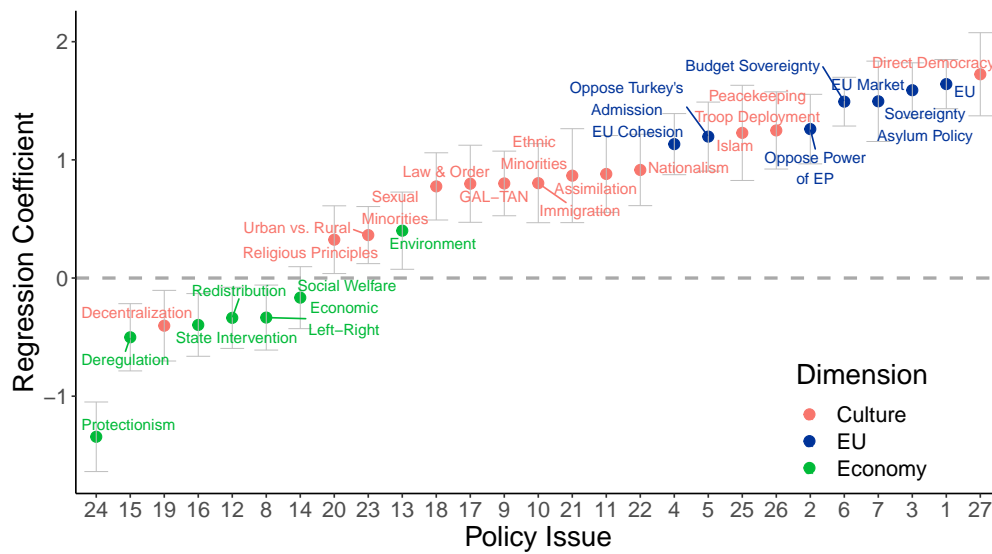
Figure 1.10. Issue-Positioning of Populist Parties Relative to Non-populist Parties
Note: This figure is based on regression Equation 1.5. Each dot shows the coefficient on the populist indicator for a different dependent variable. The text next to each variable indicates the dependent variable used. The vertical axis shows the size of the coefficient. Colors indicate policy dimensions. Coefficients are surrounded by 95% confidence intervals. Standard error are clustered at the country level.

parties are significantly more right-wing than non-populist parties regarding nearly all cultural and EU-related issues. In contrast, populists are more left-wing on all but one economic issue, while differences tend to be smaller than for cultural and EU issues. Figure 1.G.10 in the appendix shows that results change little if parties are weighted with their vote share in the last national election.

This pattern resembles the representation gap pattern documented above. Even the magnitudes of differences between 1) populists and established parties and 2) voters and established parties, on the other hand, are similar. To assess this more directly, I handly match topics from both analyses. I am able to do this for eight specific policy issues and the three policy dimensions. Figure 1.11 shows a scatter plot of the representation gap on the horizontal axis and differences between populist and non-populist parties on the vertical axis. Points refer to political issues or dimensions. The further right a point is, the further right the European Parliament was relative to Europeans on this issue in 2009. The higher a point is, the further to the right populist parties were relative to other parties five to 10 years later.

Voters are more right-wing than MEPs on all cultural issues, and populists are more right-wing than non-populist parties on these issues too. The same is true for the EU dimension. Similarly, voters are more left-wing than MEPs, and populists are also more left-wing on these issues than non-populists. The only exception is the issue of state intervention, where voters are more right-wing than MEPs, but
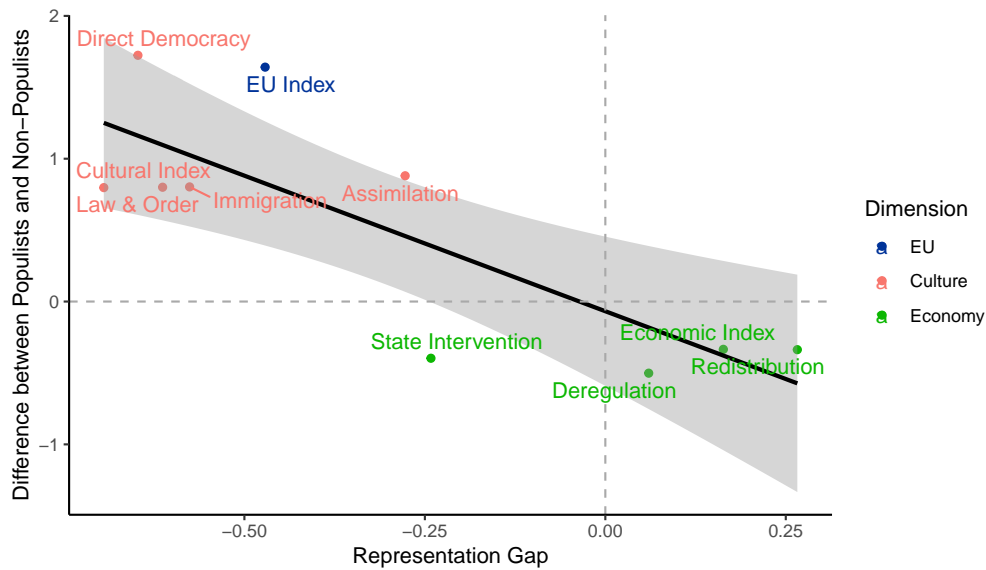
Figure 1.11. Relationship between Representation Gaps and Differences between Populists and Non-Populists
Note: representation gaps contrast those who voted in the 2009 EP election and MEPs who were elected in 2009. Higher values on the horizontal axis indicate that voters were more right-wing compared to MEPs. Differences between populists and non-populists are based on Equation 1.5 and the vertical axis shows the size of β_i. The black line visualizes the slope a linear regression and is surrounded by a 95% confidence band.

populists are more left-wing than non-populists. Still, on 10 of the 11 issues, populist parties differ from non-populist parties as if they reacted to the representation gap.

The correlation between representation gaps and the differences between populist and mainstream parties is about -0.8 and highly significant (p≈0.005). I also estimate a binary regression of the form $\beta_i = \alpha + \theta \cdot RG_i + \varepsilon_i$, using robust standard errors. The intercept turns out to be small and insignificant while θ is estimated to be about -1.9 and highly significant. Hence, higher representation gaps on a topic translate into a larger difference between populists and non-populists. Moreover, the fact that θ is larger than one shows that populists overshoot. Differences between them and mainstream parties are larger than differences between voters and parliaments.

To assess more directly whether populists are close to voters in policy space Figure 1.12 depicts mean attitudes of populist MEPs, non-populist MEPs, and voters using the main survey dataset. The vertical axis shows different policy issues. The horizontal axis shows the positions of voters and MEPs. The attitudes of the mean populists are not much closer to the attitudes of the mean voter than the attitudes of the mean non-populist are to the attitudes of the mean voter. On five topics (redistribution, abortion, same-sex marriage, EU unification, and EU membership), the mean non-populist MEPs are closer to the mean voter than the mean populist MEPs.

On six issues (state ownership, assimilation, teaching authority in schools, gender relations, immigration, and punishment for criminals), the mean populist is closer to the mean voter. On three topics (private enterprise, state intervention, and direct democracy), both types of parliamentarians are similarly close to the mean voter.
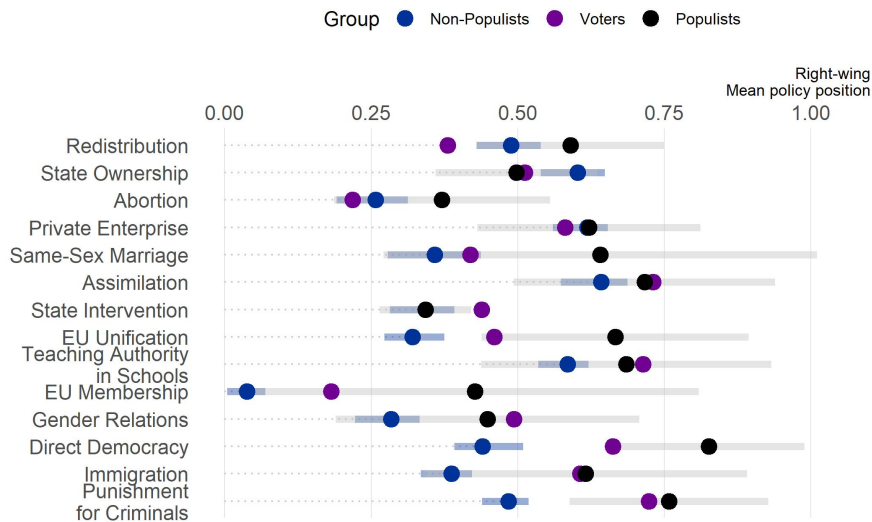
Figure 1.12. Mean Attitudes of Voters, Populist MEPs and Non-Populist MEPs by Policy Issue
Note: The vertical axis shows different policy issues. All variables are scaled to range from zero to one and such that higher values indicate a position that is more right-wing. Dots indicate the mean-attitudes of voters, MEPs who belong to populist parties and MEPs who do not belong to populist parties. Data is pooled across Europe. 95% confidence intervals are based on standard errors clustered at the country level.

On many issues, all parties differ from voters in the same direction (Figure 1.G.5). Hence, populists could increase representation even though being positioned far away from the mean voter if they differ from voters in the opposite direction than non-populists. Figure 1.12 shows that on half of the issues, populists differ from voters in the same direction. On the other half, the mean populist is positioned on the other side of the mean voter than the mean non-populist. This evidence also points to populists improving representation on some but not all issues.

Figure 1.13 shows results for the three policy dimensions. Populists' policy preferences are close to those of non-populist MEPs regarding economic variables. In the cultural dimension, the mean populist differs from the mean voters in the other direction than the mean non-populists, but he is much closer to the mean voter than the mean non-populist MEP is. In the EU dimension, attitudes of populists and non-populists differ most strongly. Mean MEPs of both types differ strongly from the mean voter whose attitudes are located between their attitudes.

Overall, evidence regarding the representation claim of populists is mixed. Gaps between MEPs from populist parties and voters are large on many issues. On some
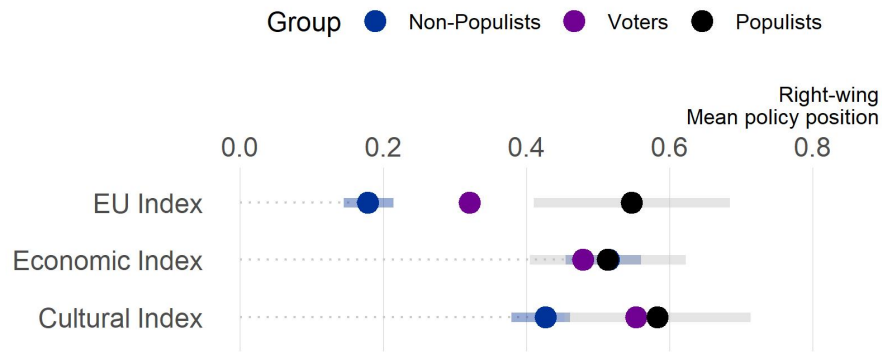
Figure 1.13. Mean Attitudes of Voters, Populist MEPs and Non-Populist MEPs by Policy Dimension
Note: The vertical axis shows different policy dimensions. All variables are scaled to range from zero to one and such that higher values indicate a position that is more right-wing. Dots indicate the mean-attitudes of voters, MEPs who belong to populist parties and MEPs who do not belong to populist parties. Data is pooled across Europe.

issues, populist MEPs are even further away from voters than mainstream MEPs, sometimes due to populists overshooting by choosing positions that are extremely far away from mainstream parties. In the EU dimension, they can help to make parliaments more representative of voter attitudes, but the mean-populist is much more opposed to EU integration than the mean voter. Economically the evidence on the position of populists is conflicting. Results based on the CHES suggests that they are a bit more left-wing than mainstream MPs, while populist MEPs have similar attitudes as mainstream MEPs. In any case, populists might not be needed to increase representation in the economic dimension, given representation gaps are relatively small, and there exist high valence parties to the right and the left of the mean voter. The representation claim of populists is strongest in the cultural dimension. The mean populist MEP is culturally close to the mean voter. This is also the case for individual issues that are important to many voters, like immigration, punishment for criminals, assimilation, and gender relations.

Moreover, given that the representation intention of populists is high, only about a third of attitude representation translates into representation gaps. This could explain the conflicting evidence on the economic position of populists. They might advocate for left-wing economic policies even though they do not believe in them.

Hence, the representation claim contains less truth than the other two claims of the populist narrative, as populists have attitudes far away from those of voters on many issues, and even 31% of populist parliamentarians seem to disregard the opinion of their voters. Still, it contains a kernel of truth. Combining evidence on representation intention and attitudes shows that populists do represent voters well

on cultural issues where representation gaps have been largest, and they provide an alternative to the uniform pro-EU stance of mainstream parties, which many voters consider to be too integration friendly.

Considering the evidence on all claims of the populist narrative in combination shows that much of it is true, even though populists often exaggerate and oversimplify it. This is consistent with the idea that the main implication of the populist narrative, their claim to be a reaction to representation gaps, is true too.

1.8 Conclusion

Representative democracies build on the idea that voters elect parliamentarians, which in turn, represent the voters' interests. In this paper, I explore the degree to which parliamentarians in Europe actually represent their voters. I document large representation gaps. For example, on an index of cultural issues, the mean parliamentarian is between half of a standard deviation and one standard deviation more left-wing than his voters in most countries. Moreover, I explore the structure of representation and find that representation gaps are largest on cultural issues, like immigration or gender relations. Their existence can largely be explained by an advantage in perceived competence that parties on the cultural left enjoy.

The question arises whether representation gaps are problematic or not. My findings are consistent with representation gaps causing vote abstention, distrust in democratic institutions, and the rise of populism, which many scientists consider problems (Otten et al., 2017; Funke, Schularick, and Trebesch, 2021; Docquier, Peluso, and Morelli, 2022; Guriev and Papaioannou, 2022). Still, the evidence presented here is far from conclusive, and, to the best of my knowledge, there has been very little research on this question. Hence, examining under what conditions representation gaps are desirable or undesirable from a welfare perspective is an exciting avenue for further research.

Assuming that policymakers try to reduce representation gaps, they could do so in two ways. First, they could change voters' policy attitudes, for instance, through information campaigns. Second, they could change policymaking directly to align with the voters' attitudes. Which alternative is preferable strongly depends on whether representation gaps result from information asymmetries or value differences between voters and parliamentarians. For instance, if a representation gap results because the public misses crucial information on a topic that politicians have, information interventions seem to be appropriate. In contrast, politicians might have different deep values or preferences, like their religious beliefs or structural parameters like risk aversion or patience. In this circumstance, a representation gap is more likely to signify a need for reform.

Interestingly, recent studies have shown that European voters hold large misperceptions about immigration, where the representation gap is particularly large (Bar-

rera et al., 2020; Grigorieff, Roth, and Ubfal, 2020; Alesina, Miano, and Stantcheva, 2023). This is consistent with representation gaps arising due to a lack of information on the part of voters. However, Kustov, Laaker, and Reller (2021) note that immigration attitudes are very stable over time and robust to major shocks, making it unlikely that they are easily susceptible to information. Consistent with this observation, many experimental studies find that providing subjects with information about immigrants does not strongly affect their immigration attitudes (Hopkins, Sides, and Citrin, 2019; Barrera et al., 2020; Alesina, Miano, and Stantcheva, 2023). Moreover, Chapter two of this dissertation (Günther, 2023) shows that most Europeans would be opposed to immigration by asylum seeker if they were informed about their true characteristics. This casts doubt on the idea that anti-immigration attitudes are mainly due to misperceptions and, thereby, the hypothesis that asymmetric information is responsible for the immigration representation gap. At the same time, Heß et al. (2018) show that German parliamentarians are more risk-loving than the average German citizen across several domains of risk-taking. This suggests that risky decisions like opening the country to large-scale immigration might result from differences in risk attitudes between voters and politicians.

Still, assessing this question directly requires data about the preferences and information of parliamentarians and voters, which, to my best knowledge, does not exist yet. Moreover, other representation gaps could be explained by information asymmetries, even if the immigration representation gap is the result of value differences. Therefore, examining whether representation gaps are due to information asymmetries or value differences seems to be a another avenue for future research.

Finally, representation gaps can help to explain the characteristics of populist parties. One example of this is the combination of right-wing cultural and left-wing economic policy positions they supply. Comparing the findings presented here to research in political science and psychology suggests that their anti-media and anti-expert stance are related to representation gaps too. The representation gaps documented here closely resemble the media biases identified by Puglisi and Snyder (2015) while Haidt and Lukianoff (2018) shows that experts tend to be more socially left-wing than the population. Thus, examining the relationship between political representation gaps, media bias, and a potential "expert bias" might be another promising starting point for further research.

Appendix 1.A   Relationship between the Estimation Framework and Previous Approaches

The two most prominent concepts of substantive representation used by political scientists are responsiveness and congruence. Policy responsiveness refers to whether the attitudes of voters and MPs change in the same direction over time. Policy responsiveness is perfect if the attitudes of voters and MPs are extremely different at any point in time but move parallel over time (Beyer and Hänni, 2018). In contrast, my measure compares policymaking to the attitudes of voters at a given point in time. Moreover, it can be calculated for consecutive points in time to make trends and levels of policymaking and voter attitudes comparable.

The literature on congruence has measured how similar the policy attitudes of politicians and voters are. In the literature, several terms are used to refer to this type of congruence: substantial (or substantive) representation, ideological congruence (Mayne and Hakhverdian, 2017), opinion congruence (Walgrave and Lefevere, 2013) or issue-congruence. To the best of my knowledge, this literature has compared the attitudes of all MPs, and often those of all MP candidates, with those of voters without considering the role of policy-motivation. This might lead to biased estimates because it ignores that MPs do not need to have the same attitudes as voters to act in their interests. For instance, in standard models of electoral competition, voters can punish MPs who do not implement their attitudes by not re-electing them, thereby incentivizing MPs to implement the voter's attitudes even if this goes against their own attitudes (Pitkin, 1967; Sappington, 1991; Przeworski, Stokes, and Manin, 1999). Consequently, I integrate policy-motivation into my framework. Moreover, I validate my estimation framework with real-world decision data.

Appendix 1.B   Are Representation Gaps are Problem?

Whether a representation gap is a problem or not likely depends on contextual factors. Hence, some RGs might be a problem, while others are not. Factors that might be important include how large and systematic RGs are. Small RGs are unlikely to have large negative welfare effects. Similarly, if there exist RGs in opposite directions on similar topics like state intervention and state ownership, RGs might effectively cancel out. On the other hand, large and systematic RGs, like those documented above for cultural issues, are more likely to be a problem.

Another factor is the degree to which some policy attitudes are objectively superior to others. For instance, if the attitudes of voters only differ from those of parliamentarians because voters are misinformed, RGs are less likely to be a problem. RGs are more likely to be a problem if they are due to differences between politicians and the population regarding deep preferences like fairness ideals (Cappelen, Falch, and Tungodden, 2020), risk aversion (Heß et al., 2018) or moral universalism

(Enke, 2020; Enke, Rodríguez-Padilla, and Zimmermann, 2022). Recent research has indeed found that the perceptions of most Europeans about immigrants are too negative (Grigorieff, Roth, and Ubfal, 2020; Facchini, Margalit, and Nakata, 2022; Alesina, Miano, and Stantcheva, 2023). Similarly, research in behavioral economics has uncovered numerous cognitive biases and shown that most people make predictable mistakes that might justify disregarding their opinions under some circumstances. However, the decision-making of politicians suffers from biases too. Sheffer et al. (2018) show that politicians in Belgium, Canada, and Israel exhibit several cognitive biases to an if anything, *larger* extent than non-politicians. This suggests that letting politicians instead of voters decide makes biased decision-making more likely. Moreover, there is evidence that deep preferences of politicians and voters differ. For example, Heß et al. (2018) show that German MPs are significantly more risk-loving than the average German citizen across several domains of risk-taking. This suggests that risky decisions like opening the country to large-scale immigration might also result from differences in risk attitudes.

Even if a representation gap results from differences in deep preferences between voters and MPs, it might not be inherently problematic. One justification could be that MPs try to protect minority rights from a "tyranny of the majority." However, Section 1.F.2 shows that this is not the case in modern Europe. Yet another justification for RGs could be that politicians act as opinion leaders or anticipate long-run trends in attitude change and make political decisions that will be congruent with the attitudes of future generations. For instance, European have become increasingly culturally liberal during the last decades (Inglehart, 1971, 2015). As a result, many policies that would have been more culturally conservative than desired by voters several decades ago are now supported by most voters. The present paper cannot rule out this possibility. Time will tell whether voters will follow politicians to close representation gaps in the long run.

Appendix 1.C Data Quality

In this section, I provide two quality checks of the EES Candidate Survey data. First, I compare the sample MEPs to the universe of MEPs (all MEPs who served between 2009 and 2014) regarding several demographic variables. Data on these variables are taken from Beauvallet, Lepaux, and Michon (2013).

In 2009 66% of all newly elected MEPs were male, and 34% were female. The proportions in the sample are nearly identical. Figure 1.C.1 visualizes the representativeness of the sample MEPs concerning the highest educational attainment. As can be seen, most MEPs are highly educated. More than 20% of all 2009 MEPs had a Doctoral degree, and only about 12% had no University-degree. This strong selection can be replicated well by the sample used in my analysis. In Figure 1.C.2, I compare the distributions of occupations previously held by the sample MEPs to the

distribution of occupations held to be the universe of MEPs prior to becoming MEPs. Most noteworthy, the vast majority of MEPs have worked in two out of the 12 occupation categories prior to becoming MEP: "higher administrative jobs," which include senior executive or political aide, and "professional and technical jobs," which incorporate scientists, journalists, and teachers. Very few MEPs have worked as manual workers, clerics, farmers, or in the sales sector. Figure 1.C.2 reveals that the sample distribution of the previous occupation is quite close to the actual one. Figure 1.C.3 shows the country of origin fractions of sample MEPs and the universe of MEPs. The sample is broadly representative of the universe. However, MEPs from some countries like Germany, Luxembourg, and Belgium are over-represented, while representatives from Spain and Poland are underrepresented in the data. I weigh to adjust for this in the main analysis. Overall, the sample represents the universe of MEPs well along key demographic dimensions.
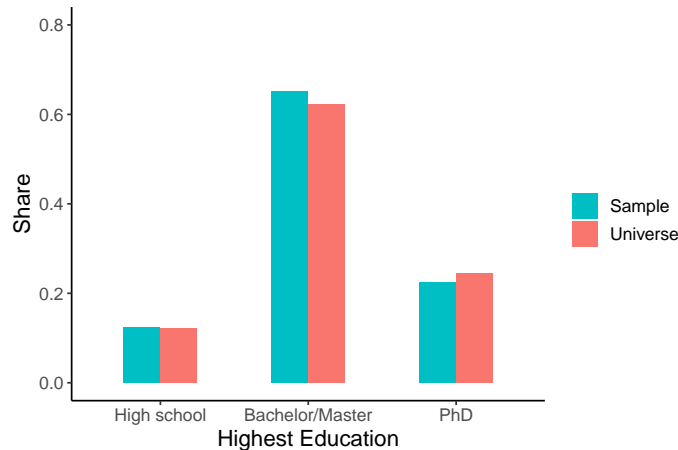
Figure 1.C.1. MEP-Sample Representativeness Regarding Education
Note: This figure compares the distribution on educational attainment of the universe of MEPs to the distribution on educational attainment of the sample of elected candidates from the Parliament Election Study 2009 (Candidate Study).

Second, I check how the EES Candidate Survey data correlates with an established and validated data source. I calculate the ideological positions of parties for various issues based on the EES Candidate Survey data by taking for each party and issue means with equal weights of the positions of all of its candidates. One established data source for the ideological positioning of parties is the Chapel Hill Expert Survey (CHES). The CHES does not enable ideological comparison between parties and voters because it does not include data on voters' ideological positions. However, it is well suited for comparing different parties (Jolly et al., 2022).

Moreover, it includes data on the ideological positioning on issues that are similar to those from the EES. Both data sources include a "redistribution" issue which I match. Moreover, I match the "position on immigration policy" issue from the CHES to the "immigration" issue from the EES, the "position on integration of immigrants
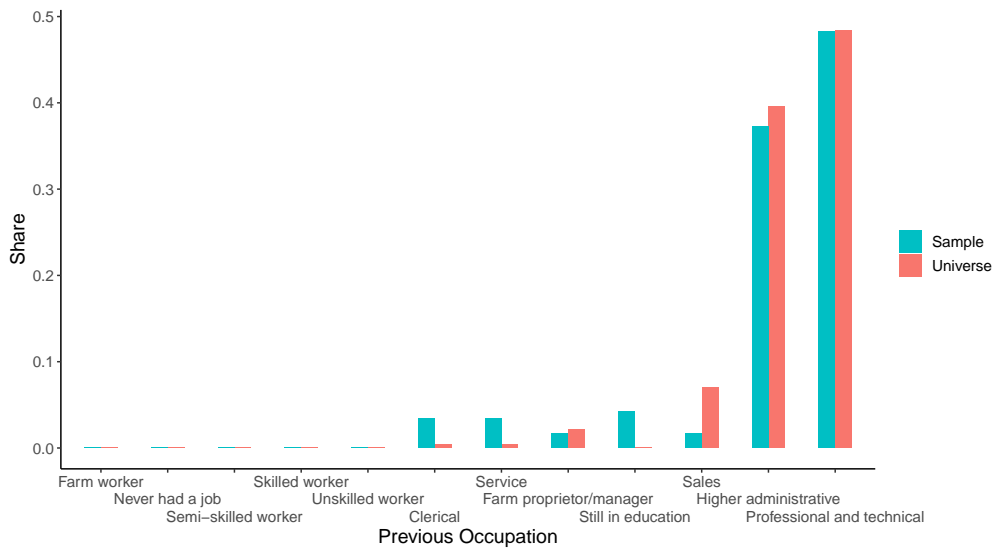
Figure 1.C.2. MEP-Sample Representativeness Regarding Occupation
Note: This figure compares the distribution of previous occupation of the universe of MEPs to the distribution on previous occupation of the sample of elected candidates from the Parliament Election Study 2009 (Candidate Study).
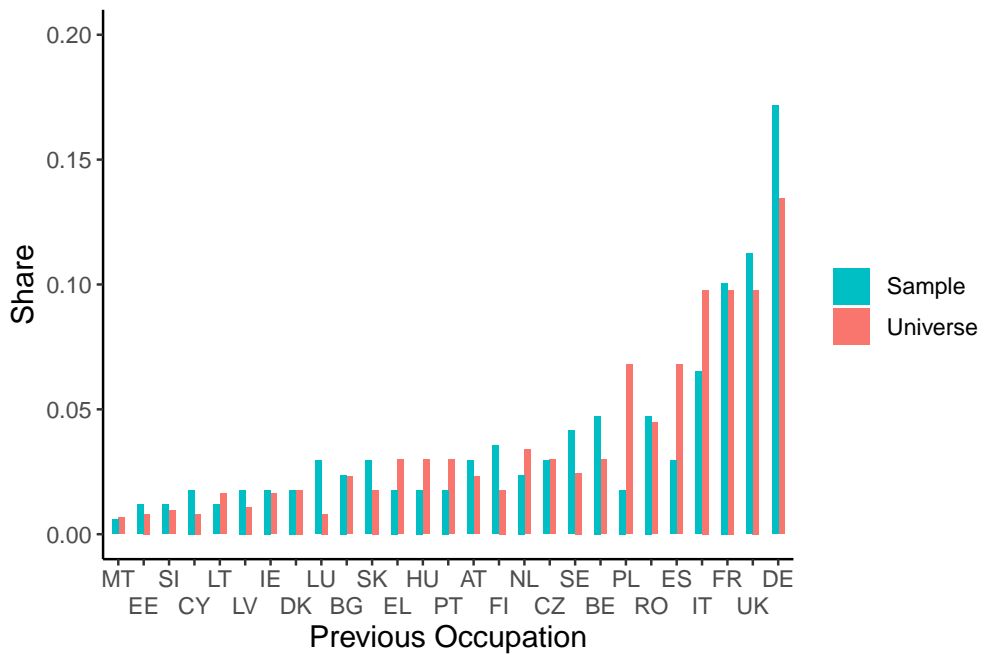
Figure 1.C.3. MEP-Sample Representativeness Regarding Country of Election
Note: This figure compares the distribution of country of election of the universe of MEPs to the distribution of country of election of the sample of elected candidates from the Parliament Election Study 2009 (Candidate Study).

and asylum seekers" issue from the CHES to the "assimilation" issue from the EES, and the "position on social lifestyle (e.g., homosexuality)" issue from the CHES to the same-sex marriage issue from the EES. The CHES contains a "position on deregulation" issue, while the EES contains positions on the related topics of "state intervention," "state ownership," and "private enterprise." I use the mean with equal weights of the three latter variables as the "deregulation" measure for the EES.

I obtain a measure based on the CHES and a measure based on the EES for each of these issues for 149 parties. If the EES data is valid, correlations between these two measures should be high on each issue. Figure 1.C.4 visualizes correlation coefficients surround by 95% confidence intervals on the horizontal axis by policy issue. Correlations are always above 0.6 and highly significant.
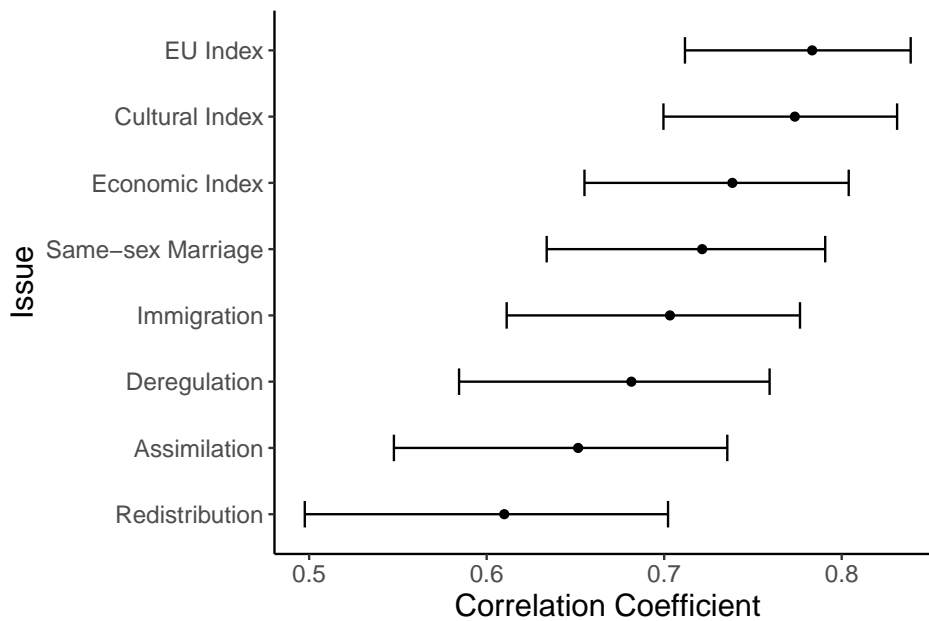
Figure 1.C.4. Correlations between MEP Survey Data and Expert Survey Data
Note: This plot shows cross-party correlation coefficients of two measures for party position by issue. Each point corresponds to the correlation regarding the political issue depicted on the horizontal axis. I also depict 95% confidence intervals.

Another concern regarding the validity of the EES data might be that it does not contain enough policy items to enable estimates of positions on broad political dimensions. The CHES contains estimates of parties' positions on broad economic, cultural, and EU dimensions. Policy experts were asked to estimate the "overall orientation of the party leadership towards European integration," the "position of the party in 2010 in terms of its ideological stance on economic issues," and the "position of the party in 2010 in terms of its ideological stance on democratic freedoms and rights." All three refer to general policy dimensions rather than specific policy issues and thereby capture the three dimensions I consider in the paper. If the in-

dexes I calculate based on the EES data are valid, they should correlate positively with these broad CHES variables. Figure 1.C.4 shows that this is the case. The correlation coefficients for all three dimensions are between 0.7 and 0.8 and are highly significant.

Appendix 1.D   Referendum Data Details

1.D.1   Descriptive Statistics

Figure 1.D.1 shows the number of all 126 referendums by policy issue contained in the referendum dataset. The topics have substantial overlap with the survey data issues. The vertical axis in Figure 1.D.1 shows the number of referendums in the dataset who belong to a topic. As can be seen, most referendums have been held on whether the state should intervene in an industry. The next frequent categories are assimilation, punishment for criminals, and immigration. For all topics except protectionism, there are at least three referendums in the dataset.

Of the 126 referendums that have been held since 1870, about 39% of the initiatives were right-wing. However, the right-wing share of referendums varies substantially by policy dimension. Among the 75 cultural referendums, about 61% were right-wing. In contrast, only 3 of the 46 referendums on economic topics (about 7%) were right-wing. Finally, none of my dataset's five referendums on EU integration was anti-EU-integration.
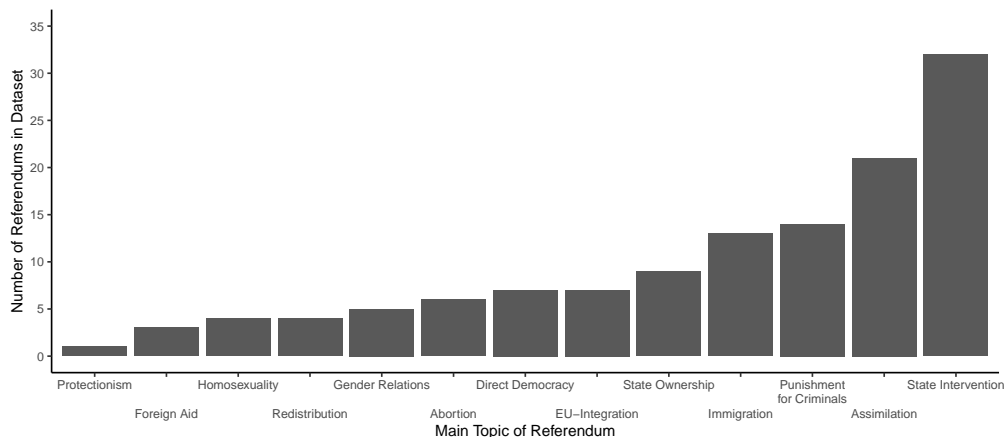
Figure 1.D.1. Number of Referendums by Topic
Note: This bar-charts shows how many referendums of each topic are contained in the referendum dataset.

Figure 1.D.2 provides information on the topics of referendums over time. For each decade (beginning in the 1890s), it shows the absolute number of cultural, economic, and EU referendums as a stacked area chart. As can be seen, the use of referendums became more frequent over time. This can be attributed to the rise of referendums on cultural topics. In contrast, referendums on economic matters have

become less frequent in recent decades. Referendums on EU integration have only been held since the existence of the EU and also became less frequent recently.
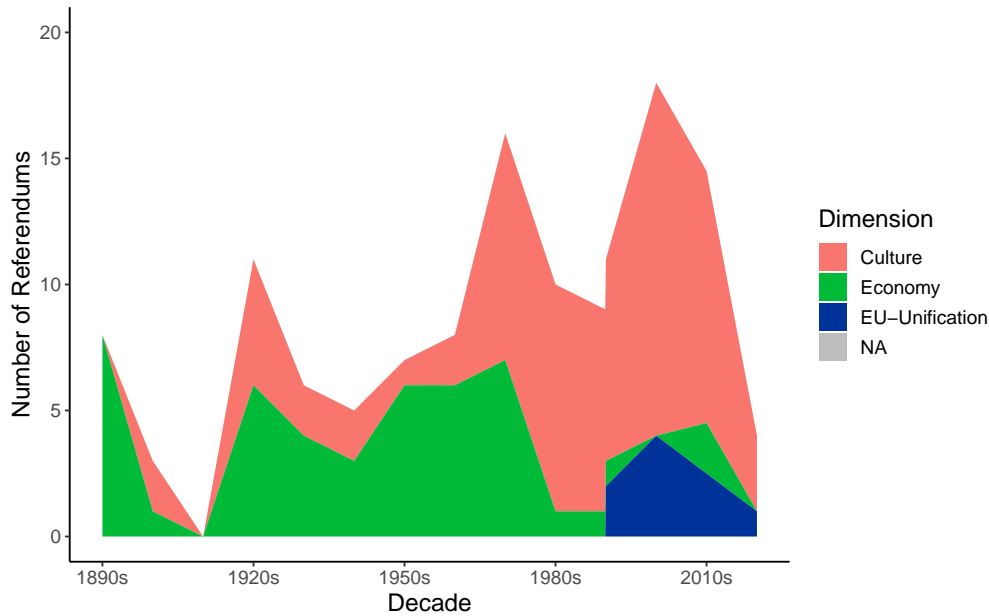
Figure 1.D.2. Number of Referendums by Dimension and Time
Note: The horizontal axis shows the decade and the vertical axis the number of referendums. Colors indicate the issue-topic. The area-chart is stacked.

1.D.2 Representation Gaps for all Policy Issues

I calculate the representation gap for a referendum r as

$$RG_r = \begin{cases} \text{share of "yes"-voting voters} - \text{share of "yes"-voting MPs}, & \text{if r is left-wing,} \\ \text{share of "yes"-voting MPs} - \text{share of "yes"-voting voters}, & \text{if r is right-wing.} \end{cases}$$

To make estimates for RGs based on referendums and survey data comparable I take the mean with equal weights of all RG_r belonging to a policy issue. Figure 1.D.3 depicts these average RGs on the vertical axis and policy issues on the horizontal axis. Higher values indicate that MPs are more right-wing compared to voters.

Voters are more right-leaning than MPs on most issues. The only exception is that voters are more in favor of redistribution than MPs. Positive RGs vary in magnitude. For some topics like gender relations, assimilation, and immigration, voters are about 25 percentage points more likely to choose the right-wing option than MPs. For other topics like abortion and direct democracy, the difference is much smaller.
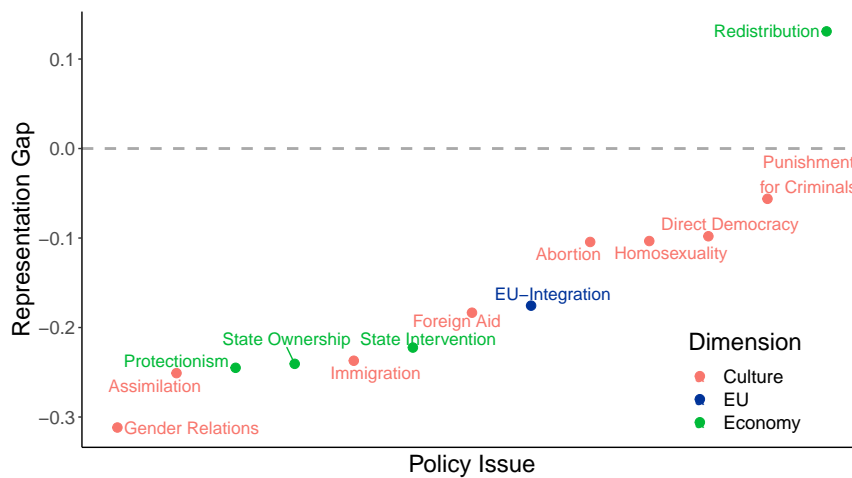
Figure 1.D.3. Representation Gaps between Swiss Voters and Swiss MPs by Issue
Note: The vertical axis depicts representation gaps. The horizontal axis refers to policy issues. Higher values indicate that the population voted more left-wing on that topic than the lower chamber of the Swiss parliament. Data is pooled over time.

1.D.3 Representation Gaps over Time

Figure 1.D.4 displays representation gaps for the three policy dimensions over time, pooling data for decades. The dataset contains data on referendums on EU issues for two decades: the 1990s and the 2000s. In both decades, voters were about 15 pp. more likely to choose the anti-EU option than MPs. Data for the cultural and economic dimensions are more extensive. As can be seen, their time trends differ strongly. In all but one decade, positive cultural RGs existed, i.e., voters were more culturally right-wing than MPs. Only in the 1920s were MPs more right-wing than their voters, and this data point relies on relatively few referendums. Hence, the cultural RG is qualitatively persistent over a long period of time. In recent decades it has decreased, but it was still positive in recent years.

In contrast, economic RGs have undergone a major transformation since the 1980s. At the end of the 19th century and during most of the 20th century, MPs were less market-oriented than their voters. In fact, the RG was similarly pronounced regarding economic as regarding cultural issues. But beginning in the 1980s, the RG switched signs. In the 80s, the 90s, and the 2010s (data for the 2000s is missing), MPs were much more likely to favor market-oriented solutions than voters. This qualitative change lines up well with anecdotal evidence. Many authors have argued that beginning in 1979, with the electoral victories of Margaret Thatcher in the UK and Ronald Reagan in the USA, a new era of economic policymaking began. In this era, market-oriented approaches became common, with former left-wing parties also adopting pro-market stances (Benedetto, Hix, and Mastrorocco, 2020).
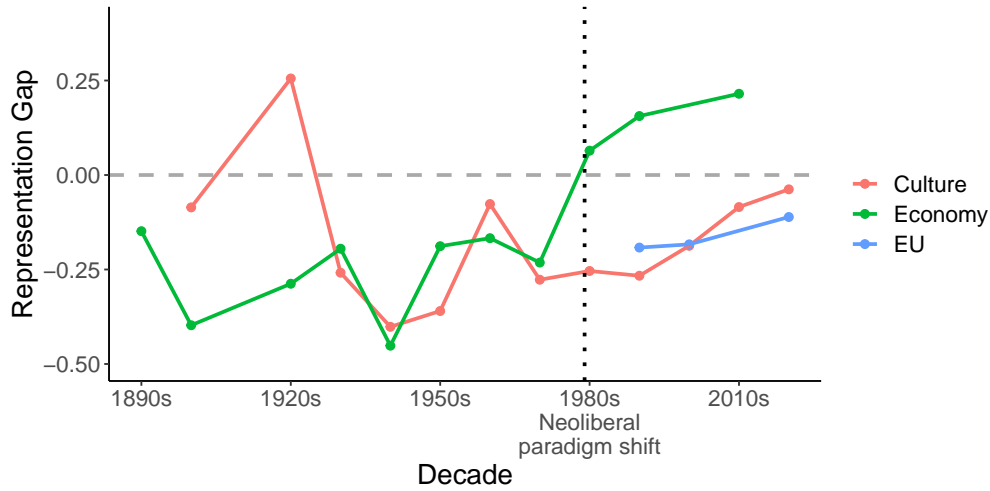
Figure 1.D.4. Representation Gaps (Voters vs MPs) Over Time
Note: The horizontal axis shows the decade. Positive values indicate that MPs voted more right-wing on referendums belonging to a given dimension within a given decade.

1.D.4   Calculating Representation Gaps by Comparing Voters with Parties

Differences between voters and MPs do not necessarily imply differences between voters and parties as a whole. For instance, party structures could give more weight to the opinions of party members which are closer to the electoral center. Hence, it is important to compare voters to parties. The referendum dataset contains information on official party recommendations for nearly all parties on all referendums. That makes it possible to compare the voting decisions of voters to party decisions. Formally, let r be a referendum with two options $\in \{yes, no\}$. Let the vote of voter k be denoted by $v(k)$. $v(k)$="yes" indicates that k is in favor of the referendum initiative and $v(k)$="no" indicates that he is opposed to it. Let $d(R) \in \{right, left\}$ be the direction of the initiative. $d(R)$ ="right" mean that the initiative aims to push policy-making to the political right. Let there a set of voters V. Let $rec(j)$ be the alternative that party j officially recommends to voters. Finally, let there be set of parties P and let $s(p)$ be the vote share in the last national election that party p got. I label RG_r^P the representation gap between voters and parties on referendum r and calculate it as:

$$
RG_r^P = \begin{cases} \dfrac{\sum_{i \in V} \mathbb{1}[v(i) = yes]}{||V||} - \sum_{p \in P} \mathbb{1}[rec(p) = yes] \cdot s(p), & \text{if r is left-wing,} \\[2em] \sum_{p \in P} \mathbb{1}[rec(p) = yes] \cdot s(p) - \dfrac{\sum_{i \in V} \mathbb{1}[v(i) = yes]}{||V||}, & \text{if r is right-wing.} \end{cases}
$$

Figure 1.D.5 shows the average RG_r^P by policy issue. Results resemble those in Figure 1.D.3, which compares voters and MPs. Figure 1.D.6 shows how RGs vary over time by dimension, comparing parties and voters. Results look similar to those

in Figure 1.D.4 where I compare voters to MPs. Overall these results highlight that the attitudes of MPs are a good indicator of their party's position and therefore illustrate the robustness of the results in the main text.
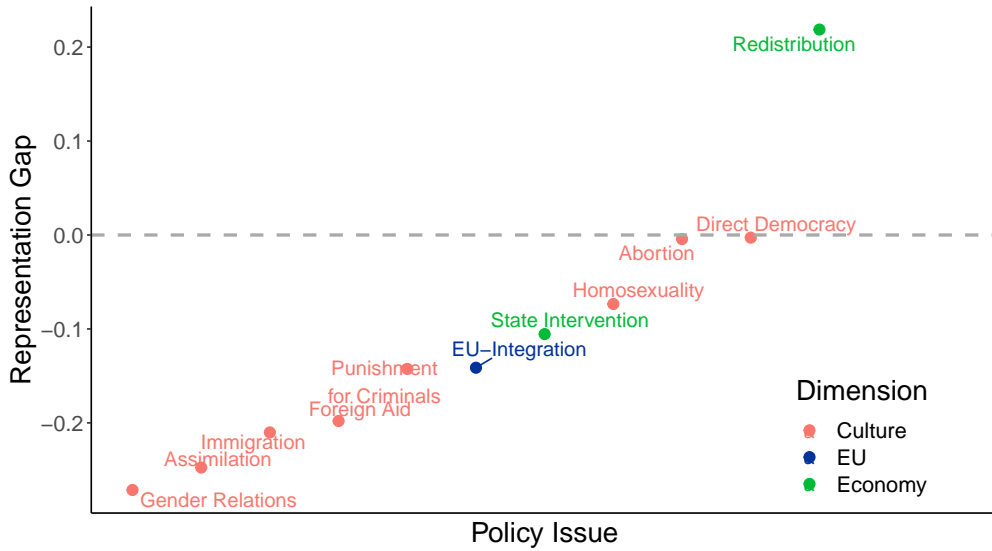
Figure 1.D.5. Representation Gaps between Swiss Voters and Swiss Parties by Political Issue
Note: The vertical axis depicts representation gaps. The horizontal axis refers to policy issues. Higher values indicate that parties, weighted with their vote share in the most recent national election, supported the left-wing alternative in referendums to a larger extend than voters in general.
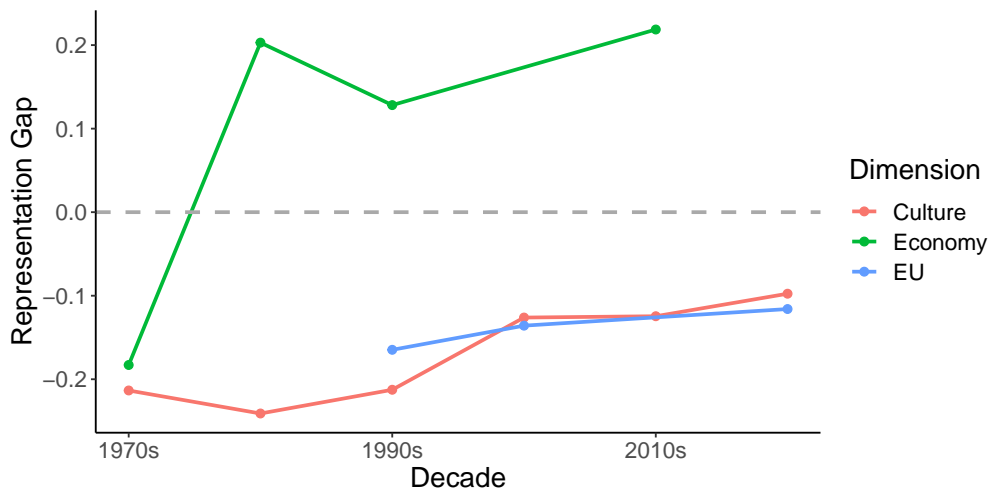
Figure 1.D.6. Representation Gaps (Voters vs. Parties) Over Time
Note: The horizontal axis shows the decade. Positive values indicate that parties, weighted with their vote share in the most recent national election, voted more right-wing on referendums than voters in general. Data is only available beginning in the 1970s.

1.D.5   Differences in Referendum Initiation Behavior between Voters and MPs

The existence of a RG would imply that MPs rather push for change to the political left than ordinary citizens. This implies that among referendum initiatives initiated by the people, the share of right-wing initiatives is higher than among those initiated by the elite. The previous analysis suggests that such a RG exists regarding cultural topics and, to a lesser degree, regarding EU issues but not regarding economic topics.

None of the EU-related referendums have been initiated by citizens. Hence, to test this prediction, I focus on the comparison between cultural and economic referendum initiatives. Figure 1.D.7 shows the share of initiatives with a right direction by originator of the initiative and dimension. The height of the bars indicates the share of right-wing initiatives in the group of referendums. The horizontal axis shows three types of originators; the elite (in most cases, the parliament, otherwise the government) and (ordinary) citizens. Red bars refer to initiatives regarding cultural issues, while blue bars refer to referendums on economic issues. I also depict 95% confidence intervals from an exact binomial test.
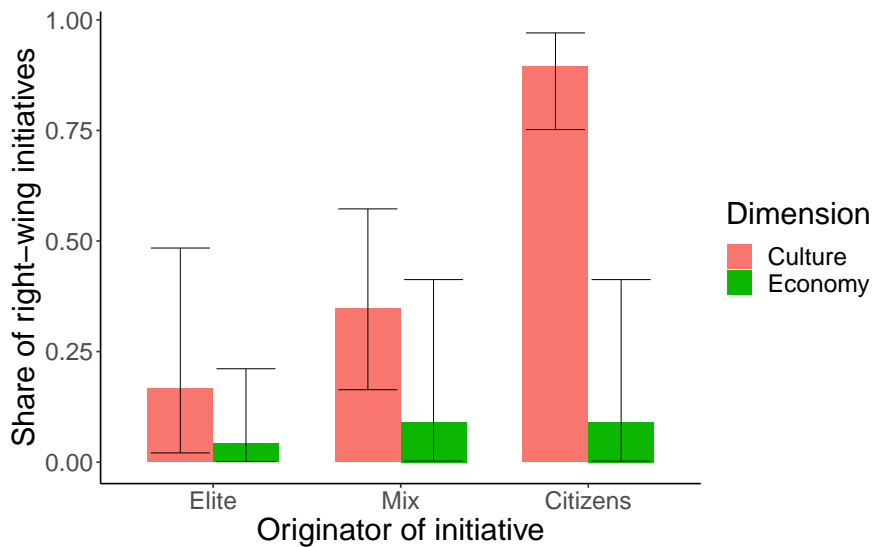
Figure 1.D.7. Share of Right-Wing Initiatives by Originator
Note: This figure shows the share of referendums with a right direction by originator. It is based on all referendums in the dataset on a cultural or economic topic. Confidence intervals are based on an exact binomial test.

Figure 1.D.7 confirms the expectations. Regarding economic issues, most initiatives are left-wing regardless of the originators. Taken literally, the point estimate suggests that citizens are only about five percentage points more likely than the elite to initiate a right-wing referendum, but this difference is far from being significant at conventional levels. Reassuringly, the share of right-wing initiatives that resulted from an interaction of the elite and citizens is similarly large. The pattern looks very different for cultural topics. Only about 16% of the referendums initiated by

the elite are right-wing. That suggests that the elite overwhelmingly proposes initiatives that are steps to the cultural left. This is qualitatively different for ordinary citizens. About 90% of the referendums initiated by the people constitute a step toward the cultural right. The difference is highly significant. Finally, the right-wing share of referendums that result from a mix of the two groups is between these two values. In sum, this evidence suggests that ordinary citizens and the elite disagree on which direction their country should be heading regarding the cultural dimension, while no such disagreement is found regarding the economic dimension.

Appendix 1.E   A Stylized Model with Policy-Motivated Candidates and Heterogeneous Valence - Proofs

There is a continuum of voters. Policy attitudes are distributed on a single continuous policy dimension according to CDF $A(\cdot)$ which I assume to be invertible. The attitude of voter i is denoted by a_i. Candidates, L, and R, announce policy positions x_L and x_R. The final policy is denoted by x. The utility of voter i is given by $w_{i,L}(x) = \lambda_L - (a_i - x)^2$ and his utility if R is elected equals $w_{i,R}(x) = \lambda_R - (a_i - x)^2$. λ denotes the valence of a candidate. I assume that $\lambda_L \geq \lambda_R$. The utility of candidate L equals $u_L = -(a_L - x)^2$ where a_L is the policy attitude of candidate L. The utility of candidate R equals $u_R = -(a_R - x)^2$ where a_R is the policy attitude of candidate R. I assume that $a_L < A(0.5)^{-1} < a_R$. I will prove that in every equilibrium of this game the final policy is biased away from the bliss point of the median voter toward the bliss point of L. Formally, in every equilibrium $x < A(0.5)^{-1}$.

Proof. Suppose for sake of contradiction that there is an equilibrium in which $x > A(0.5)^{-1}$. Suppose candidate L chooses $x_L = A(0.5)^{-1}$. If $x_R = A(0.5)^{-1}$ every voter votes for L because

$$
\begin{aligned}
\lambda_L &> \lambda_R \\
\lambda_L - (a_i - A(0.5)^{-1})^2 &> \lambda_R - (a_i - A(0.5)^{-1})^2 \\
\lambda_L - (a_i - x_L)^2 &> \lambda_R - (a_i - x_R)^2 \\
w_{i,L}(a_i) &> w_{i,R}(a_i)
\end{aligned}
$$

If $x_R > A(0.5)^{-1}$ then $\forall i$ with $a_i \leq A(0.5)^{-1}$

$$
\begin{aligned}
\lambda_L - (a_i - A(0.5)^{-1})^2 &> \lambda_R - (a_i - x_R)^2 \\
w_{i,L}(a_i) &> w_{i,R}(a_i)
\end{aligned}
$$

because $\lambda_L > \lambda_R$ and $(a_i - A(0.5)^{-1})^2 > (a_i - x_R)^2$. Similarly, if $x_R < A(0.5)^{-1}$ then $\forall i$ with $a_i \geq A(0.5)^{-1}$

$$\lambda_L - (a_i - A(0.5)^{-1})^2 > \lambda_R - (a_i - x_R)^2$$
$$w_{i,L}(a_i) > w_{i,R}(a_i)$$

because $\lambda_L > \lambda_R$ and $(a_i - A(0.5)^{-1})^2 > (a_i - x_R)^2$. Hence, choosing $x_L = A(0.5)^{-1}$ ensures that L gets elected. Moreover, $u_L(A(0.5)^{-1}) > u_L(x) \Longleftrightarrow -(a_L - A(0.5)^{-1})^2 > -(a_L - x)^2$. Hence, deviating to $A(0.5)^{-1}$ is profitable for L if $A(0.5)^{-1}$ is closer to a_L than x is to a_L. Hence, L has a profitable deviation which contradicts that $x > A(0.5)^{-1}$ in an equilibrium.

Now suppose for sake of contradiction that there is an equilibrium in which $x = A(0.5)^{-1}$. Then either $x_L = A(0.5)^{-1}$ or $x_R = A(0.5)^{-1}$, or both. I will show that in each case L has a profitable deviation. Any voter i votes for candidate L iff

$$\lambda_L - (a_i - x_L)^2 > \lambda_R - (a_i - x_R)^2$$
$$\lambda_L - \lambda_R > (a_i - x_L)^2 - (a_i - x_R)^2$$
$$\lambda_L - \lambda_R > -2a_i x_L + x_L^2 + 2a_i x_R - x_R^2$$
$$\lambda_L - \lambda_R + x_R^2 - x_L^2 > 2a_i(x_R - x_L)$$

which is equivalent to

$$a_i < \frac{\lambda_L - \lambda_R + x_R^2 - x_L^2}{2(x_R - x_L)}$$
$$a_i < \frac{\lambda_L - \lambda_R}{2(x_R - x_L)} + \frac{(x_R + x_L) \cdot (x_R - x_L)}{2(x_R - x_L)}$$
$$a_i < \frac{\lambda_L - \lambda_R}{2(x_R - x_L)} + \frac{(x_R + x_L)}{2}$$

if $x_L \leq x_R$ and

$$a_i > \frac{\lambda_L - \lambda_R}{2(x_R - x_L)} + \frac{(x_R + x_L)}{2}$$

if $x_L \geq x_R$. If $x_R = A(0.5)^{-1}$ candidate L can ensure that the majority of votes for him by choosing x_L sufficiently close to $A(0.5)^{-1}$ such that the expression on the right hand is a bit smaller than $A(0.5)^{-1}$. This is possible because the numerator of the first term is positive by assumption and the policy space is continuous.

If $x_R < A(0.5)^{-1}$, L could choose $x_L = x_R$ in which case every voter voters for him because $\forall a_i$.

$$\lambda_L > \lambda_R$$
$$\lambda_L - (a_i - x_L)^2 > \lambda_R - (a_i - x_R)^2$$
$$w_{i,L}(a_i) > w_{i,R}(a_i)$$

Finally, assume $x_R > A(0.5)^{-1}$. Then any x_L such that $|x_L - A(0.5)^{-1}| < |x_R - A(0.5)^{-1}|$ constitutes a profitable deviation. For any such x_L the median voter prefers to vote for L because x_L is closer to his attitude and $\lambda_L > \lambda_R$. The same holds for all voters with $a_i < A(0.5)^{-1}$. Hence, L wins. Moreover, L prefers $x = x_L$ to $x = x_R$ because if $A(0.5)^{-1}$ is closer to x_L than to x_R any $a < A(0.5)^{-1}$ must be closer to x_L than to x_R too and $a_L < A(0.5)^{-1}$.

Hence, L has a profitable deviation in each case when $x = A(0.5)^{-1}$ and when $x > A(0.5)^{-1}$. Therefore, in any equilibrium $x < A(0.5)^{-1}$. $\square$

Appendix 1.F Other Potential Causes of Representation Gaps

1.F.1 Perceived Importance of Political Issues

A potential explanation for representation gaps is that they arise on issues that voters do not care about. I measure perceived importance of a policy issue through the following survey item which was given to MEPs and citizens:

What do you think is the most important problem facing [COUNTRY] today?

Similar questions were also asked concerning the second and third most important problems. Answers were open-ended and recorded verbatim. They were then allocated into 146 categories. Hence, I have data on each subject's first, second, and third most important issues. Some of these issues are broad, like "Economic conditions." However, many are more specific, like "Effects of financial crisis on domestic/ EU/ global economy."

Figure 1.F.1 shows word clouds of the 40 most frequent categories of voters and the 36 most frequent categories for MEPs. A clear plurality of voters named unemployment the most important issue in their country. This makes sense, as this question was asked in 2009 when unemployment rates in many EU countries were high. Unemployment is followed by "Effects of financial crisis on domestic/ EU/ global economy," "Economic Conditions" and "National employment policies." The specific category that is next in line is "Immigration." For MEPs, "Economic Conditions" is the most frequent response. This is followed by "Effects of financial crisis on domestic/ EU/ global economy," "Unemployment," "Climate Change" and "Executive and Administrative Efficiency; Efficient government."

(a) Voters   **(b)** MEPs
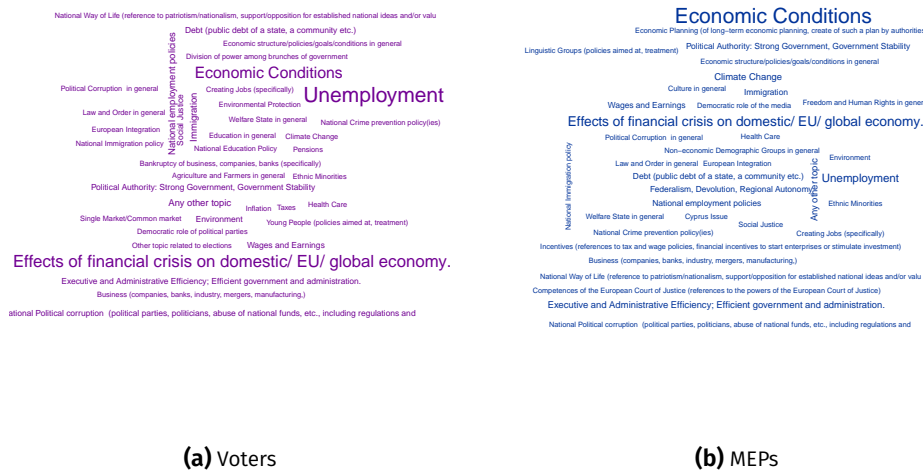
Figure 1.F.1. Word Clouds of Topics Considered to Be Most Important by Voters and MEPs
Note: This figure shows two word-clouds. Words and sentences are answers given to the following question:
"What do you think is the most important problem facing [COUNTRY] today?" The two clouds display the 40
most frequent categories of voters and the 36 most frequent categories for MEPs. More frequent responses
are displayed larger.

To compare voters and MEPs quantitatively, I construct an importance index
($II_{g,i}$) which measures how important a group g considers an issue i to be. Let
"share most important$_{g,i}$" denote the weighted share of respondents who consider
topic i most important and suppose similar definitions for the second and third most
important topic. All three shares are weighted to adjust for differences in population
between countries.

$$II_{g,i} = \frac{3 \cdot \text{share most imp.}_{g,i} + 2 \cdot \text{share 2nd most imp.}_{g,i} + \text{share 3rd most imp.}_{g,i}}{6}.$$

(1.F.1)

$II_{g,i}$ is distributed between zero and one, where one means that all subjects of group
$g \in \{\text{voters}, \text{MEPs}\}$ indicate that issue i is the first, second, and third most important
problem. It equals zero if no subject in group g considers issue i as belonging to
the three most important problems. To make the issue importance index and ARGs
comparable, I manually match issues relating to the two variables. I am able to do
this for 10 out of the 14 issues that I could calculate the ARG for.

Figure 1.F.2 compares $II_{g,i}$ to ARGs. Issues are shown on the horizontal axis. The
height of the blue bars depicts $II_{\text{MEPs},i}$, while the purple bars show $II_{\text{voters},i}$. For any
policy issue i and attitude on that issue a_i, the ARG is the estimate for β from the
following regression:

$$a_i = \alpha + \beta \cdot \mathbb{1}[\text{MEPs}]_i + \varepsilon_i.$$

The regression is weighted to adjust for population differences between countries. Due to the weighting, the figure compares a representative sample of those who voted in the 2009 European Parliament election with a representative sample of MEPs. Larger values indicate that MEPs are more right-wing relative to voters. I depict 95% confidence intervals around all values. To make ARGs and the importance index quantitatively comparable, I divide all ARGs by four.
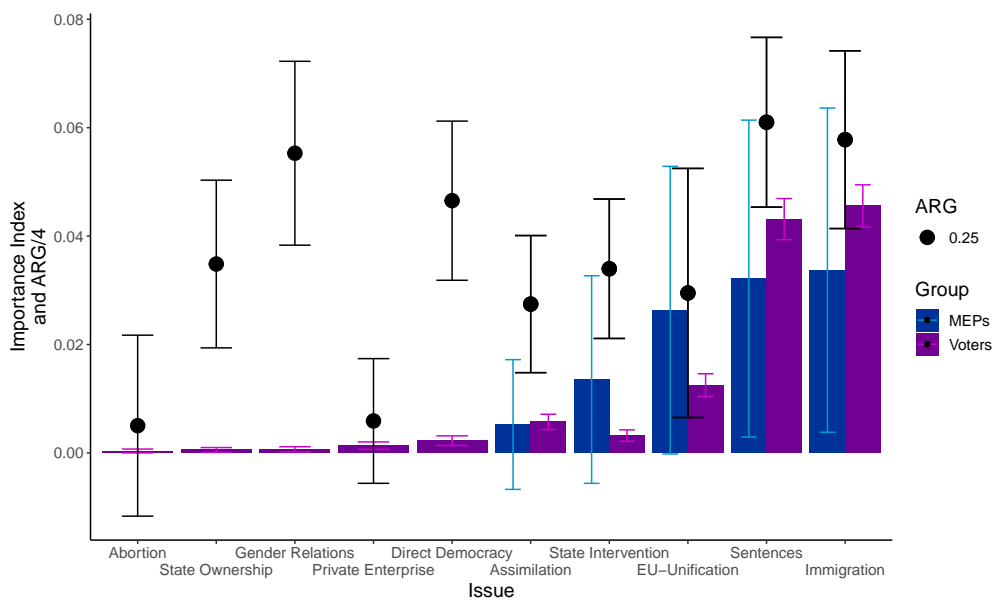
Figure 1.F.2. Perceived Importance of Policy Issues by Voters and MEPs Compared to Attitude Differences Between Voters and MEPs

Note: Bars indicate an index of perceived importance of issues ($II_{g,i}$), defined in Equation 1.F.1. Issues are shown on the horizontal axis while the importance index is shown on the vertical axis. Black dots represent the absolute value of the attitude representation gap by issue divided by four. I depict 95% confidence intervals around all values.

Figure 1.F.2 reveals that MEPs and voters tend to find the same topics important. Both groups agree that immigration, punishment of criminals, and EU unification are the most important issues. Voters find immigration and the punishment of criminals somewhat more important than MEPs, and MEPs find EU unification more important than voters, but these differences are not precisely estimated. Immigration and the punishment of criminals are the two topics where ARGs are the largest. This exemplifies the more general result that the perceived importance of topics is not negatively correlated with ARGs. If anything, the association appears to be positive.

How important are the three political dimensions relative to each other in the eyes of voters and MEPs? To answer this question, I manually classify each of the 146 categories as either cultural, economic, or EU-related. For most issues like unemployment or gender relations, this is straightforward. Some issues could be clas-
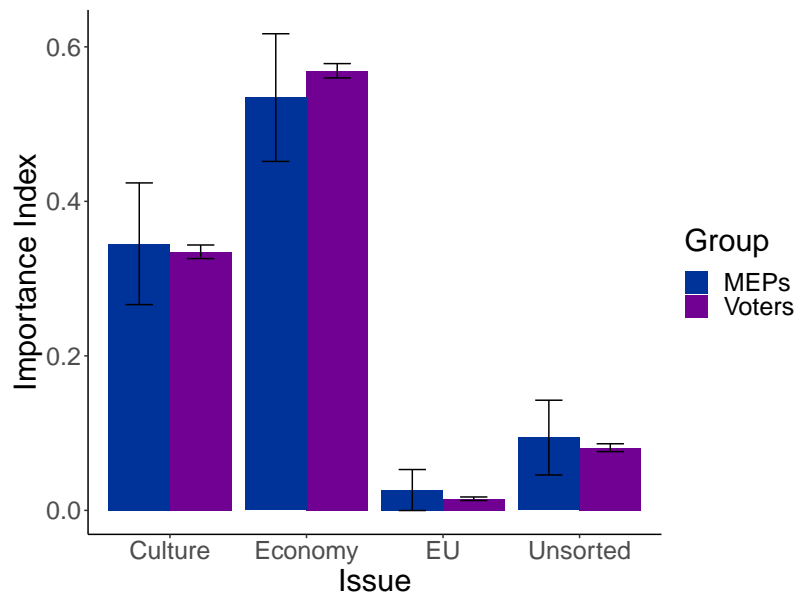
Figure 1.F.3. Perceived Importance of Policy Dimensions by Voters and MEPs
Note: Bars indicate an index of perceived importance of political dimensions ($II_{g,i}$), defined in Equation 1.F.1. Dimensions are shown on the horizontal axis while the importance index is shown on the vertical axis. I depict 95% confidence intervals around all values.

sified into several dimensions, like globalization. If a topic could be classified just as well in either dimension, I label it as "Unsorted." Figure 1.F.3 shows the importance index for the four categories. Again, results for voters and MEPs are similar. Unsorted issues are relatively unimportant to voters and MEPs. Although economic topics are more important to both groups, cultural topics are of great importance to both groups too. MEPs find topics related to the EU more important than voters, but both groups find them much less important than cultural or economic topics. This suggests that reducing the policy space in European countries to a two-dimensional economy-culture space captures most issues that are important to voters and MEPs. It also shows that the large cultural ARGs matter to voters. Overall the results suggest that voters and MEPs have similar attitudes on which topics are important while disagreeing on how one should deal with them.

1.F.2 Do Parliamentarians Try to Protect Minorities from a "Tyranny of the Majority?"

Another explanation for representation gaps is parliamentarians trying to protect groups they perceive to be vulnerable. If this were the case, one would expect groups like immigrants, women, and the poor to have very different policy attitudes than natives, men, and the rich on immigration, gender relations, and redistribution, respectively. Moreover, one would expect that parliamentarians hold attitudes

between these groups' attitudes to balance their conflicting attitudes. Finally, one would expect that parliamentarians have attitudes close to the attitudes of the group perceived to be vulnerable relative to that group's share among the population.

For instance, most natives might hold much more conservative attitudes regarding immigration and assimilation, while most immigrants might be liberal regarding these topics. Representing the mean voters' attitude would mean weighting the attitudes of natives and immigrants according to their relative size, which would put much more weight on natives' attitudes. Therefore, parliamentarians might fear representing the mean voter will effectively suppress immigrants' attitudes. To prevent this, they might overweight the attitudes of immigrants, thereby effectively shifting their policymaking away from the attitudes of the mean voter. Their policymaking would still likely be between the mean attitudes of natives and immigrants.

Inconsistent with this potential explanation, Figure 1.5 in the main text shows that the policymaking of MEPs is not between the attitudes of immigrants and natives, men and women, or the rich and the poor on any policy dimension. In contrast, policymaking is biased relative to the mean attitude of each group in the same direction.

However, Figure 1.5 provides results for aggregated policy dimensions, while parliamentarians might protect groups only on specific issues. Hence, it is more reasonable to consider topics where the conflict of interest between the structural minority and structural majority group is most obvious. To this end, Figure 1.F.4 compares the mean positions of the poor, those with a medium standard of living, the rich and MEPs regarding redistribution, the mean attitudes of natives, those with an immigration background, and MEPs on immigration and assimilation, and the mean attitudes of men, women, and MEPs on gender relations.[12]

The mean attitude of MEPs does not lie between the mean attitude of the structural minority and the mean attitude of the structural majority regarding any topic. Reassuringly, the poor are most in support of redistribution while those with a medium living standard are more opposed to it, albeit not as much as the richest third of the Europeans. However, the mean attitude of MEPs does not lie between these values. Rather, MEPs are significantly more opposed to redistribution than the rich. On immigration and assimilation, immigrants are more left-wing than natives. However, contrary to balancing the attitudes of these groups, MEPs have mean attitudes that are far more left-wing than those of immigrants. In fact, the position of immigrants is much closer to those of "ordinary" natives than to those of MEPs. On gender relations, men and women hold similar mean attitudes that do not differ significantly. If anything, women are more right-wing on that topic. MEPs are far to the left of both groups. This evidence is not consistent with the idea that MEPs deviate from the attitudes of the mean voter to protect structural minorities. The opposite is

12. Figure 1.F.4 shows attitudes instead of policymaking of MEPs for simplicity. Attitudes and policymaking of MEPs are very similar as shown in the main text.
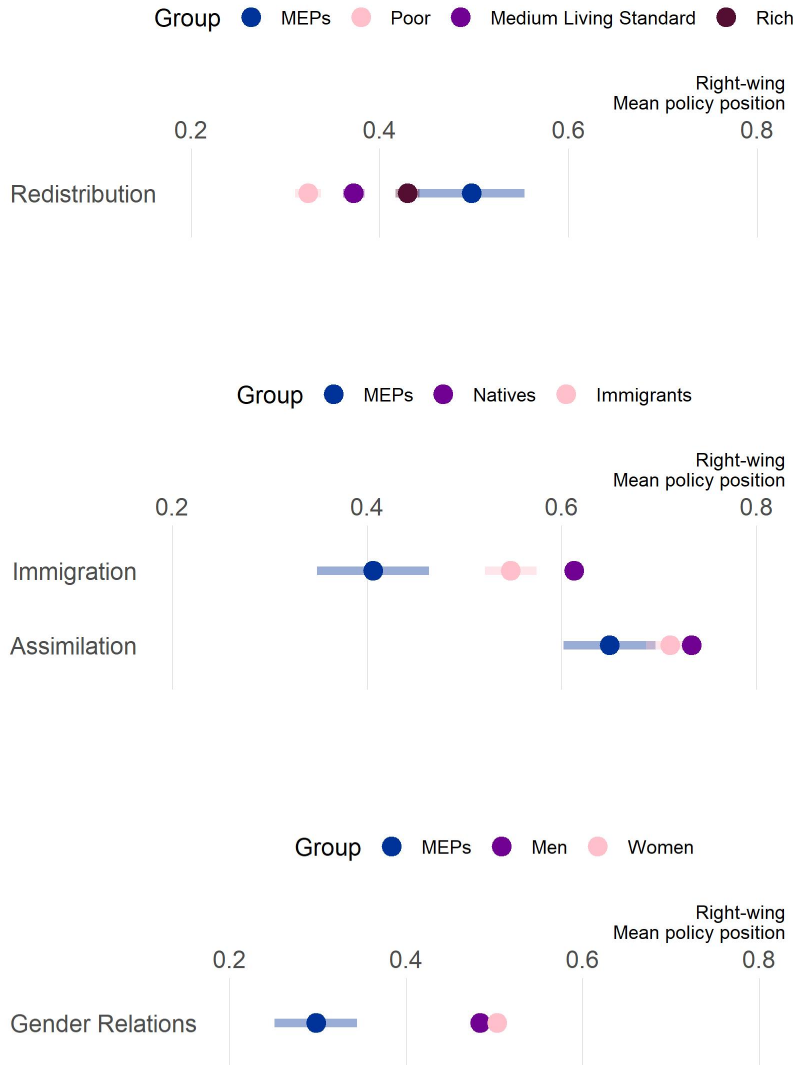
Figure 1.F.4. Mean Attitudes of Selected Voter and MEP Groups by Topic
Note: The vertical axis depicts political issues. Dots represent the mean attitudes of various groups regarding the respective issue. Higher values indicate that the mean attitude is more right-wing. The three panels show the attitudes of different groups. Groups other than MEPs constitute subsets of those who voted in the 2009 European Parliament election. I always weigh to account for differences in population size between countries. Bars indicate 95% confidence intervals.

the case. If they chose the position of the mean voter, their attitudes would be much closer to those of structural minorities than they are currently.

These results also provide evidence that in democracies, structural majorities can end up in a situation where they are disadvantaged relative to structural mi-

norities because the attitudes of MEPs are much closer to those of immigrants than those of natives. Finally, these results put the magnitude of attitude differences between voters and parliamentarians into perspective. Regarding redistribution, the difference between the mean attitude of MEPs and all voters[13] is larger than the difference between the mean attitude of the poor and the mean attitude of the rich. Regarding immigration and assimilation, differences between MEPs and the mean voter are much larger than the mean differences between natives and immigrants. Finally, mean attitudes between MEPs and voters on gender relations amount to about 20 percentage points, while mean attitudes of men and women do not differ notably.

1.F.3 Lobbyism

Organized lobby groups might influence politicians through campaign contributions, in which case it can be optimal even for office-seeking politicians to cater to the lobbyist's demands (Grossman and Helpman, 1996). To my knowledge, it has not yet been examined whether lobbyism leads to representation gaps. Most closely related to this question, the empirical literature on whether their actions reduce social welfare or not is not conclusive (Bombardini and Trebbi, 2020).

If lobbyism was a main contributor of the representation gap, one would expect that attitudes of parliamentarian (candidates) who are more closely related to lobbyists are more biased relative to the attitudes of their voters than candidates who are less closely related to lobbyists.[14]

The survey data includes six measures for relatedness to lobbyism by candidates. These measures include whether the candidate was encouraged to run for election by a lobbyist, whether the candidate himself was a lobbyist in the past, the number of hours per week the candidate spent with visits to firms or clubs, the number of hours per week the candidate's team spent with visits at firms or clubs, the amount donated to the candidate from outside his party and whether he plans to leave politics within the next 10 years. The idea behind the last measure is to measure whether a candidate has been offered positions outside politics by lobbyists akin to the "revolving door" phenomenon (Blanes i Vidal, Draca, and Fons-Rosen, 2012). While it is hard to single out candidates who made such agreements, it is unlikely that those who plan to stay in politics for the next 10 years did. Hence, the share of those who have made such agreements should be smaller among those who plan to stay in politics than those who plan to leave politics. When using the last measure,

---

13. The mean attitude of all voters is close to that of voters with a medium living standard, as all three groups are similarly large.

14. This section uses data on all candidates instead of data on elected parliamentarians only to gain precision. Attitudes of unelected and elected candidates do not differ notably as shown in Section 1.F.4.

it is necessary to condition the analysis on relatively young (I use below 50 years as the threshold) candidates.

None of these measures is a perfect measure for lobbyist influence on a candidate, particularly because they all rely on self-reports. However, if lobbyists bias policymaking away from the electorate, one would expect that at least most of these indicators correlate positively with the magnitude of the candidate's bias. I test this using OLS regressions. As these variables measure similar concepts, including several of them into one regression equation would lead to a "bad control" problem and bias estimates (Cinelli, Forney, and Pearl, 2020). Hence, I run individual regressions of the following type:

$$Bias_{i,d} = \alpha + \beta_1 \cdot Lobbyism_i + \theta \cdot C_i + \varepsilon_i, \qquad (1.F.2)$$

where

$$Bias_{i,d} = |a_{c,i} - \overline{a_i}|$$

is the bias of candidate i relative to the mean voter regarding dimension d. $Lobbyism_i$ is one of the measures described above and C_i contains a large number of demographic control variables. For each dimension I run one unconditional regression and one conditional on demographic control variables which results in six regressions for each lobbyism measure.

The results of these regressions are depicted in Table 1.H.5, Table 1.H.6, Table 1.H.7, Table 1.H.8, Table 1.H.9 and Table 1.H.10. Most of the lobbyism variables are insignificant at conventional levels, and most point estimates are very small. Only one of the 24 estimates is significant and positive (personal firm visits), but the estimate is very small. Moreover, three estimates are significant and negative. Most coefficients are precisely estimated, and one to two orders of magnitude too small to account for representation gaps if taken literally. Furthermore, there is no positive and significant estimate among the measures that are arguably most clean, like encouragement by lobbyists and whether the candidates have been lobbyists themselves. In sum, this evidence does not fit well with the explanation that lobbyism is a main driver of representation gaps in Europe.

1.F.4   Does Becoming Elected Affect the Policy Attitudes of Parliamentarians?

The political science literature on representation has usually interpreted representation gaps causally. That is, becoming an MP is assumed to alter one's policy preferences, for instance, through acquiring new information (Kertzer, 2022). Despite this interpretation, most estimates of RGs have been unconditional, which prevents causal claims (Kertzer, 2022). Recently, Kertzer (2022) made a first step in establishing causality by estimating RGs in the USA, controlling for several demographic variables. His results suggest that about half of the representation gaps, there can be explained by the selection of MPs based on demographic variables. In this section, I make a few steps forward to assess the extent to which RGs are causal.

Understanding whether the RG is causal or based on selection is important for welfare considerations. If the gap is causal, it might be driven by MPs' superior information, implying that the gap should not be seen as a problem. This is less true if the RG is just due to the selection of a particular group of people into positions of power.

In a first step of the analysis, I compare voters, elected MEPs, and unelected MEP candidates by estimating the following specification by OLS:

$$Y_i = \alpha + \beta_1 \cdot \mathbb{1}[\text{Candidate}]_i + \beta_2 \cdot \mathbb{1}[\text{Elected}]_i + \theta \cdot C_i + \varepsilon_i, \qquad (1.\text{F}.3)$$

where Y_i is a policy-attitude, $\mathbb{1}[\text{Candidate}]_i$ equals one if the individual was a *candidate* for the European Parliament and zero otherwise, $\mathbb{1}[\text{Elected}]_i$ equals one for all candidates that were elected and zero for all other subjects, including noncandidates. Finally, C_i is a vector of control variables and ε_i is the error term. The control variables include a large set of categorical variables like occupation categories, religion categories and the level of urbanisation of the place of residence.

I am interested in β_1 and β_2. For any policy issue, β_1 measures how the attitudes of unelected candidates and voters differ, while β_2 quantifies how elected MEPs differ from unelected candidates. If differences in attitudes between voters and elected MEPs result mainly causally from candidates becoming MEPs, than one would expect that β_2 is large and highly significant while β_1 should be small in comparison. In contrast, if differences result mainly from selection of citizens with particular attitudes into politics, the opposite should be the case. Most of the differences between voters and elected MEPs should be explained by β_1 while β_2 should be insignificant. Moreover, a strong role of selection would indicate that controlling for demographic characteristics decreases β_1.

Let's put these predictions to the test. Table 1.H.3 show estimates of Equation 1.F.3 for Y_i equal to the cultural index in columns (1) and (2), the economic index in columns (3) and (4) and the EU index in columns (5) and (6). For each index, the first column shows the unconditional estimates while in the second I add a large set of fine-grained controls. I only report β_1 and β_2 for readability.

Attitude differences between voters and elected MEPs can be seen by adding the MEP indicator and the MEP candidate indicator. Reassuringly, this analysis finds that MEPs are much more left-leaning on cultural issues than voters. The difference amounts to more than 11% of the range of the outcome variable and is highly significant. Moreover, note that basically all of this difference stems from the candidate indicator ($\beta_2 = -0.118$) while the elected-MEP indicator is close to zero, far from being significant at conventional levels, and even positive. Hence, MEP candidates and elected MEPs have, on average, nearly the same attitudes regarding cultural issues. This evidence is hard to bring in line with a strong causal effect of becoming an MEP. Rather, people from which MEPs are elected already hold very different attitudes than voters. This holds after including a rich set of demographic control

variables. This reduces the coefficient on the candidate indicator by about 40%, but it still stays large and highly significant. In contrast, the indicator for elected MEPs remains insignificant and close to zero. This, too, suggests that a strong causal effect of becoming an MEP on cultural attitudes is unlikely. However, even after conditioning on a very large set of demographic differences, large and highly significant attitude differences remain. This suggests that selection into politics might also occur concerning other variables, such as moral values.

As shown in the main text, attitude differences between MEPs and European voters are small in the economic dimension. This is consistent with the results displayed in column (3) in Table 1.H.3. MEP candidates are somewhat more left-leaning than voters, while the elected are significantly more right-wing than the candidates. A Wald test confirms that the sum is not significantly different from zero. After including controls, β_2 loses its significance while β_1 becomes significant. This suggests that MEPs and MEP candidates are quite similar regarding their attitudes. Unconditionally MEP's economic policy attitudes closely resemble those of the average European. But they are more left-leaning on economic issues than Europeans with otherwise similar demographic characteristics. While the highly educated and wealthy tend to be more market-oriented than society as a whole, MEP candidates are not. This suggests that selection is taking place from the higher strata of society to MEP candidates and is consistent with a well-functioning political screening process concerning economic attitudes in the European Parliament.

Columns (3) and (4) focus on EU-Unification. Unconditionally, elected MEPs are more left-leaning than voters. This can be largely attributed to elected MEPs who are much more in favor of EU integration than candidates in general. However, after accounting for demographic differences between elected and unelected candidates, the latter difference becomes insignificant. In fact, including demographic controls renders β_1 and β_2 insignificant. This again speaks against the causal hypothesis because the causal effect should be estimated more precisely if more controls are included.

In sum, the evidence for all dimensions speaks against a large causal effect of becoming an MEP. RGs do not arise once an MEP is elected but are already nearly fully present at the population of candidates from which MEPs are selected. Much of the differences between candidates and MEPs on the one side and voters on the other side can be explained through demographic differences. However, even after controlling for these differences a significant and large representation gap between MEPs and voters remains unexplained.

While the results so far speak against a large causal effect, I propose another specification which addresses two shortcomings of Equation 1.F.3. First, one might not expect an effect of election on attitudes for MEPs who already served as MEPs before. Hence, it seems reasonable to only include MEPs in the sample that have not served as MEPs before. Second, the set of controls used above misses crucial variables. For instance, it is reasonable to assume that party support raises the chances

of election. Not accounting for such differences might bias estimates of the effect of election.

To address both shortcomings, I drop voters from the analysis and focus on the comparison of elected and unelected MEPs. Moreover, instead of using attitudes directly as the dependent variable, I define a new variable: the Bias of an MEP candidate. Let $a_{c,i}$ be the attitude of MEP candidate c on issue i and let $\overline{a_i}$ be the European mean attitude of voters on issue i. I define the bias of candidate c regarding attitude a as

$$Bias_{c,a} \,=\, |a_{c,i} - \overline{a_i}|.$$

Hence, $Bias_{c,a}$ measures the political distance regarding attitude a to the European mean voter. I estimate the following equation by OLS:

$$Bias_c \,=\, \alpha + \beta_1 \cdot \mathbb{1}[\text{Elected}]_c + \theta \cdot C_c + \varepsilon_c. \tag{1.F.4}$$

I am interested in β_1. As before, C_c is a vector of control variables. I control for the following demographic variables: age, gender, categories for marital status, categories for hometown size, occupation categories, education. These variables constitute a subset of those employed in Equation 1.F.3. I do this for two reasons. First, the number of observations is substantially lower here. Second, MEP candidates are already very similar demographically. Differences between them are larger regarding their standing within their party and the support they got from other political actors. Accordingly, I prefer to include controls for these factors.

Hence, I additionally include the following controls: the number of those who helped to organize the campaign of the candidate, dummies for whether they were encouraged to run for office by a sitting MEP, a retired MEP, some other community leader, a lobbyist, their spouse, another family member, someone else, by someone from outside their party or from no one. I also control for the chances of being elected as assessed by the survey administrators. Chances are estimated based on the candidate's list position relative to the potential number of seats won by their party (Hix and Noury, 2009). Hix and Noury (2009) classified candidates with a list position below the predicted seats minus one standard deviation as safe, those positioned above the predicted seats plus one standard deviation as unpromising, and all other candidates as doubtful. I include this measure as a categorical variable in the regression. Although it is coarse, comparing candidates within one category helps to mitigate selection bias.

Table 1.H.4 shows the results. As can be seen, the results of the previous exercises are confirmed. Neither specification shows a significant coefficient on the elected indicator. Taken together, these results are more consistent with the selection of citizens with specific policy attitudes into politics than with a causal effect of getting elected. Moreover, they suggest that attitude differences are probably not an artifact of elected parliamentarians interpreting questions differently than voters. Finally, as most candidates already differ from the population, it is difficult for voters to reduce representation gaps through voting.

1.F.5   At Which Stage Of Political Selection Do Representation Gaps Arise?

One can think of citizens having to complete several stages until they become parliamentarians. First, they have to join a political party. Second, they have to rise in the ranks of this party to be nominated as a candidate for parliament. Finally, they have to get elected. At which stage of this political selection process do representation gaps emerge?

To locate where RGs arise, I focus on attitude differences. The rich survey-level data makes estimating attitude differences for various sub-groups possible. Hence, I compare the attitudes of MEPs with several sets of actors who are incrementally closer to MEPs. Formally, I estimate regressions of the following form by OLS:

$$Index_{i|i \in M \cup G} = \alpha + \beta \cdot \mathbb{1}[\text{MEP}]_{i|i \in M \cup G} + \gamma \cdot X_{i|i \in M \cup G} + \varepsilon_{i|i \in M \cup G},$$

where $Index_{i|i \in M \cup G}$ is the (cultural/economic/EU) index value of individual i which belongs either to the group M of elected MEPs or to another reference group (G), $\mathbb{1}(MEP)_{i|i \in M \cup G}$ equals one if $i \in M$ and zero otherwise and $X_{i|i \in M \cup G}$ is a vector of control variables. I am interested in β, which measures attitude differences between MEPs and the reference group.

To examine different stages of the political selection process I vary the reference group G. I consider the following groups: citizens, voters, those who are interested (but not very interested) in politics, those who are very interested in politics, and unelected MEP candidates. For each of these groups, I estimate one unconditional regression and one regression conditionally on the large set of demographic characteristics. These include age, gender, categories for marital status, occupation categories, dummies for the highest education degree, dummies for the size of the town of residence, dummies for the perceived own social class, immigration background, dummies for the religious denomination, dummies for religiosity and dummies for living standard.

Figure 1.F.5 shows estimates for β for the resulting estimates for the cultural index together with 95% confidence intervals. Attitude differences. Attitude differences between voters and parliamentarians are very similar to the representation gaps estimated in the main party of the paper. Using voters or citizens as the reference group does not lead to notably different estimates. When using voters or citizens as the reference group, including an extensive set of demographic characteristics as controls reduces attitude differences somewhat, but it stays large and significant. Hence, attitude differences regarding cultural issues do not result from demographic differences between parliamentarians and voters or citizens, like different educational attainment or age.

The first step toward becoming an MEP is to get involved in politics. To do so, one has to find an interest in politics. The third and fourth pairs of bars in Figure 1.F.5 compare the cultural attitudes of MEPs to those who stated that they were interested and very interested in politics, respectively. Focusing on unconditional
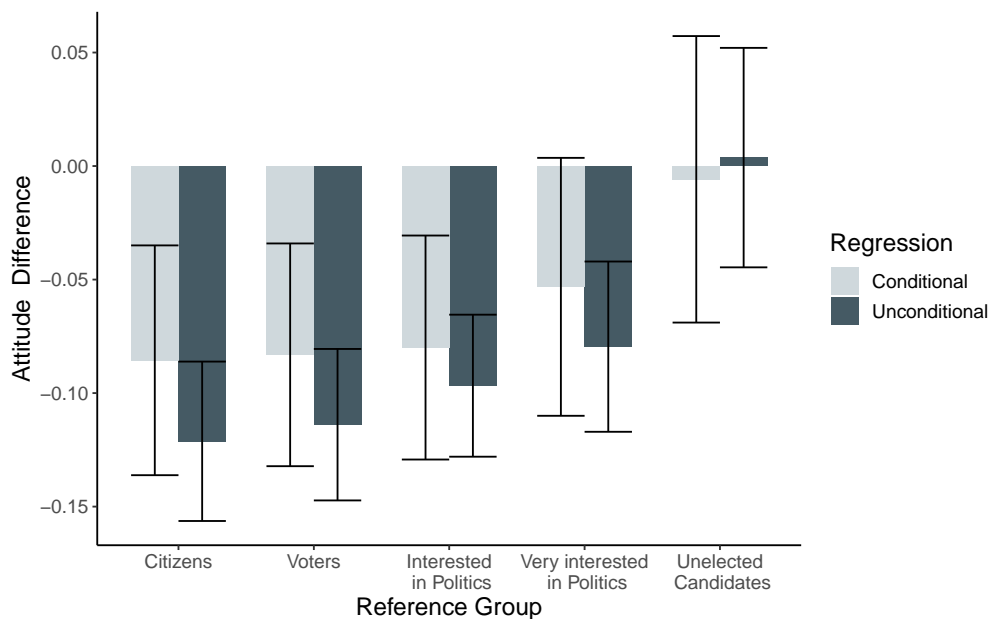
Figure 1.F.5. Attitude Differences of Selected Groups to MEPs on Cultural Issues
Note: This figure shows estimates of β from Section 1.F.5 for different reference groups *G* and the cultural index as the dependent variable. Coefficients compare the cultural attitudes of MEPs to the attitudes of reference groups displayed on the horizontal axis. The vertical axis measures attitude differences. Higher values indicate that members of the reference group are more conservative regarding cultural issues than MEPs. Control variables used to generate the conditional estimates include age, gender, the year of the election, categories for marital status, categories for town size, religion categories, and occupation categories.

estimates shows that attitude differences decrease relative to the MEP-voter comparison. Still, even those who are very interested in politics have attitudes closer to those of voters in general than members of the European Parliament. Moreover, they are significantly more conservative than MEPs. Hence, attitude differences can not be explained by parliamentarians trying to represent those interested in politics.

Those who are interested in politics and have similar demographic characteristics as voters and citizens with similar demographics as MEPs. However, those who are very interested in politics and have similar demographics as politicians are closer to MEPs. Their attitudes are about halfway between those of voters in general and MEPs.

Among those who are very interested in politics and join a party, a few rise high enough in the ranks of their party to become candidates for the parliament. The last two bars in Figure 1.F.5 compare the attitudes of elected MEPs and candidates who were not elected. Unelected candidates do not differ notably from elected MEPs, neither unconditionally nor after accounting for (small) demographic differences.

In sum, Figure 1.F.5 shows that attitude differences to MEPs decrease monotonically as the reference group gets closer to the group of MEPs. By far, the largest decrease happens between those very interested in politics and the MEP candidates.
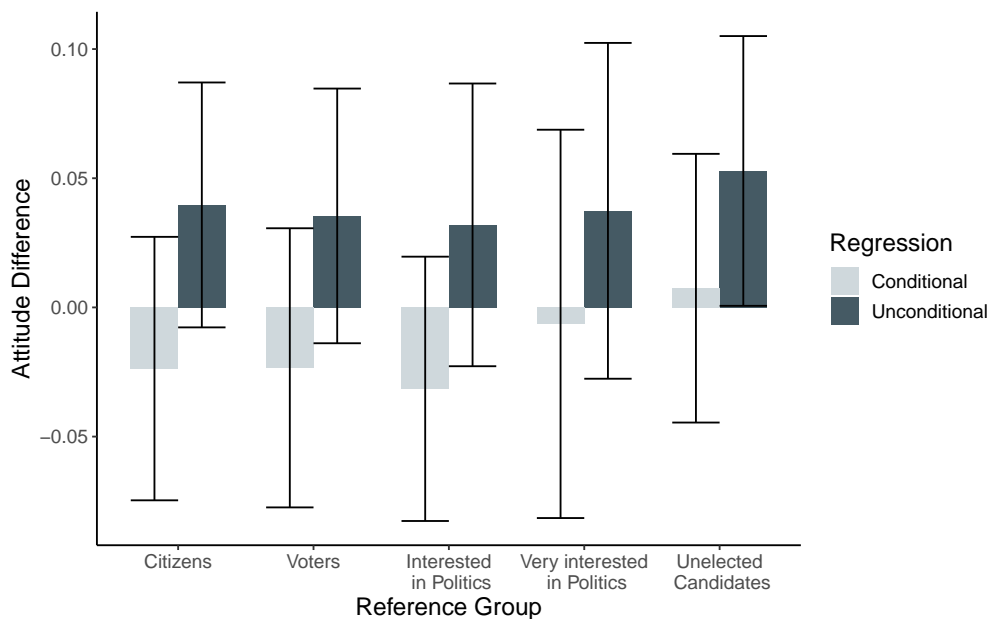
Figure 1.F.6. Attitude Differences of Selected Groups to MEPs on Economic Issues

Note: This figure shows estimates of β from Section 1.F.5 for different reference groups G and the economic index as the dependent variable. Coefficients compare the economic attitudes of MEPs to the attitudes of reference groups displayed on the horizontal axis. The vertical axis measures attitude differences. Higher values indicate that members of the reference group are more right-wing regarding economic issues than MEPs. Control variables used to generate the conditional estimates include age, gender, the year of the election, categories for marital status, categories for town size, religion categories, and occupation categories.

This suggests that parties are responsible for a large part of attitude differences and representation gaps, as their role in politics is to turn those interested in politics into politicians.[15]

Figure 1.F.6 shows a similar analysis for the economic dimension. The pattern looks very different. As shown in the main text, the attitudes of voters and parliamentarians do not differ significantly in the economic dimension overall. This is still true after including demographic controls in the regression. The same is true for voters and all other reference groups considered. If anything, unelected MEP candidates are more left-leaning on economic topics unconditionally, but this difference vanishes after accounting for demographic differences.

Finally, Figure 1.F.7 shows results for the EU dimension. The pattern resembles the pattern concerning the cultural dimension. Overall, accounting for demographic differences reduces attitude differences even though they remain significant regarding most reference groups. The unconditional attitude differences decrease monotonically as reference groups become more similar to elected MEPs. Insignificant attitude differences are only found for unelected candidates and those who are in-

---

15. Performing a similar analysis using the sample of national MPs yields very similar results.
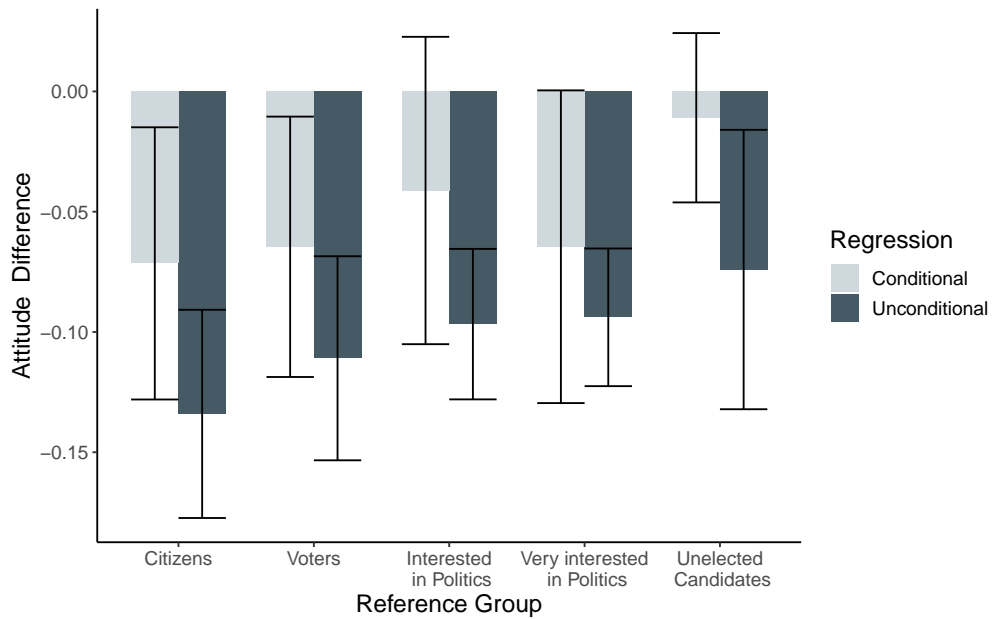
Figure 1.F.7. Attitude Differences of Selected Groups to MEPs on EU-Related Issues

Note: This figure shows estimates of *β* from Section 1.F.5 for different reference groups *G* and the EU index as the dependent variable. Coefficients compare EU-related attitudes of MEPs to the attitudes of reference groups displayed on the horizontal axis. The vertical axis measures attitude differences. Higher values indicate that members of the reference group are more anti-EU than MEPs. Control variables used to generate the conditional estimates include age, gender, the year of the election, categories for marital status, categories for town size, religion categories, and occupation categories.

terested in politics with similar demographics as the elected MEPs. The most crucial difference to the results for the cultural dimension is that, unconditionally, the EU attitudes of elected and unelected MEP candidates differ strongly. Hence, attitude differences in the EU dimension do not arise within parties but after candidates get elected. This difference appears to be driven by selection based on observable demographic characteristics because controlling for them renders the attitudes of elected and unelected candidates very similar.

Appendix 1.G   Additional Figures
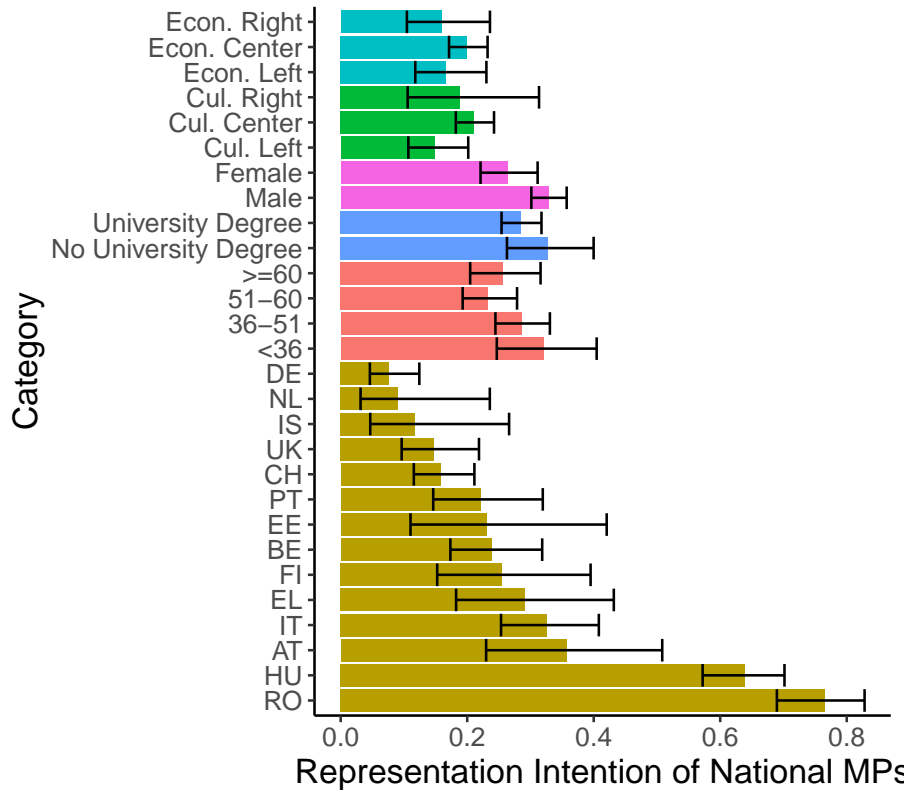
Figure 1.G.1. Representation Intention of National MPs by Demographic Group

Note: This bar-charts illustrate the responses of an MP-sample to the following question: "An MP in a conflict between own opinion and the constituency voters should follow:" Possible answers included "own opinion" and "voter opinion." Bars indicate the share that chose "voters opinion." The vertical axis shows different demographic groups of MPs. I also depict 95% confidence intervals.
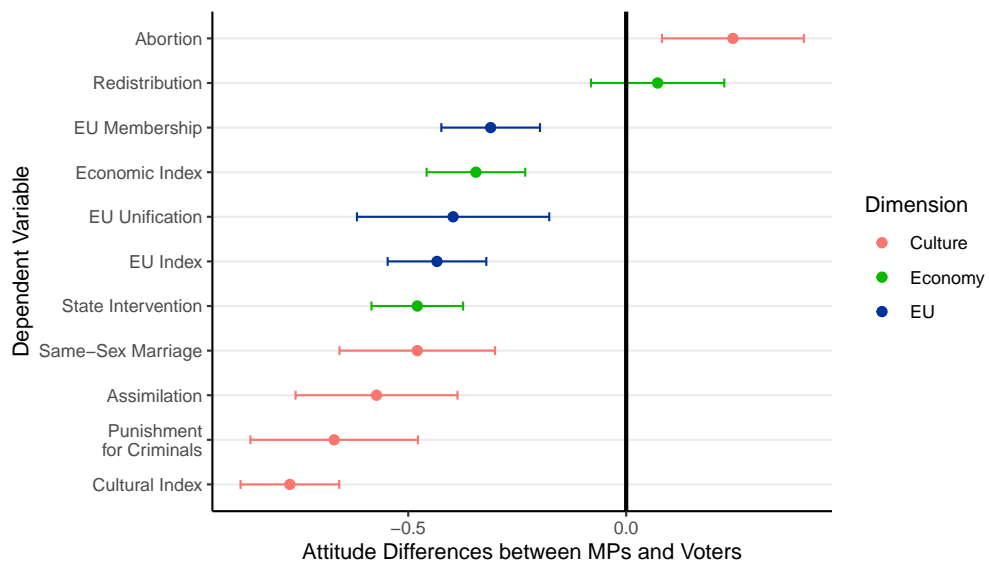
Figure 1.G.2. Attitude Differences between Voters and National MPs by Issue
Note: The vertical axis shows different policy issues or policy dimensions. The horizontal axis shows OLS estimates for β's from Equation 1.2 but using attitudes of national MPs instead of MEPs. Regressions compare national MPs with a representative sample of those who voted in the national election. All variables are scaled to have a standard deviation in terms of citizen attitudes of one. Higher values indicate that voters are more right-wing on an issue than MEPs. 95% confidence intervals are based on standard errors clustered at the country level.
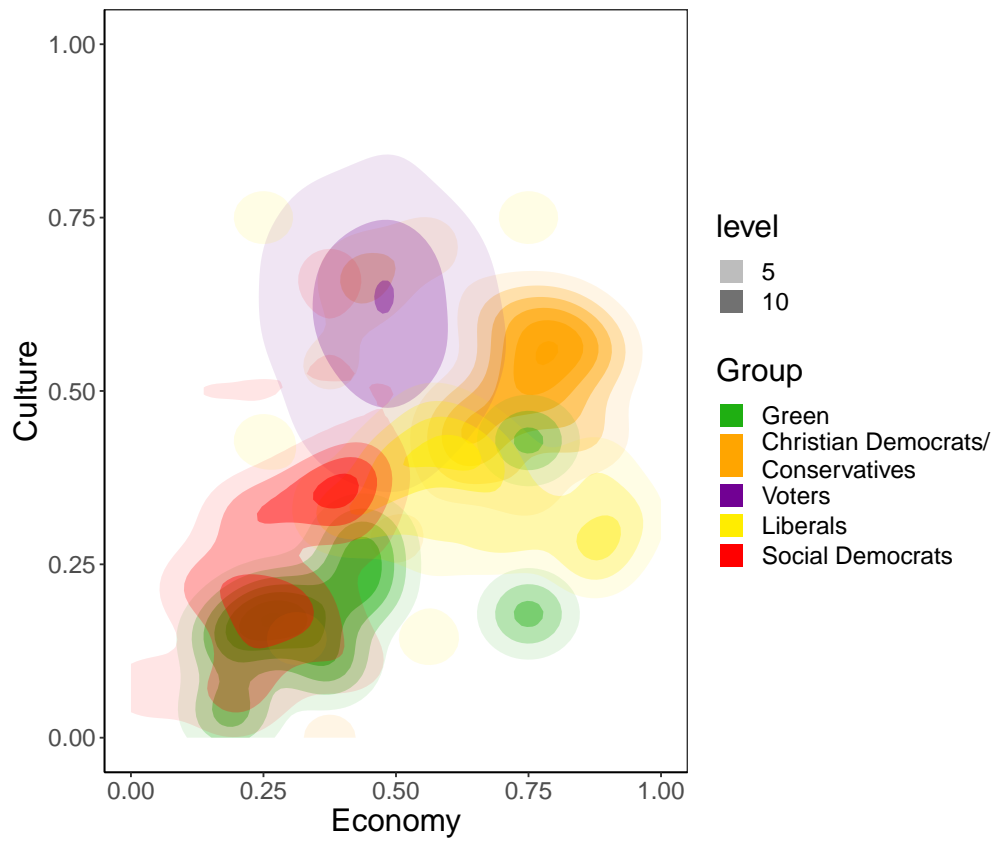
Figure 1.G.3. Attitude Distributions of Voters and MEPs of Several Party Groups in the Culture-Economy Policy Space

Note: Shades illustrate 2D densities. The density is higher in less transparent areas. Colors indicate the group that the density relates to. Attitudes in the cultural dimension are shown on the vertical axis, while the horizontal axis shows attitudes regarding the economic dimension. Higher values indicate a more right-wing stance. Data is pooled across Europe.
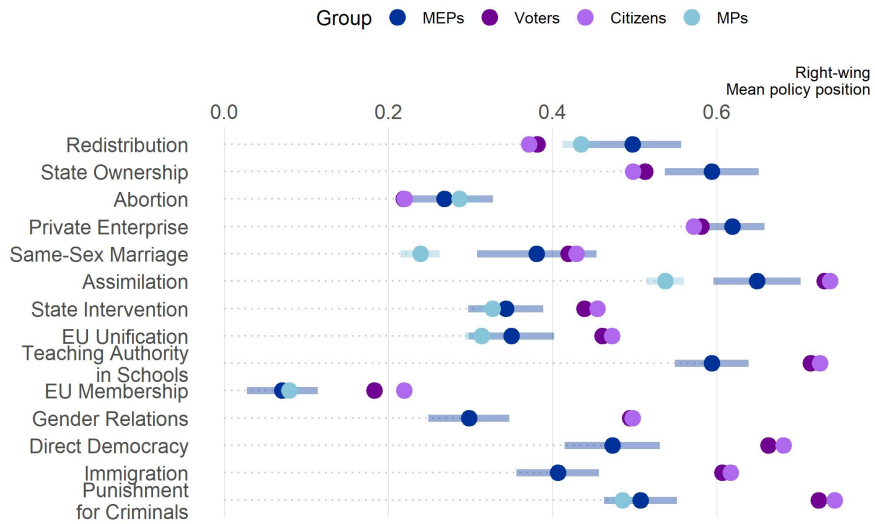
Figure 1.G.4. Mean Attitudes of Voters, MEPs, Citizens and National MPs by Issue
Note: This figure shows a dumbbell plot. Different policy issues are shown on the vertical axis. The horizontal axis shows the mean positions of three groups. All variables are scaled to range from zero to one and such that higher values indicate a more right-wing position. Data is pooled across Europe.
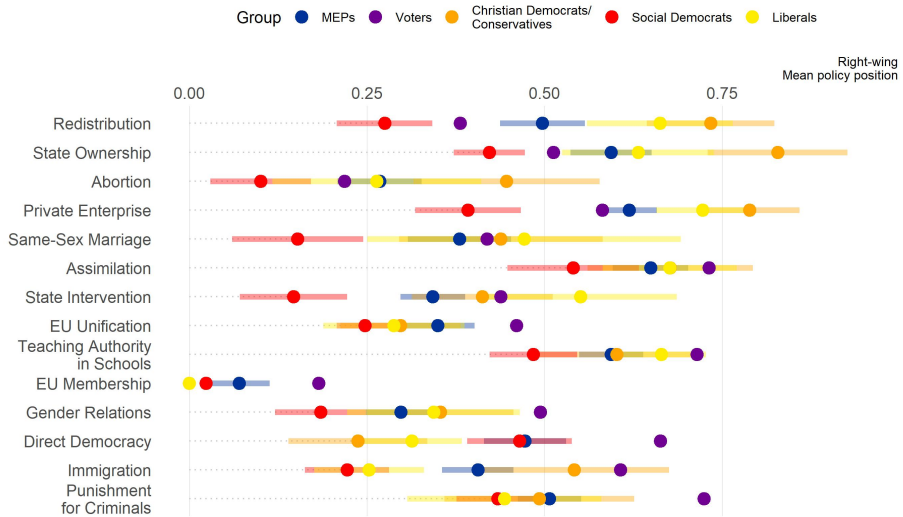
Figure 1.G.5. Mean Attitudes of Voters and MEPs of Several Party Families by Issue
Note: This figure shows a dumbbell plot. Different policy issues are shown on the vertical axis. The horizontal axis shows the mean attitudes of voters and politicians from various party families. All variables are scaled to range from zero to one and such that higher values indicate a position that is more right-wing.
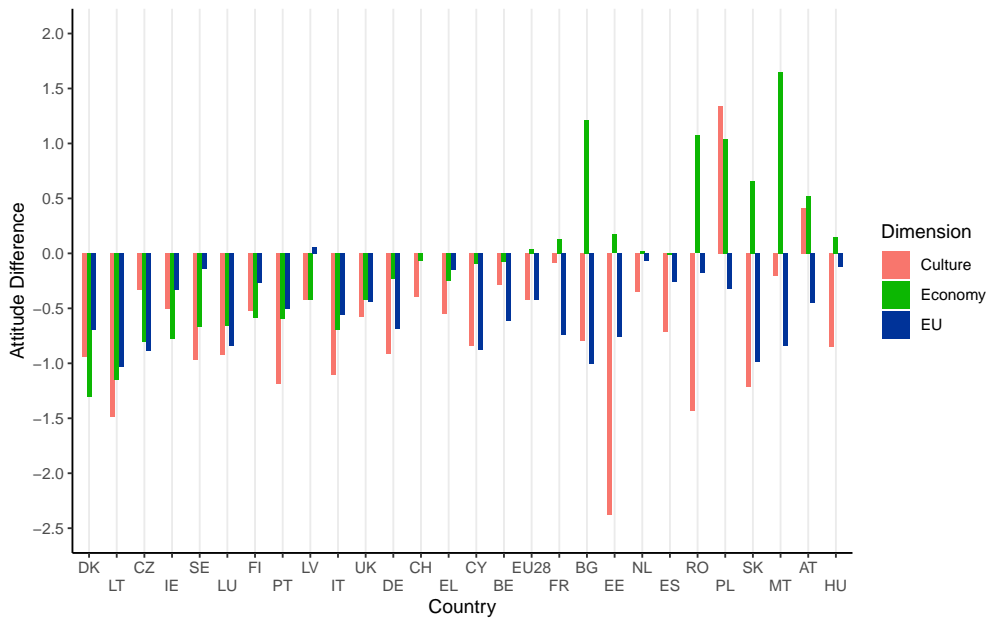
Figure 1.G.6. Attitudes Differences between Voters and Parliamentarians by Country
Note: Bars show attitude differences by country and policy dimension between voters and parliamentarians (MPs and MEPs) from the same country. The larger the value, the more right-wing parliamentarians are compared to voters from their country. Data for Slovenia is missing.
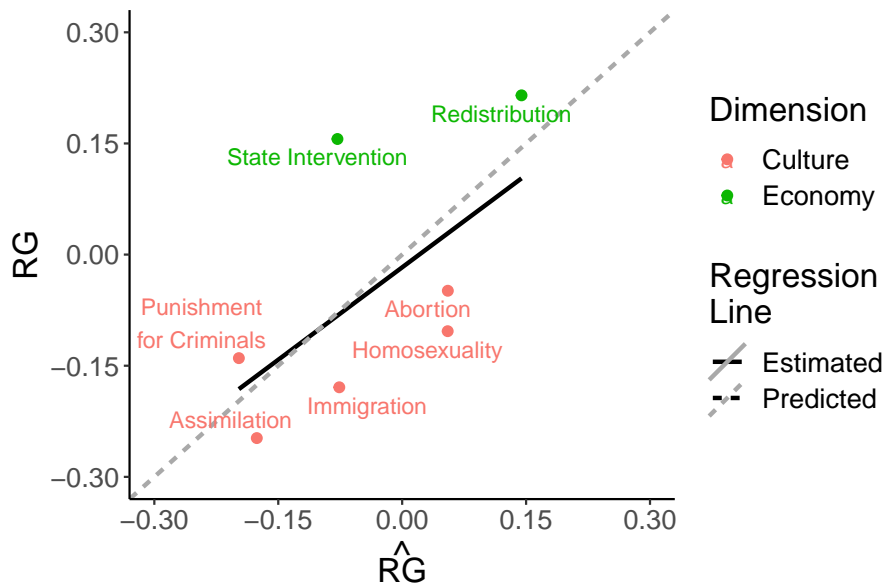
Figure 1.G.7. Estimates for Representation Gaps Based on Survey and Referendum Data (1992–2022)
Note: The horizontal axis shows RGs calculated from 2007 Swiss survey data. The vertical axis depicts RGs calculated from referendum votes of MPs and ordinary citizens. I take the unweighted means of RGs in all referendums belonging to a topic between 1992 and 2022. The figure also shows a dashed 45° line and a fitted line from a binary regression.
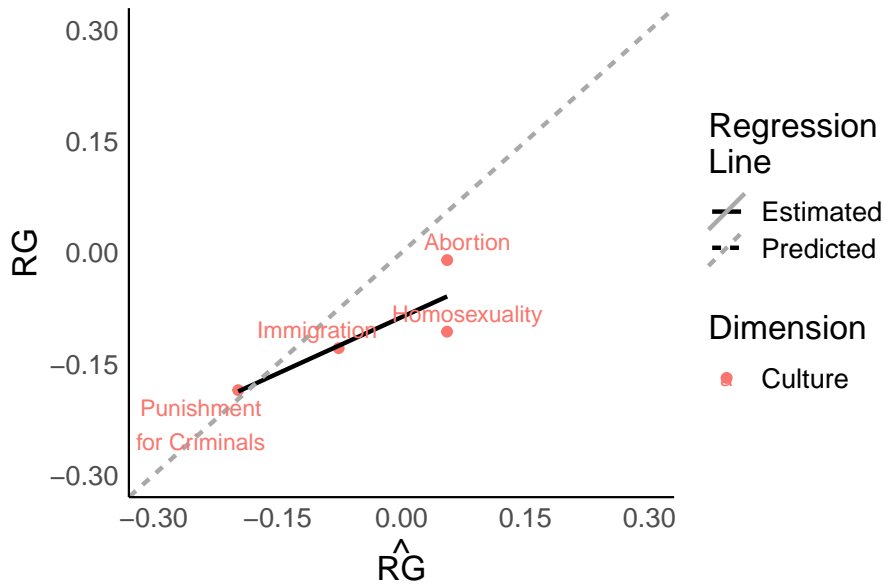
Figure 1.G.8. Estimates for Representation Gaps Based on Survey and Referendum Data (2002–2012)

Note: The horizontal axis shows RGs calculated from 2007 Swiss survey data. The vertical axis depicts RGs calculated from referendum votes of MPs and ordinary citizens. I take the unweighted means of RGs in all referendums belonging to a topic between 2002 and 2012. The figure also shows a dashed 45° line and a fitted line from a binary regression.
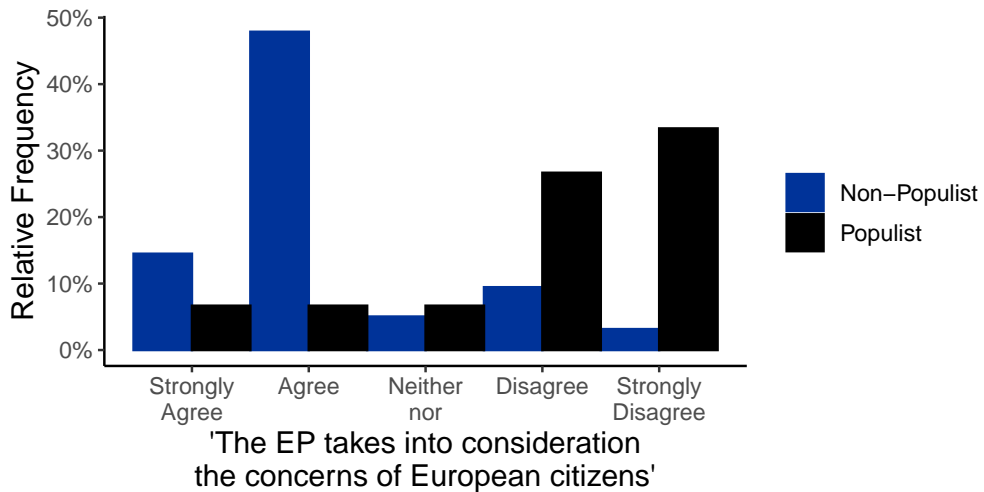
Figure 1.G.9. Assessment of Political Representation by Populist and Non-populist MEPs

Note: This figure depicts two histograms referring to answers of MEPs to the following statement: "The European Parliament takes into consideration the concerns of European citizens."
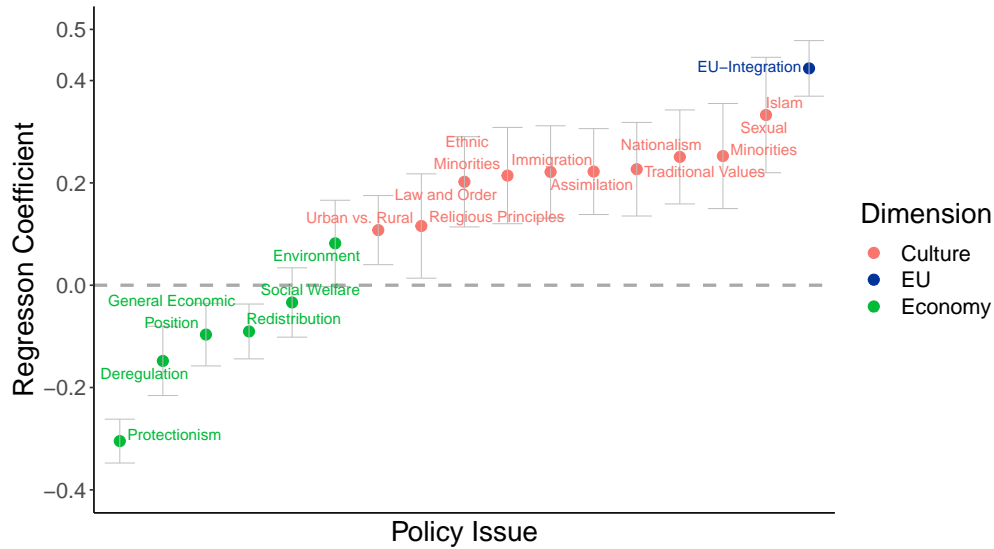
Figure 1.G.10. Issue-Positioning of Populist Parties Relative to Non-populist Parties Weighted With Vote Shares

Note: This figure is based on Equation 1.5. Each dot shows β_i for a different dependent variable/policy issue. These issues are arranged on the horizontal axis. Parties are weighted with their vote-share in the last national election before the observation. Higher values mean that populists are more right-wing than non-populist parties. Coefficients are surrounded by 95% confidence intervals.
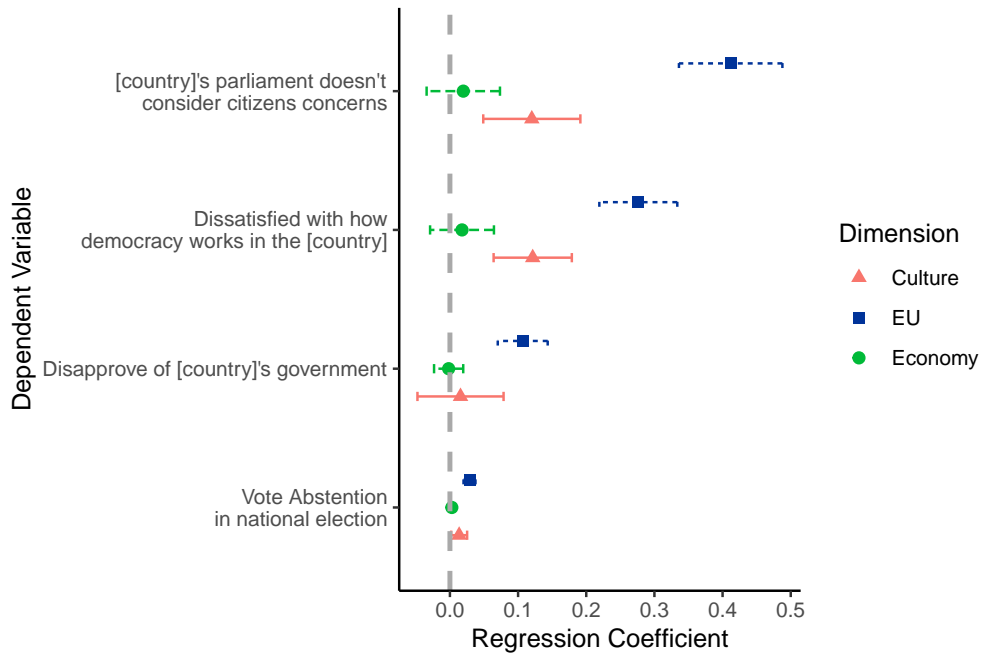
Figure 1.G.11. Association between Representation and Political Trust at the National Level

Note: This figure is based on Equation 1.5 and shows a coefficient plot. It depicts estimates for β_d from versions of Equation 1.4 with different dependent variables, which are arranged on the vertical axis. The horizontal axis shows the magnitude of coefficients. All regressions condition on a large set of demographic characteristics. Data is pooled for 15 EU countries. Standard errors are clustered at the country level.

Appendix 1.H   Additional Tables

Table 1.H.1. Information on Policy Attitude Variables

Variable Name	Question Wording	Question Type	Dimension	Included in
Assimilation	Immigrants should be required to adapt to the customs of [COUNTRY].	5 point Likert	Culture	EES/CCS
Private Enterprise	Private enterprise is the best way to solve [COUNTRY]'s economic problems.	5 point Likert	Economy	EES
Same-Sex Marriage	Same-sex marriages should be prohibited by law.	5 point Likert	Culture	EES/CCS
State Ownership	Major public services and industries ought to be in state ownership.	5 point Likert	Economy	EES
Abortion	Women should be free to decide on matters of abortion.	5 point Likert	Culture	EES/CCS
State Intervention	Politics should abstain from intervening in the economy.	5 point Likert	Economy	EES/CCS
Punishment for Criminals	People who break the law should be given much harsher sentences than they are these days.	5 point Likert	Culture	EES/CCS
Redistribution	Income and wealth should be redistributed towards ordinary people.	5 point Likert	Economy	EES/CCS
Teaching Authority in Schools	Schools must teach children to obey authority.	5 point Likert	Culture	EES
Direct Democracy	EU treaty changes should be decided by referendum.	5 point Likert	Culture	EES
Gender Relations	A woman should be prepared to cut down on her paid work for the sake of her family.	5 point Likert	Culture	EES
Immigration	Immigration to [COUNTRY] should be decreased significantly.	5 point Likert	Culture	EES
EU Unification	Some say European unification should be pushed further. Others say it already has gone too far. What is your opinion?	10 point from "has gone too far" to "should be pushed further"	EU	EES/CCS
EU Membership	Generally speaking, do you think that [COUNTRY]'s membership of the European Union is a good thing, a bad thing, or neither good nor bad?	3 Options: 1) "Good thing" 2) "Bad thing" 3) "Neither"	EU	EES/CCS

Note: Column one shows the wording for each policy attitude variable that I use in the paper. The wording is taken from the English version of the study. Questions were translated into the national language for other versions. [COUNTRY] is an placeholder for the name of the country the version of the survey was administered in. "Variable Name" refers to the names I use in the paper when referring to the items. Wording was identical in the EES and CCS surveys for all items with one exception. In the CCS the question for the "Punishment for Criminals" variable read as: "People who break the law should be given stiffer sentences."

Table 1.H.2. Using Valence to Predict Representation Gaps

	Dependent variable: Right-wing Representation Gap				
	(1)	(2)	(3)	(4)	(5)
Valence Advantage Right	0.614***	0.611***	0.611***	0.620***	0.604***
	(0.121)	(0.124)	(0.120)	(0.131)	(0.136)
Democracy Index		−0.012		0.082	0.090
		(0.083)		(0.354)	(0.356)
Control of Corruption			−0.015	0.280	0.281
			(0.094)	(0.390)	(0.397)
Constant	−0.423***	−0.374	−0.405*	0.440	0.156
	(0.141)	(0.386)	(0.219)	(0.768)	(0.755)
Other Controls				✓	✓
Dimension Dummies					✓
Observations	77	77	77	77	77
R^2	0.377	0.377	0.377	0.445	0.490

Note: This table shows results from OLS regressions. Standard errors (in parenthesis) are clustered at the country level. The dependent variable is positive if policymaking is more right-wing than preferred by voters. Valence Advantage Right is positive if right-wing parties are seen as more competent by voters than left-wing parties. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.3. Comparing Attitudes of Voters, MEPs and MEP-Candidates

	Dependent variable:					
	Cultural Index		Economic Index		EU Index	
	(1)	(2)	(3)	(4)	(5)	(6)
𝟙[MEP Candidate]	−0.118***	−0.070***	−0.022	−0.041**	−0.042	−0.005
	(0.018)	(0.012)	(0.018)	(0.017)	(0.028)	(0.024)
𝟙[Elected]	0.004	−0.018	0.057**	0.012	−0.069***	−0.028
	(0.023)	(0.026)	(0.027)	(0.031)	(0.027)	(0.020)
Demographic Controls		✓		✓		✓
Observations	17,311	15,599	17,705	15,928	18,864	16,928
R^2	0.063	0.404	0.003	0.207	0.007	0.233

Note: This table shows results from OLS regressions. All variables are standardized such that they range between zero and one and such that higher values indicate an attitude that is more right-wing. 𝟙[MEP Candidate] equals one if the respondent was a candidate for the 2009 EP election. 𝟙[Elected] equals one if he was elected to the European Parliament. Data includes responses from elected and unelected MEPs after elected MEPs officially started working as MEPs after the 2009 European Parliamentary election and those who voted at the 2009 European Parliament election. Standard errors (in parenthesis) are clustered at the country level. Regressions are weighted for population and country share of representatives in the European Parliament. Demographic controls include age, gender, categories for marital status, occupation, highest education degree, size of the town of residence, perceived own social class, immigration background, religious denomination, religiosity, living standard, and country. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.4. Comparing Attitudes of MEPs and MEP-Candidates

	Cultural Bias		Economic Bias		EU Bias	
	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$[Elected]	0.005	0.005	−0.020	−0.020	0.006	0.006
	(0.010)	(0.010)	(0.023)	(0.023)	(0.004)	(0.004)
Age	−0.0004***	−0.0004***	−0.0004	−0.0004	−0.0002	−0.0002
	(0.0002)	(0.0002)	(0.0004)	(0.0004)	(0.001)	(0.001)
Gender	−0.025*	−0.025*	0.012*	0.012*	0.013	0.013
	(0.015)	(0.015)	(0.007)	(0.007)	(0.025)	(0.025)
Only Inexperienced Candidates		✓		✓		✓
Observations	1,084	906	1,104	926	1,120	936
R^2	0.077	0.095	0.040	0.045	0.101	0.123

Note: This table shows results from OLS regressions. All variables are standardized such that they range between zero and one and such that higher values indicate an attitude that is more right-wing. Data includes responses from elected and unelected MEPs after elected MEPs officially started working as MEPs after the 2009 European Parliament election. Standard errors (in parenthesis) are clustered at the country level. Regressions are weighted for population and country share of representatives in the European Parliament. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.5. The Relationship between Campaign Donations and Bias of MEPs

	Cultural Bias		Economic Bias		EU Bias	
	Dependent variable:					
	(1)	(2)	(3)	(4)	(5)	(6)
Campaign Donations in Mill.	0.0005	0.0001	−0.001	−0.001	−0.0001	−0.0001
	(0.002)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)
Demographic Controls		✓		✓		✓
Observations	756	679	759	681	763	688
R^2	0.0001	0.401	0.0004	0.289	0.00000	0.317

Note: This table shows results from OLS regressions. All variables are standardized such that they range between zero and one and such that higher values indicate an attitude that is intuitively more right-wing. Data includes responses from elected and unelected MEPs. Data only includes data from countries where data is available for all three groups. Standard errors (in parenthesis) are clustered at the country level. Demographic controls include age, gender, categories for marital status, occupation, highest education degree, size of the town of residence, perceived own social class, immigration background, religious denomination, religiosity, living standard, and country. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.6. Comparing the Biases of MEPs Who Were and Were Not Encouraged to Run by Lobbyists

	Dependent variable:					
	Cultural Bias		Economic Bias		EU Bias	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$[encouraged by lobbyist]	0.008	0.005	0.012	0.007	−0.015	−0.020
	(0.021)	(0.018)	(0.012)	(0.010)	(0.015)	(0.014)
Demographic Controls		✓		✓		✓
Observations	1,266	1,129	1,288	1,144	1,304	1,158
R^2	0.0003	0.382	0.001	0.260	0.001	0.282

Note: This table shows results from OLS regressions. All variables are standardized such that they range between zero and one and such that higher values indicate an attitude that is more right-wing. Data includes responses from elected and unelected MEPs. Standard errors (in parenthesis) are clustered at the country level. Demographic controls include age, gender, categories for marital status, occupation, highest education degree, size of the town of residence, perceived own social class, immigration background, religious denomination, religiosity, living standard, and country. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.7. Relationship between Biases of MEPs and Frequency of Personal Visits at Firms and Clubs

	Dependent variable:					
	Cultural Bias		Economic Bias		EU Bias	
	(1)	(2)	(3)	(4)	(5)	(6)
Number firm visits	0.001	0.001**	0.001	0.001	−0.001	−0.001
of MEP	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Demographic Controls		✓		✓		✓
Observations	1,152	1,018	1,166	1,027	1,169	1,043
R^2	0.004	0.391	0.001	0.256	0.003	0.291

Note: This table shows results from OLS regressions. All variables are standardized such that they range between zero and one and such that higher values indicate an attitude that is more right-wing. Data includes responses from elected and unelected MEPs. Standard errors (in parenthesis) are clustered at the country level. Demographic controls include age, gender, categories for marital status, occupation, highest education degree, size of the town of residence, perceived own social class, immigration background, religious denomination, religiosity, living standard, and country. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.8. Relationship between Biases of MEPs and Frequency of Visits of Their Teams at Firms and Clubs

	Dependent variable:					
	Cultural Bias		Economic Bias		EU Bias	
	(1)	(2)	(3)	(4)	(5)	(6)
Number of firm visits by campaign team	−0.00002***	0.0004	−0.00001	0.001	−0.0001***	−0.001
	(0.00001)	(0.001)	(0.00001)	(0.001)	(0.00001)	(0.0005)
Demographic Controls		✓		✓		✓
Observations	873	766	879	769	878	781
R^2	0.0001	0.433	0.0001	0.310	0.001	0.324

*Note:*This table shows results from OLS regressions. All variables are standardized such that they range between zero and one and such that higher values indicate an attitude that is more right-wing. Data includes responses from elected and unelected MEPs. Standard errors (in parenthesis) are clustered at the country level. Demographic controls include age, gender, categories for marital status, occupation, highest education degree, size of the town of residence, perceived own social class, immigration background, religious denomination, religiosity, living standard, and country. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.9. Comparing the Bias of MEPs Who Used to Be Lobbyists with That of Other MEPs

	Dependent variable:					
	Cultural Bias		Economic Bias		EU Bias	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$[MPE was a lobbyist]	0.006	−0.0004	−0.034**	−0.025	−0.004	0.014
	(0.022)	(0.022)	(0.015)	(0.022)	(0.017)	(0.020)
Demographic Controls		✓		✓		✓
Observations	1,086	981	1,101	993	1,116	1,005
R^2	0.0001	0.390	0.004	0.284	0.00003	0.297

Note: This table shows results from OLS regressions. All variables are standardized such that they range between zero and one and such that higher values indicate an attitude that is more right-wing. Data includes responses from elected and unelected MEPs. Standard errors (in parenthesis) are clustered at the country level. Demographic controls include age, gender, categories for marital status, occupation, highest education degree, size of the town of residence, perceived own social class, immigration background, religious denomination, religiosity, living standard, and country. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.10. Comparing Biases of Young MEPs That Plan and Not Plan to Leave Politics Soon

	Dependent variable:					
	Cultural Bias		Economic Bias		EU Bias	
	(1)	(2)	(3)	(4)	(5)	(6)
𝟙[MEP plans to leave politics]	0.015	0.013	−0.031	−0.030*	0.008	0.004
	(0.017)	(0.017)	(0.019)	(0.017)	(0.021)	(0.026)
Demographic Controls		✓		✓		✓
Observations	649	589	664	601	672	605
R^2	0.001	0.488	0.003	0.333	0.0002	0.319

Note: This table shows results from OLS regressions. All variables are standardized such that they range between zero and one and such that higher values indicate an attitude that is more right-wing. Data includes responses from elected and unelected MEPs. Standard errors (in parenthesis) are clustered at the country level. Demographic controls include age, gender, categories for marital status, occupation, highest education degree, size of the town of residence, perceived own social class, immigration background, religious denomination, religiosity, living standard, and country. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1.H.11. List of Populist Parties

Partyname	Abbreviation	Country
Vlaams Belang	VB	Belgium
Dutch Labour Party	PVDA	Netherlands
People's Movement Against the EU	FolkB	Denmark
National Democratic Party of Germany	NPD	Germany
Alternative for Germany	AfD	Germany
Human Environment Animal Protection	DieTier	Germany
Coalition of the Radical Left	SYRIZA	Greece
Communist Party of Greece	KKE	Greece
Popular Orthodox Rally	LAOS	Greece
Independent Greeks	ANEL	Greece
Popular Association—Golden Dawn	XA	Greece
We Can	Podemos	Spain
National Rally (formerly Front National)	FN	France
Left Party	PG	France
We Ourselves	SF	Ireland
Socialist Party	SP	Ireland
People Before Profit Alliance	PBPA	Ireland
Communist Refoundation Party	RC	Italy
Northern League	LN	Italy
Five Star Movement	M5S	Italy
Party for Freedom	PVV	Netherlands
Green Party	GREEN	United Kingdom
UK Independency Party	UKIP	United Kingdom
Earth Party	MPT	Portugal
Freedom Party of Austria	FPO	Austria
Team Stronach for Austria	TeamStronach	Austria
True Finns	PS	Finland
Sweden Democrats	SD	Sweden
Pirate Party	PIRAT	Sweden
VMRO—Bulgarian National Movement	VMRO-BND	Bulgaria
Attack	ATAKA	Bulgaria
National Front for the Salvation of Bulgaria	NFSB	Bulgaria
Bulgaria without Censorship	BBT	Bulgaria
ANO 2011, Action of Dissatisfied Citizens	ANO2011	Czech Republic
Dawn of Direct Democracy	USVIT	Czech Republic
Estonian Free Party	EVE	Estonia
Jobbik—Movement for a Better Hungary	JOBBIK	Hungary
Latvian Russian Union	LKS	Latvia
For Latvia from the Heart	NSL	Latvia
Latvian Association of Regions	LRA	Latvia
The Way of Courage	DK	Lithuania
Congress of the New Right	KNP	Poland
United Poland	SP	Poland

Continued on next page

Table 1.H.11 – *Continued from previous page*

Partyname	Abbreviation	Country
People's Party—Dan Diaconescu	PP-DD	Romania
Ordinary People and Independent Personalities	OLaNO	Slovakia
Croatian Labourists–Labour Party	HL-SR	Croatia
Croatian Party of Rights dr. Ante Starcevic	HSP-AS	Croatia
Alternative Democratic Reform Party	ADR	Luxembourg
The Left	DL	Luxembourg

References

Adams, James. 1999. "Multiparty Spatial Competition with Probabilistic Voting." en. *Public Choice* 99 (3): 259–74. DOI: 10.1023/A:1018362719595. (Visited on 04/12/2023). [7, 13]

Ajzenman, Nicolás, Cevat Giray Aksoy, and Sergei Guriev. 2022. "Exposure to Transit Migration: Public Attitudes and Entrepreneurship." en. *Journal of Development Economics* 158 (9): 102899. DOI: 10.1016/j.jdeveco.2022.102899. (Visited on 04/17/2023). [8]

Albanese, Giuseppe, Guglielmo Barone, and Guido de Blasio. 2022. "Populist Voting and Losers' Discontent: Does Redistribution Matter?" en. *European Economic Review* 141 (1): 104000. DOI: 10.1016/j.euroecorev.2021.104000. (Visited on 04/12/2023). [8]

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva. 2023. "Immigration and Redistribution." *Review of Economic Studies* 90 (1): 1–39. DOI: 10.1093/restud/rdac011. (Visited on 03/27/2023). [46, 48]

Algan, Yann, Sergei Guriev, Elias Papaioannou, and Evgenia Passari. 2017. "The European Trust Crisis and the Rise of Populism." *Brookings Papers on Economic Activity* 2017 (2): 309–400. DOI: 10.1353/eca.2017.0015. (Visited on 04/12/2023). Publisher: Johns Hopkins University Press. [8]

Andeweg, Rudy B. 2012. "3. Approaching Perfect Policy Congruence: Measurement, Development, and Relevance for Political Representation." In *How Democracy Works*. Amsterdam University Press, 39–52. [5, 10]

Andreadis, Ioannis, and Yannis Stavrakakis. 2017. "European Populist Parties in Government: How Well are Voters Represented? Evidence from Greece." en. *Swiss Political Science Review* 23 (4): 485–508. DOI: 10.1111/spsr.12255. (Visited on 03/19/2023). [10]

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi. 2020. "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure." en. *American Economic Review* 110 (10): 3139–83. DOI: 10.1257/aer.20170011. (Visited on 03/19/2023). [8]

Bakker, Ryan, Seth Jolly, and Jonathan Polk. 2012. "Complexity in the European Party Space: Exploring Dimensionality with Experts." en. *European Union Politics* 13 (2): 219–45. DOI: 10.1177/1465116512436995. (Visited on 03/19/2023). [13, 18]

Bakker, Ryan, Seth Jolly, and Jonathan Polk. 2018. "Multidimensional Incongruence and Vote Switching in Europe." en. *Public Choice* 176 (1-2): 267–96. DOI: 10.1007/s11127-018-0555-z. (Visited on 03/19/2023). [8]

Bakker, Ryan, Seth Jolly, and Jonathan Polk. 2020. "Multidimensional Incongruence, Political Disaffection, and Support for Anti-establishment Parties." en. *Journal of European Public Policy* 27 (2): 292–309. DOI: 10.1080/13501763.2019.1701534. (Visited on 03/19/2023). [36]

Banerjee, Abhijit V., and Rohini Pande. 2007. "Parochial Politics: Ethnic Preferences and Politician Corruption." en. SSRN Scholarly Paper. Rochester, NY. URL: https://papers.ssrn.com/abstract=1136706 (visited on 04/12/2023). [10]

Banks, Jeffrey S. 1995. "Singularity Theory and Core Existence in the Spatial Model." en. *Journal of Mathematical Economics* 24 (6): 523–36. DOI: 10.1016/0304-4068(94)00704-E. (Visited on 03/19/2023). [13]

Barone, Guglielmo, Alessio D'Ignazio, Guido de Blasio, and Paolo Naticchioni. 2016. "Mr. Rossi, Mr. Hu and Politics. The Role of Immigration in Shaping Natives' Voting Behavior." en. *Journal of Public Economics* 136 (4): 1–13. DOI: 10.1016/j.jpubeco.2016.03.002. (Visited on 03/19/2023). [8]

Barrera, Oscar, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya. 2020. "Facts, Alternative Facts, and Fact Checking in Times of Post-truth Politics." en. *Journal of Public Economics* 182 (2): 104123. DOI: 10.1016/j.jpubeco.2019.104123. (Visited on 03/19/2023). [45, 46]

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *Quarterly Journal of Economics* 124 (4): 1497–540. DOI: 10.1162/qjec.2009.124.4.1497. (Visited on 04/12/2023). [10]

Beauvallet, Willy, Victor Lepaux, and Sébastien Michon. 2013. "Who Are The MEPs? A Statistical Analysis Of The Backgrounds Of Members Of European Parliament (2004-2014) And Of Their Transformations." en. *Etudes européennes. La revue permanente des professionnels de l'Europe*, URL: https://hal.science/hal-02060086 (visited on 04/17/2023). [48]

Becker, Sascha O., and Thiemo Fetzer. 2016. "Does Migration Cause Extreme Voting." *Center for Competitive Advantage in the Global Economy and The Economic & Social Research Council*, 1–54. [8]

Bellodi, Luca, Massimo Morelli, and Matia Vannoni. 2021. "A Costly Commitment: Populism, Government Performance, and the Quality of Bureaucracy." en. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.3982033. (Visited on 04/12/2023). [9]

Benedetto, Giacomo, Simon Hix, and Nicola Mastrorocco. 2020. "The Rise and Fall of Social Democracy, 1918–2017." en. *American Political Science Review* 114 (3): 928–39. DOI: 10.1017/S0003055420000234. (Visited on 03/19/2023). [54]

Berger, Suzanne. 2017. "Populism and the Failures of Representation." en. *French Politics, Culture & Society* 35 (2): DOI: 10.3167/fpcs.2017.350203. (Visited on 03/19/2023). [8]

Berman, Sheri. 2021. "The Causes of Populism in the West." en. *Annual Review of Political Science* 24 (1): 71–88. DOI: 10.1146/annurev-polisci-041719-102503. (Visited on 03/19/2023). [8]

Berman, Sheri, and Hans Kundnani. 2021. "The Cost of Convergence." en. *Journal of Democracy* 32 (1): 22–36. DOI: 10.1353/jod.2021.0013. (Visited on 03/19/2023). [8]

Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne. 2017. "Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden." en. *American Economic Review* 107 (8): 2204–42. DOI: 10.1257/aer.20160080. (Visited on 03/19/2023). [10]

Beyer, Daniela, and Miriam Hänni. 2018. "Two Sides of the Same Coin? Congruence and Responsiveness as Representative Democracy's Currencies: Democracy's Currencies." en. *Policy Studies Journal* 46 (5): S13–S47. DOI: 10.1111/psj.12251. (Visited on 03/19/2023). [47]

Black, Duncan. 1948. "On the Rationale of Group Decision-making." en. *Journal of Political Economy* 56 (1): 23–34. DOI: 10.1086/256633. (Visited on 03/19/2023). [13, 30]

Blanes i Vidal, Jordi, Mirko Draca, and Christian Fons-Rosen. 2012. "Revolving Door Lobbyists." en. *American Economic Review* 102 (7): 3731–48. DOI: 10.1257/aer.102.7.3731. (Visited on 03/19/2023). [66]

Bó, Ernesto Dal, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne. 2017. "Who Becomes A Politician?" *Quarterly Journal of Economics* 132 (4): 1877–914. DOI: 10.1093/qje/qjx016. (Visited on 04/12/2023). [10]

Bó, Ernesto Dal, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne. 2023. "Economic and Social Outsiders but Political Insiders: Sweden's Populist Radical Right." en. *Review of Economic Studies* 90 (2): 675–706. DOI: 10.1093/restud/rdac037. (Visited on 03/19/2023). [8, 10, 39]

Bombardini, Matilde, and Francesco Trebbi. 2020. "Empirical Models of Lobbying." *Annual Review of Economics* 12 (1): 391–413. DOI: 10.1146/annurev-economics-082019-024350. (Visited on 04/14/2023). _eprint: https://doi.org/10.1146/annurev-economics-082019-024350. [66]

Bonomi, Giampaolo, Nicola Gennaioli, and Guido Tabellini. 2021. "Identity, Beliefs, and Political Conflict." *Quarterly Journal of Economics* 136 (4): 2371–411. DOI: 10.1093/qje/qjab034. (Visited on 04/12/2023). [9]

Bühlmann, Marc, Antoinette Feh Widmer, and Lisa Schädel. 2010. "Substantive and Descriptive Representation in Swiss Cantons." en. *Swiss Political Science Review* 16 (3): 565–95. DOI: 10.1002/j.1662-6370.2010.tb00441.x. (Visited on 03/19/2023). [10]

Calvert, Randall L. 1985. "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence." en. *American Journal of Political Science* 29 (1): 69. DOI: 10.2307/2111212. (Visited on 03/19/2023). [30]

Caplin, Andrew, and Barry Nalebuff. 1991. "Aggregation and Social Choice: A Mean Voter Theorem." en. *Econometrica* 59 (1): 1. DOI: 10.2307/2938238. (Visited on 03/19/2023). [13]

Cappelen, Alexander W., Ranveig Falch, and Bertil Tungodden. 2020. "Fair and Unfair Income Inequality." en. In *Handbook of Labor, Human Resources and Population Economics*. Edited by Klaus F. Zimmermann. Cham: Springer International Publishing, 1–25. DOI: 10.1007/978-3-319-57365-6_130-1. (Visited on 03/19/2023). [47]

CCS. 2016. "Comparative Candidates Survey Module I - 2005-2013 [Dataset - cumulative file]." Edited by FORS. [14]

Chattopadhyay, Raghabendra, and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." en. *Econometrica* 72 (5): 1409–43. DOI: 10.1111/j.1468-0262.2004.00539.x. (Visited on 03/19/2023). [10]

Cinelli, Carlos, Andrew Forney, and Judea Pearl. 2020. "A Crash Course in Good and Bad Controls." en. *SSRN Electronic Journal*, DOI: 10.2139/ssrn.3689437. (Visited on 03/19/2023). [67]

Cohen, L., and S. Matthews. 1980. "Constrained Plott Equilibria, Directional Equilibria and Global Cycling Sets." en. *Review of Economic Studies* 47 (5): 975–86. DOI: 10.2307/2296927. (Visited on 03/19/2023). [13]

Colantone, Italo, and Piero Stanig. 2018a. "Global Competition and Brexit." en. *American Political Science Review* 112 (2): 201–18. DOI: 10.1017/S0003055417000685. (Visited on 03/19/2023). [8]

Colantone, Italo, and Piero Stanig. 2018b. "The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe." en. *American Journal of Political Science* 62 (4): 936–53. DOI: 10.1111/ajps.12358. (Visited on 04/12/2023). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12358. [8]

Costello, Rory, Jacques Thomassen, and Martin Rosema. 2012. "European Parliament Elections and Political Representation: Policy Congruence between Voters and Parties." en. *West European Politics* 35 (6): 1226–48. DOI: 10.1080/01402382.2012.713744. (Visited on 03/19/2023). [10]

Costello, Rory, Dimiter Toshkov, Barend Bos, and André Krouwel. 2021. "Congruence between Voters and Parties: The Role of Party-Level Issue Salience." en. *European Journal of Political Research* 60 (1): 92–113. DOI: 10.1111/1475-6765.12388. (Visited on 03/19/2023). [10]

Coughlin, Peter, and Shmuel Nitzan. 1981. "Electoral Outcomes With Probabilistic Voting and Nash Social Welfare Maxima." en. *Journal of Public Economics* 15 (1): 113–21. DOI: 10.1016/0047-2727(81)90056-6. (Visited on 03/19/2023). [13]

Dalton, Russell J. 2017. "Party Representation across Multiple Issue Dimensions." en. *Party Politics* 23 (6): 609–22. DOI: 10.1177/1354068815614515. (Visited on 03/19/2023). [10]

Danieli, Oren, Noam Gidron, Shinnosuke Kikuchi, and Ro'ee Levy. 2022. "Decomposing the Rise of the Populist Radical Right." en. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.4255937. (Visited on 04/12/2023). [9]

Dinas, Elias, Konstantinos Matakos, Dimitrios Xefteris, and Dominik Hangartner. 2019. "Waking Up the Golden Dawn: Does Exposure to the Refugee Crisis Increase Support for Extreme-Right Parties?" en. *Political Analysis* 27 (2): 244–54. DOI: 10.1017/pan.2018.48. (Visited on 04/12/2023). Publisher: Cambridge University Press. [8]

Docquier, Frédéric, Eugenio Peluso, and Massimo Morelli. 2022. "Democracies in Danger: How Can We Break the Vicious Circle of Populism?" en. [9, 45]

Downs, Anthony. 1957. "An Economic Theory of Political Action in a Democracy." *Journal of Political Economy* 65 (2): 135–50. DOI: 10.1086/257897. (Visited on 03/31/2023). Publisher: The University of Chicago Press. [7, 9, 12, 13, 29, 30, 36]

Drago, Francesco, Roberto Galbiati, and Francesco Sobbrio. 2020. "The Political Cost of Being Soft on Crime: Evidence from a Natural Experiment." en. *Journal of the European Economic Association* 18 (6): 3305–36. DOI: 10.1093/jeea/jvz063. (Visited on 03/19/2023). [8]

Duflo, Esther. 2012. "Women Empowerment and Economic Development." en. *Journal of Economic Literature* 50 (4): 1051–79. DOI: 10.1257/jel.50.4.1051. (Visited on 03/19/2023). [10]

Dustmann, Christian, Kristine Vasiljeva, and Anna Piil Damm. 2019. "Refugee Migration and Electoral Outcomes." en. *Review of Economic Studies* 86 (5): 2035–91. DOI: 10.1093/restud/rdy047. (Visited on 03/19/2023). [8]

Eaton, B. Curtis, and Richard G. Lipsey. 1975. "The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition." en. *Review of Economic Studies* 42 (1): 27. DOI: 10.2307/2296817. (Visited on 03/19/2023). [13]

Edo, Anthony, Yvonne Giesing, Jonathan Öztunc, and Panu Poutvaara. 2019. "Immigration and Electoral Support for the Far-Left and the Far-Right." en. *European Economic Review* 115 (6): 99–143. DOI: 10.1016/j.euroecorev.2019.03.001. (Visited on 03/19/2023). [8]

Egmond, Marcel van, Wouter van der Brug, Sara Hobolt, Mark Franklin, and Eliyahu V. Sapir. 2017. "European Parliament Election Study 2009, Voter Study." DOI: 10.4232/1.12732. Published: GESIS Datenarchiv, Köln. ZA5055 Datenfile Version 1.1.1, https://doi.org/10.4232/1.12732. [14]

Enke, Benjamin. 2020. "Moral Values and Voting." *Journal of Political Economy* 128 (10): 3679–729. DOI: 10.1086/708857. (Visited on 04/14/2023). Publisher: The University of Chicago Press. [48]

Enke, Benjamin, Ricardo Rodríguez-Padilla, and Florian Zimmermann. 2022. "Moral Universalism: Measurement and Economic Relevance." *Management Science* 68 (5): 3590–603. DOI: 10.1287/mnsc.2021.4086. (Visited on 03/31/2023). Publisher: INFORMS. [48]

Esaiasson, Peter, Jacob Sohlberg, Marina Ghersetti, and Bengt Johansson. 2021. "How the Coronavirus Crisis Affects Citizen Trust in Institutions and in Unknown Others: Evidence from 'the Swedish Experiment'." en. *European Journal of Political Research* 60 (3): 748–60. DOI: 10.1111/1475-6765.12419. (Visited on 03/19/2023). [8]

Evans, Georgina, and Peter A. Hall. 2019. "Representation Gaps: Changes in Popular Preferences and Party Positions over the Long Term in the Developed Democracies." In *Annual Meeting of the American Political Science Association, August 2019*. Washington DC. [10]

Facchini, Giovanni, Yotam Margalit, and Hiroyuki Nakata. 2022. "Countering Public Opposition to Immigration: The Impact of Information Campaigns." en. *European Economic Review* 141 (1): 103959. DOI: 10.1016/j.euroecorev.2021.103959. (Visited on 04/02/2023). [48]

Ferland, Benjamin. 2016. "Revisiting the Ideological Congruence Controversy." en. *European Journal of Political Research* 55 (2): 358–73. DOI: 10.1111/1475-6765.12127. (Visited on 04/12/2023). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-6765.12127. [10]

Fetzer, Thiemo. 2019. "Did Austerity Cause Brexit?" en. *American Economic Review* 109 (11): 3849–86. DOI: 10.1257/aer.20181164. (Visited on 03/19/2023). [8]

Funke, Manuel, Moritz Schularick, and Christoph Trebesch. 2016. "Going to Extremes: Politics After Financial Crises, 1870–2014." en. *European Economic Review* 88 (9): 227–60. DOI: 10.1016/j.euroecorev.2016.03.006. (Visited on 03/19/2023). [8, 9]

Funke, Manuel, Moritz Schularick, and Christoph Trebesch. 2021. "The Cost of Populism: Evidence from History." eng. Research Report 015. ECONtribute Policy Brief. URL: https://www.econstor.eu/handle/10419/268568 (visited on 03/25/2023). [45]

Gabriel, Ricardo Duque, Mathias Klein, and Ana Sofia Pessoa. 2022. "The Political Costs of Austerity." en. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.4160971. (Visited on 04/12/2023). [8]

Gethin, Amory, Clara Martínez-Toledano, and Thomas Piketty. 2021. "Brahmin Left Versus Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948–2020." en. *Quarterly Journal of Economics* 137 (1): 1–48. DOI: 10.1093/qje/qjab036. (Visited on 03/19/2023). [8]

Golder, Matt, and Benjamin Ferland. 2017. "Electoral Systems and Citizen-Elite Ideological Congruence." In *The Oxford Handbook of Electoral Systems*. Edited by Erik S. Herron, Robert J. Pekkanen, and Matthew S. Shugart. Oxford University Press, 0. DOI: 10.1093/oxfordhb/9780190258658.013.11. (Visited on 04/12/2023). [10]

Golder, Matt, and Jacek Stramski. 2010. "Ideological Congruence and Electoral Institutions." en. *American Journal of Political Science* 54 (1): 90–106. DOI: 10.1111/j.1540-5907.2009.00420.x. (Visited on 03/19/2023). [5]

Grigorieff, Alexis, Christopher Roth, and Diego Ubfal. 2020. "Does Information Change Attitudes Toward Immigrants?" en. *Demography* 57 (3): 1117–43. DOI: 10.1007/s13524-020-00882-8. (Visited on 03/19/2023). [46, 48]

Grossman, Gene M., and Elhanan Helpman. 1996. "Electoral Competition and Special Interest Politics." *Review of Economic Studies* 63 (2): 265–86. DOI: 10.2307/2297852. (Visited on 04/14/2023). [66]

Grzymala-Busse, Anna. 2019. "The Failure of Europe's Mainstream Parties." en. *Journal of Democracy* 30 (4): 35–47. DOI: 10.1353/jod.2019.0067. (Visited on 03/19/2023). [8]

Günther, Laurenz. 2023. "Would Europeans Accept Immigrants If They Knew Them?" en. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.4316180. (Visited on 04/13/2023). [46]

Guriev, Sergei, and Elias Papaioannou. 2022. "The Political Economy of Populism." en. *Journal of Economic Literature* 60 (3): 753–832. DOI: 10.1257/jel.20201595. (Visited on 03/25/2023). [8, 37, 45]

Haidt, Jonathan, and Greg Lukianoff. 2018. *The Coddling of the American Mind: How Good Intentions and Bad Ideas Are Setting Up a Generation for Failure*. en. Penguin UK. Google-Books-ID: SV5HDwAAQBAJ. [46]

Hakhverdian, Armen, and Wouter Schakel. 2022. "The Political Representation of Left-Nationalist Voters." en. *Acta Politica* 57 (3): 489–509. DOI: 10.1057/s41269-021-00205-8. (Visited on 03/19/2023). [10]

Halla, Martin, Alexander F Wagner, and Josef Zweimüller. 2017. "Immigration and Voting for the Far Right." en. *Journal of the European Economic Association* 15 (6): 1341–85. DOI: 10.1093/jeea/jvx003. (Visited on 03/19/2023). [8]

Hangartner, Dominik, Elias Dinas, Moritz Marbach, Konstantinos Matakos, and Dimitrios Xefteris. 2019. "Does Exposure to the Refugee Crisis Make Natives More Hostile?" en. *Ameri-*

can Political Science Review 113 (2): 442–55. DOI: 10.1017/S0003055418000813. (Visited on 03/19/2023). [8]

Harmon, Nikolaj A. 2018. "Immigration, Ethnic Diversity, and Political Outcomes: Evidence from Denmark." en. *Scandinavian Journal of Economics* 120 (4): 1043–74. DOI: 10.1111/sjoe.12239. (Visited on 03/19/2023). [8]

Heckman, James J., Tomáš Jagelka, and Tim Kautz. 2021. "Some Contributions of Economics to the Study of Personality." In *Handbook of Personality: Theory and Research*. 4th edition. New York, NY, US: The Guilford Press, 853–92. [26]

Heß, Moritz, Christian von Scheve, Jürgen Schupp, Aiko Wagner, and Gert G. Wagner. 2018. "Are Political Representatives More Risk-Loving than the Electorate? Evidence from German Federal and State Parliaments." en. *Palgrave Communications* 4 (1): 1–7. DOI: 10.1057/s41599-018-0112-x. (Visited on 03/19/2023). [46–48]

Hinich, Melvin J, John O Ledyard, and Peter C Ordeshook. 1972. "Nonvoting and the Existence of Equilibrium under Majority Rule." en. *Journal of Economic Theory* 4 (2): 144–53. DOI: 10.1016/0022-0531(72)90145-7. (Visited on 03/19/2023). [13]

Hix, Simon, and Abdul Noury. 2009. "After Enlargement: Voting Patterns in the Sixth European Parliament." en. *Legislative Studies Quarterly* 34 (2): 159–74. DOI: 10.3162/036298009788314282. (Visited on 03/19/2023). [70]

Holmberg, Sören. 2012. "4. Dynamic Representation from Above." In *How Democracy Works*. Amsterdam University Press, 53–76. [10]

Hopkins, Daniel J., John Sides, and Jack Citrin. 2019. "The Muted Consequences of Correct Information about Immigration." en. *Journal of Politics* 81 (1): 315–20. DOI: 10.1086/699914. (Visited on 03/19/2023). [46]

Inglehart, Ronald. 1971. "The Silent Revolution in Europe: Intergenerational Change in Post-Industrial Societies." en. *American Political Science Review* 65 (4): 991–1017. DOI: 10.2307/1953494. (Visited on 05/08/2023). Publisher: Cambridge University Press. [48]

Inglehart, Ronald. 2015. *The Silent Revolution: Changing Values and Political Styles Among Western Publics*. en. Princeton University Press. Google-Books-ID: wX59BgAAQBAJ. [48]

Jaime-Castillo, Antonio M., and Xavier Coller. 2022. "Congruence between Voters and Representatives in Preferences for Social Policies in Spain." *Journal of Legislative Studies* 0 (0): 1–21. DOI: 10.1080/13572334.2022.2146864. (Visited on 04/14/2023). Publisher: Routledge _eprint: https://doi.org/10.1080/13572334.2022.2146864. [10]

Jolly, Seth, Ryan Bakker, Liesbet Hooghe, Gary Marks, Jonathan Polk, Jan Rovny, Marco Steenbergen, and Milada Anna Vachudova. 2022. "Chapel Hill Expert Survey Trend File, 1999–2019." en. *Electoral Studies* 75 (2): 102420. DOI: 10.1016/j.electstud.2021.102420. (Visited on 04/14/2023). [15, 49]

Kertzer, Joshua D. 2022. "Re-Assessing Elite-Public Gaps in Political Behavior." en. *American Journal of Political Science* 66 (3): 539–53. DOI: 10.1111/ajps.12583. (Visited on 03/19/2023). [10, 12, 67]

Kritzinger, Sylvia, Martial Foucault, Romain Lachat, Julia Partheymüller, Carolina Plescia, and Sylvain Brouard. 2021. "'Rally round the flag': the COVID-19 crisis and trust in the national government." en. *West European Politics* 44 (5-6): 1205–31. DOI: 10.1080/01402382.2021.1925017. (Visited on 03/19/2023). [8]

Kübler, Marie, and Armin Schäfer. 2022. "Closing the Gap? The Populist Radical Right and Opinion Congruence between Citizens and MPs." en. *Electoral Studies* 80 (12): 102527. DOI: 10.1016/j.electstud.2022.102527. (Visited on 04/02/2023). [10]

Kustov, Alexander, Dillon Laaker, and Cassidy Reller. 2021. "The Stability of Immigration Attitudes: Evidence and Implications." *Journal of Politics* 83 (4): 1478–94. DOI: 10.1086/715061. (Visited on 03/27/2023). Publisher: The University of Chicago Press. [46]

Lesschaeve, Christophe. 2022. "Bringing Representatives Back In: How Political Parties Moderate Patterns of Inequality in Opinion Representation." en. *European Political Science* 21 (2): 255–73. DOI: 10.1057/s41304-021-00332-y. (Visited on 03/19/2023). [10]

Lin, Tse-min, James M. Enelow, and Han Dorussen. 1999. "Equilibrium in Multicandidate Probabilistic Spatial Voting." en. *Public Choice* 98 (1): 59–82. DOI: 10.1023/A:1018391410733. (Visited on 04/12/2023). [7, 13]

Lindner, Attila, Filip Novokmet, Thomas Piketty, and Tomasz Zawisza. 2020. "Political Conflict, Social Inequality and Electoral Cleavages in Central-Eastern Europe, 1990-2018." en. URL: https://shs.hal.science/halshs-03047677 (visited on 03/31/2023). [8]

Mayne, Quinton, and Armen Hakhverdian. 2017. "Ideological Congruence and Citizen Satisfaction: Evidence From 25 Advanced Democracies." en. *Comparative Political Studies* 50 (6): 822–49. DOI: 10.1177/0010414016639708. (Visited on 03/19/2023). [47]

McKelvey, Richard D. 1979. "General Conditions for Global Intransitivities in Formal Voting Models." en. *Econometrica* 47 (5): 1085. DOI: 10.2307/1911951. (Visited on 03/19/2023). [13, 21]

McKelvey, Richard D., and John W. Patty. 2006. "A Theory of Voting in Large Elections." en. *Games and Economic Behavior* 57 (1): 155–80. DOI: 10.1016/j.geb.2006.05.003. (Visited on 03/19/2023). [13]

McKelvey, Richard D., and Norman Schofield. 1987. "Generalized Symmetry Conditions at a Core Point." en. *Econometrica* 55 (4): 923. DOI: 10.2307/1911036. (Visited on 03/19/2023). [13, 21]

McKelvey, Richard D., and Richard E. Wendell. 1976. "Voting Equilibria in Multidimensional Choice Spaces." en. *Mathematics of Operations Research* 1 (2): 144–58. DOI: 10.1287/moor.1.2.144. (Visited on 03/19/2023). [13, 21]

Morelli, Massimo, Antonio Nicolò, and Paolo Roberti. 2021. "A Commitment Theory of Populism." en. SSRN Scholarly Paper. Rochester, NY. URL: https://papers.ssrn.com/abstract=3846138 (visited on 04/12/2023). [9]

Mudde, Cas, and Cristobal Rovira Kaltwasser. 2017. *Populism: A Very Short Introduction*. en. Oxford University Press. Google-Books-ID: zVjODQAAQBAJ. [7, 8, 37, 38]

Mueller, John E. 1970. "Presidential Popularity from Truman to Johnson." en. *American Political Science Review* 64 (1): 18–34. DOI: 10.2307/1955610. (Visited on 03/19/2023). [8]

Munshi, Kaivan, and Mark Rosenzweig. 2015. "Insiders and Outsiders: Local Ethnic Politics and Public Goods Provision." en. Working paper w21720. Cambridge, MA: National Bureau of Economic Research, w21720. DOI: 10.3386/w21720. (Visited on 03/19/2023). [10]

Norris, Pippa, and Ronald Inglehart. 2019. *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. en. Cambridge University Press. Google-Books-ID: I8aGDwAAQBAJ. [18]

Otten, S., C. Dustmann, B. Eichengreen, A. Sapir, G. Tabellini, and G. Zoega. 2017. *Europe's Trust Deficit: Causes and Remedies*. eng. Vol. 1, London, UK: CEPR Press. URL: https://voxeu.org/content/europes-trust-deficit-causes-and-remedies (visited on 04/12/2023). Publication Title: Monitoring International Integration: Vol.1. (1st ed.). CEPR Press: London, UK. (2017). [45]

Pande, Rohini. 2003. "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities? Theory and Evidence from India." en. *American Economic Review* 93 (4): 1132–51. DOI: 10.1257/000282803769206232. (Visited on 04/12/2023). [10]

Persson, Torsten, and Guido Tabellini. 2002. *Political Economics: Explaining Economic Policy*. en. MIT Press. Google-Books-ID: XC4MEAAAQBAJ. [9, 13, 29, 30]

Pitkin, Hanna F. 1967. *The Concept of Representation*. en. University of California Press. Google-Books-ID: AgUVWLswTNEC. [5, 47]

Plott, Charles R. 1967. "A Notion of Equilibrium and its Possibility Under Majority Rule." *American Economic Review* 57 (4): 787–806. URL: https://www.jstor.org/stable/1815369 (visited on 04/12/2023). Publisher: American Economic Association. [13, 21]

Polk, Jonathan, Jan Rovny, Ryan Bakker, Erica Edwards, Liesbet Hooghe, Seth Jolly, Jelle Koedam, Filip Kostelka, Gary Marks, Gijs Schumacher, Marco Steenbergen, Milada Vachudova, and Marko Zilovic. 2017. "Explaining the Salience of Anti-elitism and Reducing Political Corruption for Political Parties in Europe with the 2014 Chapel Hill Expert Survey Data." en. *Research & Politics* 4 (1): 205316801668691. DOI: 10.1177/2053168016686915. (Visited on 03/19/2023). [15]

Przeworski, Adam, Susan C. Stokes, and Bernard Manin. 1999. *Democracy, Accountability, and Representation*. en. Cambridge University Press. Google-Books-ID: aid08OpvF5gC. [47]

Puglisi, Riccardo, and James M. Snyder. 2015. "The Balanced US Press." en. *Journal of the European Economic Association* 13 (2): 240–64. DOI: 10.1111/jeea.12101. (Visited on 03/19/2023). [46]

Saari, Donald G. 1997. "The Generic Existence of a Core for Q-Rules." en. *Economic Theory* 9 (2): 219–60. DOI: 10.1007/BF01213800. (Visited on 04/17/2023). [13]

Sappington, David E. M. 1991. "Incentives in Principal-Agent Relationships." en. *Journal of Economic Perspectives* 5 (2): 45–66. DOI: 10.1257/jep.5.2.45. (Visited on 03/19/2023). [47]

Schakel, Wouter, and Armen Hakhverdian. 2018. "Ideological Congruence and Socio-Economic Inequality." en. *European Political Science Review* 10 (3): 441–65. DOI: 10.1017/S1755773918000036. (Visited on 03/19/2023). [10]

Schofield, N. 1978. "Instability of Simple Dynamic Games." en. *Review of Economic Studies* 45 (3): 575–94. DOI: 10.2307/2297259. (Visited on 03/19/2023). [13]

Schofield, Norman. 1983. "Generic Instability of Majority Rule." en. *Review of Economic Studies* 50 (4): 695. DOI: 10.2307/2297770. (Visited on 03/19/2023). [13]

Schofield, Norman. 2004. "Equilibrium in the Spatial 'Valence' Model of Politics." en. *Journal of Theoretical Politics* 16 (4): 447–81. DOI: 10.1177/0951629804046150. (Visited on 03/19/2023). [7]

Schofield, Norman. 2007. "The Mean Voter Theorem: Necessary and Sufficient Conditions for Convergent Equilibrium." en. *Review of Economic Studies* 74 (3): 965–80. DOI: 10.1111/j.1467-937X.2007.00444.x. (Visited on 03/19/2023). [7, 13, 21, 29, 30]

Schofield, Norman, and Alexei Zakharov. 2010. "A Stochastic Model of the 2007 Russian Duma Election." en. *Public Choice* 142 (1-2): 177–94. DOI: 10.1007/s11127-009-9483-2. (Visited on 03/19/2023). [7, 13]

Selects. 2009. "Enquête auprès des candidats - 2007 [Dataset]." Edited by FORS. URL: www.selects.ch%20https://doi.org/10.23662/FORS-DS-531-1. [14]

Shapiro, R. Y. 2011. "Public Opinion and American Democracy." en. *Public Opinion Quarterly* 75 (5): 982–1017. DOI: 10.1093/poq/nfr053. (Visited on 03/19/2023). [10]

Sheffer, Lior, Peter John Loewen, Stuart Soroka, Stefaan Walgrave, and Tamir Sheafer. 2018. "Nonrepresentative Representatives: An Experimental Study of the Decision Making of Elected Politicians." en. *American Political Science Review* 112 (2): 302–21. DOI: 10.1017/S0003055417000569. (Visited on 03/19/2023). [48]

Stokes, Donald. 1992. "Valence Politics." *Electoral politics*, 141–64. [30]

Swissvotes. 2021. "Swissvotes – the Database on the Swiss Popular Votes: Dataset." Edited by Année Politique Suisse, University of Bern. URL: www.swissvotes.ch. [16]

Tabellini, Marco. 2020. "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration." en. *Review of Economic Studies* 87 (1): 454–86. DOI: 10.1093/restud/rdz02 7. (Visited on 03/19/2023). [8]

Walgrave, Stefaan, and Jonas Lefevere. 2013. "Ideology, Salience, and Complexity: Determinants of Policy Issue Incongruence between Voters and Parties." *Journal of Elections, Public Opinion and Parties* 23 (4): 456–83. DOI: 10.1080/17457289.2013.810630. (Visited on 04/13/2023). Publisher: Routledge _eprint: https://doi.org/10.1080/17457289.2013.810630. [47]

Weßels, Bernhard. 2013. "European Parliament Election Study 2009, Candidate Study (Europäische Kandidatenstudie 2009) - Vollversion." Edited by Wissenschaftszentrum Berlin für Sozialforschung (WZB), Berlin. DOI: 10.4232/1.12961. Published: GESIS Datenarchiv, Köln. ZA5053 Datenfile Version 2.0.0, https://doi.org/10.4232/1.12961. [14]

Wittman, Donald. 1977. "Candidates with Policy Preferences: A Dynamic Model." en. *Journal of Economic Theory* 14 (1): 180–89. DOI: 10.1016/0022-0531(77)90091-6. (Visited on 04/17/2023). [30]

Chapter 2

Would Europeans Accept Immigrants If They Knew Them?*

2.1 Introduction

Immigration has been a contested topic in Europe during the last decades and became particularly salient during the European refugee crisis when several million asylum seekers entered Europe. In surveys, most Europeans indicate that they favor reducing immigration to their countries, and majorities in many European countries even agree with the statement that their country should not accept any refugees at all (Esipova et al., 2015; Skinner and Gottfried, 2017). A possible explanation for these policy attitudes is that natives hold biased beliefs about the characteristics of immigrants, like their education or work skills. However, while research has confirmed that the beliefs of natives are biased (Alesina, Miano, and Stantcheva, 2023), it is still unknown whether these biases can account for anti-immigration attitudes or whether Europeans would hold anti-immigration attitudes even if they were well-informed.

It matters whether anti-immigration attitudes result from biased beliefs. If biased beliefs drove anti-immigration attitudes, Europeans would favor immigration if they were well informed. In this case, informing natives about the true characteristics of immigrants is a plausible policy response. First, such information campaigns would meaningfully affect attitudes. Second, they might increase the well-being of natives by enabling them to make better-informed decisions (Grigorieff, Roth, and Ubfal, 2020). Policy implications are very different if biased beliefs are not responsible for anti-immigration attitudes. In this case, information interventions are un-

likely to affect attitudes. Moreover, high immigration rates are truly inconsistent with natives' preferences.

This paper examines whether Europeans would favor immigration if they were well-informed about the true characteristics of immigrants. To that end, it proposes a new method to estimate immigration attitudes that subjects with biased beliefs would have if their beliefs were *not* biased. The key idea is to elicit subjects' attitudes toward immigrant vignettes, which include a detailed description of the immigrants' characteristics. Because subjects are informed about the immigrants' characteristics, this limits the effect of biased beliefs about these characteristics. Such vignette studies have so far only been done with hypothetical vignettes. Consequently, my method matches the resulting attitude data with data on the actual characteristics of immigrants.

I apply this method by combining attitude data from a vignette study performed in multiple European countries by Bansak, Hainmueller, and Hangartner (2016) with data from various administrative sources about the actual characteristics of asylum seekers in Europe. I find that even if Europeans had been well-informed about the characteristics of the asylum seekers who came to Europe during the refugee crisis, most asylum seekers would have been denied entry by a majority of Europeans. For instance, the average asylum seeker would have been rejected by 64% of Europeans.

This "informed" opposition to immigration is common across many demographic dimensions. For instance, majorities in most European countries, political ideologies, occupation categories, all age groups, and education levels would have rejected the average asylum seeker if they had been informed about his characteristics. Thus, anti-immigration attitudes are widespread even if biased beliefs about many characteristics are eliminated.

Still, the characteristics of asylum seekers matter greatly for the immigration attitudes of Europeans. One key distinction is between those who apply for asylum because of fear of war or persecution and asylum seekers who come for other reasons, like economic opportunities (Wright, Levy, and Citrin, 2016; Verkuyten, Altabatabaei, and Nooitgedagt, 2018; Verkuyten, Mepham, and Kros, 2018). This distinction is important because asylum systems were explicitly designed to provide shelter to the former "humanitarian asylum seekers" but not to the latter "other asylum seekers." Hence, a key task of asylum ministries is to classify asylum seekers into these categories (Dustmann et al., 2017).[1] Moreover, according to data from European asylum ministries, about half of those who applied for asylum dur-

---

1. Asylum ministries decide on the legal status an asylum seeker is granted. These include several positive statuses that correspond to what I summarize as "humanitarian asylum seekers." Alternatively, they can reject an application, meaning they think the asylum seeker did not migrate due to fear of war or persecution. I use the terms "humanitarian asylum seeker" and "other asylum seeker" because they directly relate to the attitude data I use.

ing the refugee crisis were humanitarian asylum seekers. My results show that the "informed" attitudes of the average European are consistent with asylum laws. Anti-immigration attitudes are extremely strong for asylum seekers who did not come for humanitarian reasons. Furthermore, I show that attitudes toward this group are strikingly similar between European countries and various demographic groups. In any demographic group I consider, and even among those who place themselves on the far left, an often overwhelming majority opposes immigration by members of this group. This suggests that policies like the deportation of rejected asylum seekers find broad "informed" support across all European countries and the political spectrum.

In contrast, about half of Europeans favor letting in those who flee war or persecution. Furthermore, different demographic groups tend to have very different opinions regarding humanitarian asylum seekers. For example, in about half of the countries considered, majorities favor taking in humanitarian asylum seekers, while majorities oppose it in the other half of the countries. Similarly, those who identify as liberals predominantly favor providing shelter for those who flee war or persecution, while those on the center or the right are predominantly against it. Hence, while there is agreement that asylum seekers who did not migrate for humanitarian reasons should not be allowed to immigrate across all countries and demographic groups, there is disagreement regarding humanitarian asylum seekers.

This matters for the political implementability of asylum policies. The European Union has struggled to design asylum policies that all its member states can agree on (Dustmann et al., 2017). It is presumably easier for the government of a country to implement a policy if the majority of its population is in favor of it. Hence, policies that large majorities within all European countries agree with are relatively easy to implement. My results indicate that policies to reject and deport asylum seekers who did not come to Europe due to persecution or war would still have such support if Europeans were well-informed about the characteristics of asylum seekers. In addition, such policies find broad support across the political spectrum within countries. In contrast, policies about hosting asylum seekers who flee war or persecution do find support in some countries and among left-leaning Europeans, while they are primarily rejected in other countries and by those who are not politically left-leaning. Such disagreement might make it hard to implement such policies.

This paper relates to the extensive literature on immigration attitudes (Hainmueller and Hopkins, 2014). Immigration attitudes of European voters and citizens are much more negative than parliamentarians' attitudes (Andreadis and Stavrakakis, 2017; Dalton, 2017; Günther, 2022; Kübler and Schäfer, 2022) and policy decisions (Günther, 2022). Similarly, large majorities demand lower immigration rates, particularly regarding asylum seekers (Ceobanu and Escandell, 2010; Blinder, 2015; Kaufmann, 2017; Hix, Kaufmann, and Leeper, 2021). This opposition to immigration is remarkably stable over time (Kustov, Laaker, and Reller, 2021).

A potential explanation is that Europeans hold biased beliefs about the characteristics of immigrants. Indeed, most studies find that Europeans overestimate the frequency of negative characteristics, like unemployment (Barrera et al., 2020; Grigorieff, Roth, and Ubfal, 2020; Alesina, Miano, and Stantcheva, 2023). This suggests that natives would be more in favor of immigration than they indicate in polls if they were well-informed about the true characteristics of immigrants, which raises the question of how immigration attitudes would look if natives were well-informed. Would most Europeans accept immigrants if they knew them? Or would they still reject most immigrants, only to a slightly lower extent?

To find out, social scientists have tested how informing citizens about immigrant characteristics changes immigration attitudes. Results are mixed. Some studies find insignificant effects (Hopkins, Sides, and Citrin, 2019; Barrera et al., 2020) while others, containing the most comprehensive studies in this field, find that information interventions lead to more favorable views toward immigration (Kaufmann, 2019; Grigorieff, Roth, and Ubfal, 2020; Facchini, Margalit, and Nakata, 2022; Alesina, Miano, and Stantcheva, 2023).

Such information interventions are a great tool to analyze how subjects react to specific information. However, from the perspective of the research question of this paper, they have a major shortcoming: they only correct subjects' beliefs regarding a single characteristic. Informing subjects about many characteristics at once might overstrain the ability of subjects to incorporate all new information. Furthermore, subjects might need time to fully comprehend the new information they are presented with. Overcharged with incorporating new pieces of information and quickly forming new attitudes, they might state their original preferences.

In this paper, I propose a complementary approach in which biases are not eliminated. Instead, attitudes are directly elicited for the immigrants who actually immigrated using a combination of vignette studies and data on the characteristics of immigrants. I find that most Europeans would still oppose immigration by most asylum seekers if they were well informed about their characteristics. Consistent with the limited effect of information interventions, this suggests that biased beliefs are not the main driver of anti-immigration attitudes.

2.2   Data

The key idea behind my estimation approach is to combine data about informed attitudes toward hypothetical asylum seekers (AS) with information on the characteristics of real asylum seekers. Consequently, my data comes from two types of sources.

Data about attitudes toward hypothetical asylum seekers come from Bansak, Hainmueller, and Hangartner (2016). The authors conducted surveys in 15 European countries with at least 1,000 participants per country in early 2016, at the

height of the so-called "refugee crisis." Each subject was shown 10 vignettes of hypothetical asylum seekers. The vignettes described the characteristics of a hypothetical asylum seeker regarding nine demographic variables: the asylum seeker's age, gender, country of origin, previous occupation, religion, skills with the language of the host country, the consistency of his asylum testimony, his vulnerability, i.e., what would happen to them if they returned to their country of origin (e.g., torture), and the reason for his migration decision. These variables were categorical and could take on two to seven different values. For instance, each vignette stated whether the asylum seeker was a Christian, a Muslim, or an Agnostic. Table 2.E.1 and Table 2.E.2 show the full list of variables with all of their possible realizations, and Figure 2.D.1 provides two example vignettes. The authors randomly assigned the characteristics and resulting vignettes to the participants. For each variable, each characteristic was chosen with the same probability. The large number of variables and categories per variable makes it possible to identify attitudes toward 136,080 different hypothetical asylum seekers. For each asylum seeker, subjects were asked whether the asylum seeker should be accepted to live in their country or whether he or she should be sent back.

The second type of data I use provides information on the characteristics of real asylum seekers. Because I want to link data on the attitudes toward the asylum seeker vignettes with data on the characteristics of actual asylum seekers, I compile data on asylum seeker characteristics about the same variables that subjects of Bansak, Hainmueller, and Hangartner (2016) were informed about. I compile such data from several sources. If possible, I rely on Eurostat, which collects data from the migration ministries of EU member states. Eurostat provides data on the joint distribution of age, sex, country of origin, and the protection status of asylum seekers. I use the latter variable to distinguish between "humanitarian asylum seekers" (HAS) and "other asylum seekers" (OAS). International asylum laws require European countries to host HAS while they do not have to host OAS. Hence, a key task of migration ministries is to distinguish between these two types of asylum seekers. To that end, ministries try to identify HAS and grant them positive protection status but reject asylum applications by OAS. Thus, whether they have granted positive protection status to an asylum seeker is likely indicative of whether he is a HAS. Moreover, it indicates whether the asylum testimony is consistent or not because this is the major factor that migration ministries use when deciding on asylum applications. When using data from Eurostat, I pool data for the years 2015-2020 to focus on the refugee crisis.[2]

Eurostat does not provide information about asylum seekers' previous occupations, language skills, religious denominations, or vulnerability. To assess the reli-

---

2. The main influx of asylum seekers during the refugee crisis happened in 2015 and 2016. However, many decisions of migration ministries on the legal status of asylum seekers were made years later.

gious denomination of actual asylum seekers, I combine data on the countries of origin with data on the distribution of religious denominations of asylum seekers from these countries of origin. I take estimates for the latter from Siegert (2020), who calculates them based on the IAB-BAMF-SOEP survey among asylum seekers (Brücker, Rother, and Schupp, 2018). The sample is based on asylum seekers who made their application in Germany between 2013 and 2016, at the time of the refugee crisis. A distinguishing feature of this study is that asylum seekers are interviewed independently of the result of their application. This makes it possible to derive separate estimates for OAS and HAS. Moreover, the dataset includes a large sample of about 10,000 asylum seekers and was designed to be representative of the universe of asylum seekers in Germany (Babka von Gostomski et al., 2016; Brücker, Rother, Schupp, et al., 2016; Jacobsen et al., 2019). Germany was by far the most frequent destination for asylum seekers in absolute numbers. Still, estimates for the characteristics of those who apply for asylum in other countries are biased if these people systematically differ from those who apply in Germany. I use data from Schröder, Zok, and Faulbaum (2018) and Brücker, Rother, Schupp, et al. (2016) to infer the distribution of previous occupations and knowledge of the host country's language of asylum seekers. Both studies present survey-based evidence for asylum seekers in Germany. To estimate the frequencies of the vulnerability categories distinguished by Bansak, Hainmueller, and Hangartner (2016) for HAS and OAS separately, I use the survey data from Schröder, Zok, and Faulbaum (2018) in which asylum seekers self-reported their vulnerabilities. Section 2.A provides an overview of the asylum seeker characteristics regarding all variables.

2.3 The Method

Let there be a group A containing k asylum seekers $(1, 2, ..., k)$. An electorate decides whether to host (accept) or return the whole group A to their country of origin. To this end, each member of the electorate either votes for "accept A" or "return A." I use the share of the electorate that votes for acceptance as a measure of the acceptance of asylum seekers and label this measure "acceptance rate." Compared to qualitative survey items which ask subjects whether "many," "some," or "few" immigrants should be allowed to enter or whether immigration should be "reduced significantly," this measure is arguably easier to interpret and more comparable between countries and points in time. Terms like "many" and "few" might differ in meaning between countries and over time. In contrast, the meaning of "send back" and "accept" are more clearly defined.

2.3.1 Estimating Acceptance Rates for Individual Asylum Seekers

To consider acceptance rates for an individual asylum seeker, let the group of asylum seekers A be a singleton $(k = 1)$. Let there be l characteristics of the single

asylum seeker a: $c_a^1, c_a^2, ..., c_a^l \in \{0, 1\}$. The asylum seeker either has characteristic c_a^i ($c_a^i = 1$) or a reference characteristic ($c_a^i = 0$). For instance, $c_a^i = 1$ might denote that the asylum seeker is male while $c_a^i = 0$ denotes that the asylum seeker is female. I summarize these characteristics in the type-vector $t_a = (c_a^1, ..., c_a^l)$. Furthermore, let α denote the true acceptance rate for an asylum seeker with type $t_a = (0, 0, ..., 0)$ and $\beta = (\beta_1, ..., \beta_l)$ the vector of true effects of these characteristics relative to the reference characteristic on the acceptance rate.

I estimate the acceptance rate, ar_a, for asylum seeker a of type t_a as

$$ar_a = a + b \cdot t_a, \tag{2.1}$$

where a and b are estimates for α and β respectively, estimated by OLS from a regression equation of the form

$$\mathbb{1}[r \text{ accepts } a]_{r,a} = a + b \cdot X_{r,a} + \varepsilon_{r,a}. \tag{2.2}$$

Here, $\mathbb{1}[r \text{ accepts } a]_{r,a}$ equals one if respondent r accepted asylum seeker a to immigrate and zero else and $X_{r,a}$ includes all characteristics Bansak, Hainmueller, and Hangartner (2016) employ, excluding one reference category for each variable. Hence, to get the acceptance rate for a particular asylum seeker, I sum up all the effects of the acceptance rates of the characteristics that define him. This estimation framework relies on several assumptions to yield reliable estimates. First, most relevant variables have to be included in X. Second, the estimates for α and β have to be unbiased. Third, interaction effects between asylum seeker characteristics have to be negligible.

Regarding assumption one, most essential variables are included because this is how they were chosen. Bansak, Hainmueller, and Hangartner (2016) carefully reviewed the existing literature on determinants of attitudes towards immigrants to select the characteristics that are most important in determining acceptance towards them. This suggests that including additional variables in X would not strongly alter the results.

Concerning assumption two, the type of vignette study I employ has been shown to be indicative of real-world decision-making by Hainmueller, Hangartner, and Yamamoto (2015).[3] They find that conjoint experiments yield valid estimates of β, the

---

3. The authors use the fact that naturalizations of immigrants in many Swiss municipalities have been decided by plebiscites after voters are informed about the applicants' characteristics. They run conjoint experiments in these municipalities and provide subjects with the same type of information as they would see in a real referendum. They then compare the results of real referendums and conjoint studies through OLS regressions of the form

$$y_a = \alpha + \beta \cdot X_a + \varepsilon_a,$$

where y_i is an indicator variable that equals one if applicant a is accepted by the plebiscite and zero else, and X_a is a matrix that contains applicant characteristics.

coefficients on the characteristics. However, they also find α to be biased upwards. Hence, while conjoint experiments allow researchers to assess how characteristics translate into the likelihood of acceptance, the level of acceptance is overestimated. The latter finding implies that the acceptance rates I estimate below are biased upwards. However, since the estimates for the acceptance rates turn out to be very low, correcting this bias would not alter but rather strengthen the conclusions.

Regarding assumption three, Section 2.B shows that 1) coefficients on interaction effects are small compared to main effects and 2) the number of significant interaction effects is only slightly larger than expected by chance while half of all main effects are significant at the 5% level. I conclude that omitting interaction effects does not strongly bias estimates for the acceptance rates while simplifying the analysis substantially.

2.3.2 Estimating Acceptance Rates for Groups of Asylum Seekers

Let there be a group of asylum seekers A consisting of $k > 1$ individuals $(1,...,k)$. Let there be l characteristics of the group $c_A^1, c_A^2, ..., c_A^l$, and each of these characteristics is obtained in A with frequency $f_A^1, ..., f_A^l \in [0, 1]$. Let t be a vector that includes all of these frequencies, i.e., $t_A = (f_A^1, ..., f_A^l)$. Furthermore, let β_j be the true effect on the acceptance rate of $f_A^j = 1$ compared to $f_A^j = 0$ for any $j \in \{1, ..., k\}$. Consequently, let $\beta = (\beta_1, ..., \beta_l)$ be the vector of true effects of the characteristics on the acceptance rate and α the true acceptance rate for a group of asylum seekers who all have the reference characteristics regarding each variable ($t_A = (0, ..., 0)$).

This paper focuses on three groups of asylum seekers. First, it considers the case where A contains all adult Asylum Seekers that migrated to Europe during the refugee crisis (group AS). Second, it distinguishes between those who migrated for humanitarian reasons and those who came to Europe for other reasons. To make this distinction, I use the fact that each asylum seeker vignette included a "reason for migrating" variable. There, subjects were directly told whether the asylum seeker migrated due to persecution or economic reasons. I define an asylum seeker whose vignette states that he migrated due to persecution humanitarian asylum seeker (HAS), while I label all asylum seekers who have been stated to have migrated because of economic reasons as other asylum seekers (OAS).

I estimate the acceptance rate for group $A \in \{AS, HAS, OAS\}$ with type t_A as

$$ar_A = a + t_A \cdot b, \tag{2.3}$$

where a and b are estimates for α and β respectively and obtained from a regression of the form

$$\mathbb{1}[r \text{ accepts } a]_{r,a} = a + b \cdot X_{r,a} + \varepsilon_{r,a}. \tag{2.4}$$

Here, r is an index for the respondent in the vignette study who either accepts or rejects immigration by an asylum seeker a belonging to group A. $X_{r,a}$ includes all

characteristics Bansak, Hainmueller, and Hangartner (2016) employ, excluding the reference categories. For example, about 70% of all HAS are men. Hence, to calculate ar_{HAS}, I multiply the coefficient on the male indicator by 0.7 and add the result to the intercept.[4]

There are two ways to interpret this measure. First, one could interpret it as the acceptance rate for the average HAS/OAS. Second, this measure can be interpreted as the acceptance rate for the whole group of HAS/OAS seen as one entity under the assumption that the acceptance rate for a group of asylum seekers is the linear combination of the acceptance rates of the individual asylum seeker contained in A. Kaufmann (2018) argues that when faced with an individual migrant, people rather empathize with him and, thereby, are more likely to accept him. On the other hand, groups are rather perceived as abstract entities with which it is much harder to identify or empathize. This implies that the acceptance rate for a group of migrants is weakly lower than the average acceptance rate for all migrants individually. Therefore, the following estimates might alternatively be viewed as an upper bound for the true acceptance rates that would result if Europeans were to choose between taking in or returning all asylum seekers of a certain group.

2.4   Results

2.4.1   Attitudes toward Hypothetical Asylum Seekers

Figure 2.1 summarizes acceptance rates for all hypothetical asylum seeker vignettes by the asylum seeker type they belong to. The boxplot on the left summarizes acceptance rates for all asylum seeker vignettes while the other two boxplots refer only to HAS or OAS respectively.[5]

The boxplot regarding AS in total reveals that more than three-quarters of asylum seeker vignettes were rejected by a majority of Europeans. However, the range of acceptance rates is large because Bansak, Hainmueller, and Hangartner (2016) considered quite different asylum seekers. The most positively evaluated asylum seeker vignettes were accepted by nearly 70% of respondents, while acceptance rates for the most negatively evaluated asylum seekers are close to 10%.

The other two boxplots show that much of this variation can be explained by a distinction between HAS and OAS. Even though there are huge differences in demographic characteristics between OAS, all types of OAS would be rejected by a Europe-wide referendum according to the point estimates. This is even true for

---

4. Section 2.A provides an overview of the distribution of asylum seeker characteristics by type and Section 2.C discusses in detail how I matched observational data on characteristics to vignettes.

5. The boxplots for HAS and OAS are identical except for a constant difference. This is driven by the fact that in this analysis, I plot the acceptance rates for hypothetical AS with otherwise identical demographics. I adjust for group differences between HAS and OAS in Section 2.4.2.

.
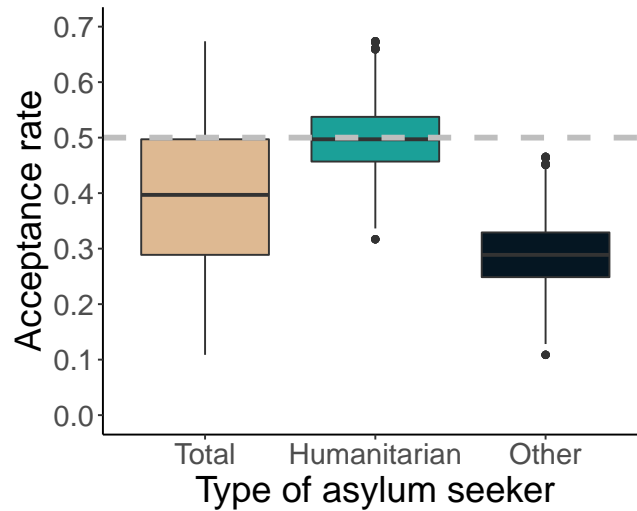
Figure 2.1. Boxplots of Acceptance Rates for Different Asylum Seekers by Type
Note: Boxplots summarize acceptance rates for all asylum seekers, humanitarian asylum seekers, and other asylum seekers with various characteristics. Acceptance rates are estimated through Equation 2.1

types of OAS that likely elicit higher acceptance rates than most real OAS. One example is a 21-year-old female Christian who speaks the host country's language fluently and has worked as a doctor before leaving her country of origin. This finding strongly suggests that nearly all OAS that apply would have been rejected if the European population had been asked. The fact that the depicted acceptance rates serve as upper bounds for the true acceptance rates does not qualitatively change this conclusion but strengthens it.

Let's now turn to the HAS. The median lies close to the 50% threshold, indicating that about half of the HAS vignettes would have been accepted in a Europe-wide plebiscite while the other half would have been rejected. Hence, while Europeans vehemently oppose OAS immigration, there is more disagreement regarding HAS. It matters for the implementability of a policy whether majorities oppose or accept it. My results indicate that it is hard to implement policies that aim to host OAS. If undertaken nonetheless, policymaking will be at odds with voters' attitudes which have been shown to be associated with distrust in democratic institutions and the rise of populist parties (Stecker and Tausendpfund, 2016; Bakker, Jolly, and Polk, 2022; Günther, 2022; Kübler and Schäfer, 2022).

However, what acceptance rates correspond to the asylum seekers that actually immigrated to Europe during the refugee crisis? To answer this question, I combine data on attitudes toward asylum-seekers with data on asylum-seeker characteristics. Section 2.A discusses the true characteristics of asylum seekers who immigrated during the European refugee crisis in detail. To sum up, according to Europeans, most actual asylum seekers have less desirable characteristics than asylum seeker

vignettes. For instance, in about one in seven vignettes, it was indicated that the asylum seeker had worked as a doctor. Similarly, about a third of hypothetical asylum seekers speak the language of the hist country fluently upon arriving. As shown in Section 2.A, the actual share that has worked as doctors or in similarly prestigious occupations is much lower, and good skills in the host country's language were virtually non-existent. Because Europeans are more likely to accept highly skilled and those who speak their language, acceptance rates for real asylum seekers tend to be lower than those shown in Figure 2.1. Hence, acceptance rates for most real asylum seekers lie in the lower part of the boxplots. This further strengthens the result that Europeans generally reject OAS. I now go beyond this qualitative description by developing a quantitative measure for the acceptance rates toward asylum seekers, which combines information on the preferences of Europeans and information on the actual characteristics of asylum seekers.

2.4.2 Attitudes toward Real Asylum Seekers

Figure 2.2 shows the acceptance rate toward the average real AS, HAS, and OAS, respectively, based on Equation 2.3. Preference data is pooled across all European countries in my dataset. The actual share of HAS among all asylum-seekers during the refugee crisis was about 48% based on national European asylum ministries data, as further discussed in Section 2.A. The acceptance rate for the average asylum seeker is about 36%. Hence, 64% of Europeans would have rejected immigration by the average asylum seeker if they had been well informed about his true characteristics. This estimate is precisely estimated and significantly different from 50%.
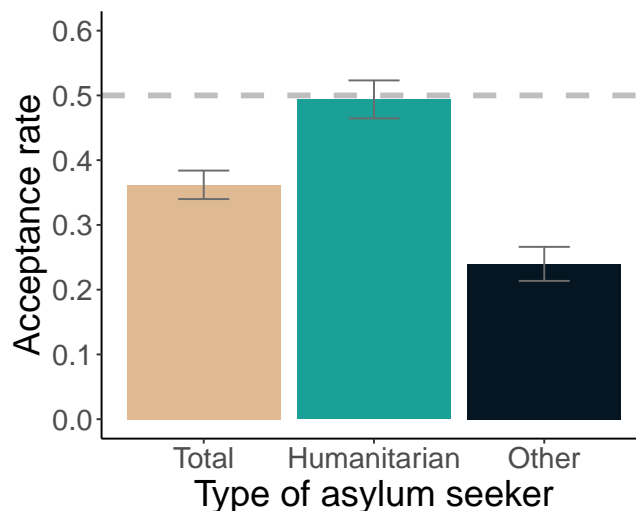
Figure 2.2. Acceptance Rates for the Average Asylum Seeker by Type
Note: The vertical axis depicts acceptance rates for the average total of asylum seekers, humanitarian asylum seekers, and other asylum seekers, respectively, and 95% confidence intervals based on robust standard errors. Estimates are derived from Equation 2.3.

However, acceptance depends strongly on whether the asylum seeker is a HAS or an OAS. The difference in the acceptance rates between the average HAS and the average OAS equals more than 25 percentage points. My estimates suggest that about 49% of European citizens reject hosting the average HAS. However, I cannot reject the hypothesis that a majority favors HAS immigration at the 5% level of significance. Either way, the general population has no strong tendency to accept or reject the average humanitarian asylum seeker. In stark contrast, the acceptance rate for the average OAS is estimated to equal just under 24%. Hence, a huge majority of 76% favor sending back the average OAS, and I can reject the hypothesis that a majority favor OAS-immigration.

These results confirm that most Europeans oppose immigration by OAS, while the average European appears indifferent toward hosting HAS. These preferences broadly align with the fundamental idea behind modern asylum laws, which have been designed to distinguish between HAS and OAS and oblige countries to host HAS but not to host OAS. However, because about half of all asylum seekers are OAS, most Europeans are opposed to immigration by the average asylum seeker or viewed differently by the total of asylum seekers.

2.4.2.1 Attitudes toward Real Asylum Seekers by Country of Respondent

So far, I have pooled responses from all European countries. Do these Europe-level results mask important between-country heterogeneity? Figure 2.3 depicts acceptance rates for all AS, HAS, and OAS for 15 countries. Bars indicate the estimated acceptance rate for AS, HAS, and OAS by the respondent's country.
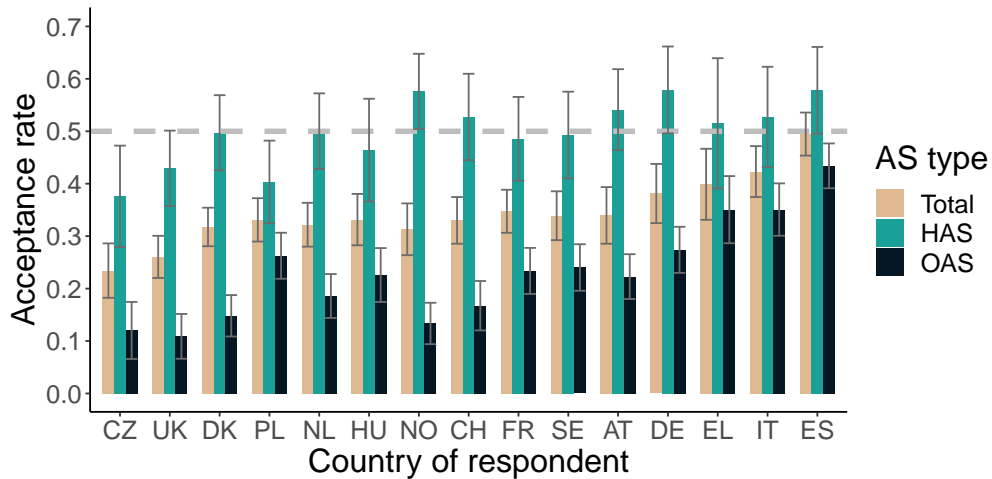
Figure 2.3. Acceptance Rates for the Average Asylum Seeker by Type and Country
Note: The horizontal axis shows the country of the respondent. The vertical axis depicts acceptance rates for the average total of asylum seekers, humanitarian asylum seekers, and other asylum seekers, respectively. Estimates come from Equation 2.3. The figure also depicts 95% confidence intervals. Standard errors are clustered at the spectator level.

Regarding the total average asylum seeker, the acceptance rate varies between a minimum of about 23% in the Czech Republic and a maximum of just under 50% in Spain. Except for Spain, estimates are significantly below 50% in every country. Hence, in every country considered, majorities would be opposed to immigration by the average asylum seeker if they were well informed about his characteristics.

Regarding HAS, all acceptance rates are between 38% and 59%. Spain and Germany have the highest acceptance rates, while Poland and the Czech Republic have the lowest. Acceptance rates are significantly larger than 50% Norway and close to being significantly larger in Spain and Germany. At the same time, estimates are significantly lower than 50% in the Czech Republic and Poland and close to being significantly lower in the UK. In all other countries, I cannot reject the hypotheses that estimates are above or below 50%. Hence, majorities in some European countries are opposed to asylum seekers who flee war or persecution, while majorities in other European countries accept them.

Acceptance rates for OAS are significantly lower than those for HAS in every country. Moreover, I can reject the hypothesis that citizens of any of the 15 countries were willing to let in the average OAS at the 5% level of significance. Even in the countries that took in the most asylum seekers and are typically seen as very open toward immigration, like Sweden and Germany, the acceptance rates for OAS are quite low (between 24% and 27%). Hence, despite acceptance rates for OAS differing quantitatively between countries, majorities in all countries agree that OAS should not be allowed to immigrate.

It matters for the political implementability of policies whether majorities oppose or reject a policy. Immigration by asylum seekers is a cross-national European issue. Hence, implementing effective immigration policies requires the approval of each, or at least most, countries. Consequently, the implementability of Europe-wide immigration policies strongly depends on whether majorities in most countries can agree on it. Hence, these country-level results suggest that policies preventing OAS from entering Europe are easier to implement than policies denying or grating entry to HAS.

2.4.2.2 Attitudes toward Real Asylum Seekers by Political Orientation of Respondent

Figure 2.4 depicts attitudes toward asylum seekers by political orientation. Respondents were asked to place themselves on a 5-point political left-right spectrum from "Far Left" to "Far Right". I pool respondents from all countries. Reassuringly, acceptance rates tend to be lower for subjects who identify as more right-wing.

Focusing on acceptance for the total average asylum seeker reveals that majorities on the political left, the center, the right, and the far right would all oppose immigration by the average asylum seeker if they were well-informed about his characteristics. Only those on the far left predominantly accept immigration by the average asylum seeker and only by a small margin.
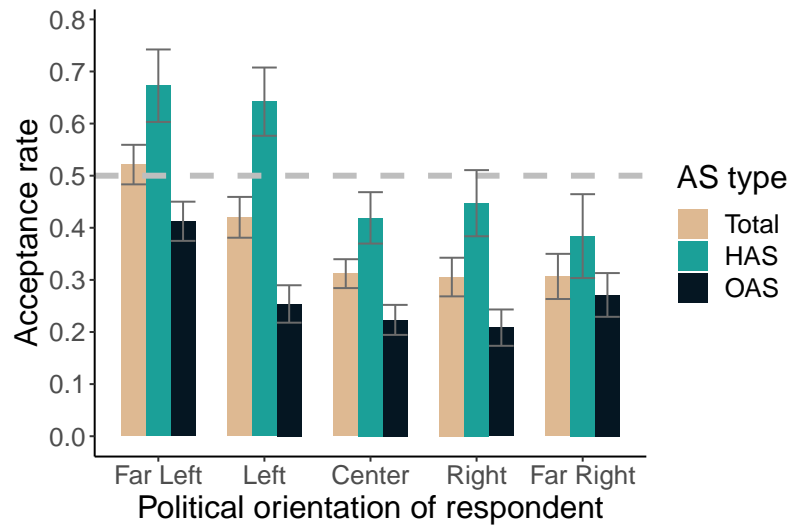
Figure 2.4. Acceptance Rates for the Average Asylum Seeker by Type and Political Attitude
Note: The vertical axis depicts acceptance rates for the average total of asylum seekers, humanitarian asylum seekers, and other asylum seekers, respectively by political attitude of the survey respondent. Data is pooled across Europe.

Acceptance rates toward HAS reveal that the country- and Europe-level acceptance rates around 50% mask opposing attitudes by those on the political left and all other citizens. About 67% of those on the far left and 64% on the left accept immigration by the average HAS. In contrast, estimates for acceptance rates among those on the center, the right, and the far right are below 50%. Thus, HAS-immigration is a question that Europeans with different political orientations are divided on. Consequently, it might be hard to reach a consensus between political camps on whether to accept immigration by humanitarian asylum seekers or not.

Regarding OAS, attitudes differ much less between citizens with different political orientations. All estimates are between 20% and 42%. Hence, although there is more acceptance among left-leaning individuals concerning OAS, even those on the far left predominantly oppose immigration by OAS. This shows that opposition to OAS migration is not confined to the political right but is shared across the political spectrum. Consequently, policies that aim at denying entry to OAS might be relatively easy to implement politically.

2.4.2.3   Attitudes toward Real Asylum Seekers along Other Dimensions

How do acceptance rates for HAS and OAS depend on other demographic characteristics? Instead of dividing the subject sample by country or political orientation, I split it along age groups (Figure 2.D.2), immigrant status (Figure 2.D.3), occupation categories (Figure 2.D.4), and sex (Figure 2.D.5) in the appendix. The results displayed in these figures are consistent with the previous literature on demographic

differences regarding immigration attitudes. For example, the very young are the age group most supportive of asylum seekers, and females are more likely to accept asylum seekers than males.

Despite these differences, the pattern documented above is remarkably similar across all sample splits. In all of these splits for all groups considered, the acceptance rate for the average total asylum seeker is below 50%. This is even true for subgroups that have been found to be very pro-immigrant-like students. Acceptance rates are even lower and more similar between subgroups regarding OAS. In contrast, most subgroups have acceptance rates close to 50% regarding HAS. Moreover, in all dimensions, point estimates for HAS acceptance rates are below 50% for some subgroups while those are above 50% for others. Hence, again, I find that majorities are united in all subgroups of all considered dimensions to keep OAS out of their countries, while groups are divided on whether to host or send back HAS.

Some additional findings are noteworthy. First, while I find that younger individuals tend to have higher acceptance rates, I also find the relationship to be non-linear. Both the very old and the very young seem to be most welcoming toward AS, while the middle-aged favor a more restrictive immigration policy. Age differences regarding divisive cultural topics like immigration have been argued to reflect a "silent revolution" in which each generation tends to be more progressive than the previous one, ultimately resulting in progressive cultural values replacing traditional ones. The rise of populist parties across western countries is seen as a cultural backlash by the old against this development (Norris and Inglehart, 2019). This prominent theory has strong policy implications. For instance, Norris and Inglehart (2019) have argued based on this theory that the rise of populism is a temporary phenomenon and will disappear as the very old die. Following this logic, simply riding out the rise of populism might be the best response for non-populist parties. While it is far beyond the scope of this article to test the cultural backlash hypothesis, the evidence presented here is not fully consistent with it. Contrary to what the theory predicts, the very old seem to be rather welcoming toward all groups of asylum seekers compared to most other age groups. In particular, this suggests that as the very old die, the electorate might become more resistant to immigration, the opposite of what the cultural backlash hypothesis predicts.

Second, while much research has been done on attitudes toward immigrants, much less research has been conducted on the attitudes toward immigrants. Figure 2.D.3 compares the acceptance rates of those who were born themselves in their country of residence and those that were not. Surprisingly, the acceptance rates are very close on average.

Finally, Figure 2.D.4 reveals variation in asylum seeker attitudes regarding occupation categories. Students are much more welcoming toward all groups of asylum seekers. Acceptance rates are also high among those who are unemployed and not searching for a job, while those who are unemployed but searching have much lower acceptance rates. The disabled have the lowest acceptance rates.

In sum, a broad informed consensus exists across countries, political camps, age groups, occupations, and other demographic dimensions that asylum seekers who did not migrate for humanitarian reasons should not be allowed to immigrate to Europe. In contrast, the average members of different demographic groups disagree on immigration by asylum seekers who migrated for humanitarian reasons. In particular, those on the political right and the center oppose it, while the vast majority on the political left favor it.

2.5 Conclusion

Most Europeans state in surveys that they prefer fewer immigrants coming to their countries. Recent evidence indicates that these attitudes might be driven by misperceptions Europeans hold about the characteristics of immigrants. This paper develops an estimation technique to estimate attitudes toward immigrants that uninformed people would have if they were well-informed about the actual characteristics of immigrants. Moreover, the present paper employed this method to assess "informed" attitudes toward asylum seekers who immigrated during the refugee crisis to Europe.

The results indicate that misperceptions cannot explain opposition to immigration by asylum seekers. Even when well-informed, a large majority of Europeans oppose the immigration of most who apply for asylum. Still, acceptance strongly depends on the characteristics of asylum seekers. A crucial distinction is whether asylum seekers migrated due to fear of war or other reasons, like economic opportunities. In some countries, a slight majority of citizens accept immigration from those fleeing wars, while in other countries, a slight majority opposes such immigration. In contrast, most citizens of all countries and social groups that I considered oppose immigration by asylum seekers who do not flee wars or persecution.

Hence, policies meant to hinder immigration by asylum seekers who do not flee wars, or persecution are relatively likely to receive much support from majorities in all European countries and social groups. Thus, such policies might be relatively easy to implement. In contrast, policies aimed at taking in war refugees are likely controversial and hard to implement.

Appendix 2.A   Demographic Characteristics of Asylum Seekers

This section assesses differences in the demographic composition of accepted and rejected asylum seekers (AS). Although the latter groups are not necessarily identical to HAS and OAS, respectively, they are related to them. Accepted AS are considered HAS by the host country, while rejected AS are considered OAS by the host country. To assess differences in the demographic structure of accepted and rejected asylum seekers, I use data from Eurostat if available. I pool data for the years 2015-2020 to focus on applications made by those who immigrated during the refugee crisis. During this time period, about 48% of applicants were granted some form of positive protection status, while the remaining applications were rejected.

Figure 2.A.1 depicts the shares of males and females among accepted and rejected, respectively, in the form of a stacked bar chart. The majority of AS are male. The fraction of males among the accepted AS is slightly lower among the accepted than among the rejected AS (62% vs. 67%).
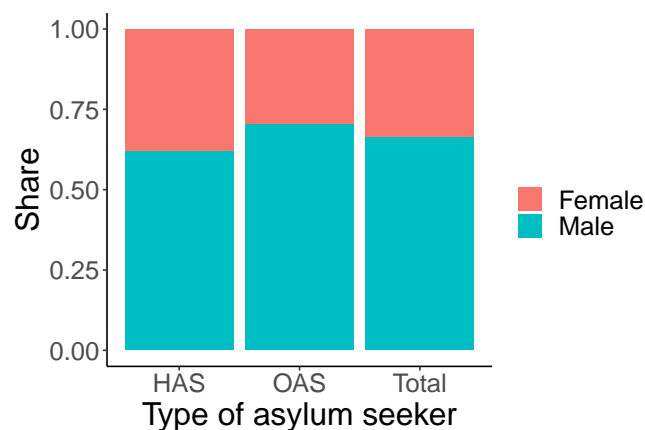
Figure 2.A.1. Shares of Males and Females among Accepted and Rejected Asylum Seekers
Note: The vertical axis depicts the shares of males and females among accepted and rejected asylum seekers who applied for asylum in Europe between 1.1.2015 and 31.12.2020.

Figure 2.A.2 depicts information on the age distributions. As can be seen from the pooled distribution on the left, most AS are young, the majority being not older than 34, and only very few being older than 64. Moreover, a large share of AS was underage when their asylum application was handed in. Age distributions of rejected and accepted asylum seekers are similar. There are more minors among the accepted AS than the rejected AS (40% of the accepted and 31% of the rejected). Conversely, more young adults (18-34) are among the rejected than among the accepted (49% vs. 42%). The share of older people among both groups is again similar.

In Figure 2.A.3 I compare religious denominations by AS group. Because Eurostat does not provide data on religious beliefs, this data is taken from a large-scale survey among AS in Germany. More details on the calculation can be found in Sec-
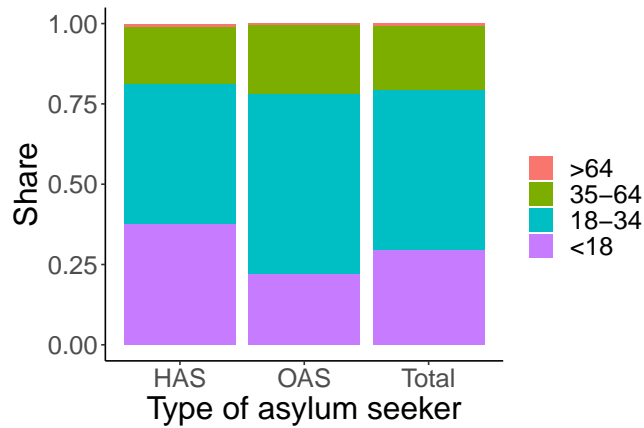
Figure 2.A.2. Age Distributions of Accepted and Rejected Asylum Seekers
Note: Stacked bar charts illustrate the age distributions of accepted and rejected asylum seekers who applied for asylum in Europe between 1.1.2015 and 31.12.2020.

tion 2.C. As apparent from the figure, Muslims are a majority among both accepted and rejected AS. However, they make up a larger share of the accepted (about 73%) than the rejected (about 60%). Sizable shares of either group are Christians (about 15% of the accepted and 31% of the rejected), while those without a confession form small minorities in either group. Hence, the two groups of AS differ more regarding religious beliefs than age and gender.
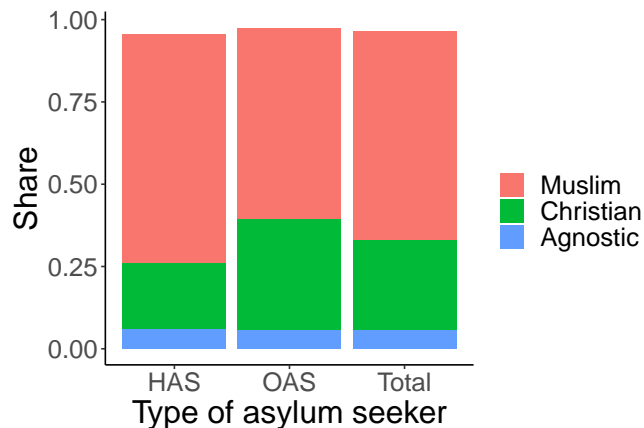
Figure 2.A.3. Frequencies of Religious Beliefs of Accepted and Rejected Asylum Seekers
Note: Stacked bar charts illustrate the distributions of religious beliefs of accepted and rejected asylum seekers who applied for asylum in Europe between 1.1.2015 and 31.12.2020.

Finally, Figure 2.A.4 illustrates distributions regarding countries of origin. More than half of all accepted AS came from Syria. Other notable countries of origin of the accepted AS include Afghanistan, Iraq, and Eritrea. In total, more than 85% of accepted AS originate from these four countries. Notably, these countries either suffered from civil wars (Syria, Afghanistan, and Iraq) or repressive dictatorships

(Eritrea) in the time frame under consideration. In contrast, most rejected AS do not originate from these four countries. For instance, only about 5% of rejected AS came from Syria. In general, the distribution of rejected AS seems to be much less concentrated in a few countries. Instead, most of them come from many (mostly developing) countries, and the number of rejected AS from these countries is often only in the thousands. These findings are consistent with European Migration Agencies classifying AS correctly. Civil wars do not necessarily involve the whole country, especially in countries that traditionally lack a strong central authority as many middle eastern countries do. Hence, a civil war in one's country does not necessarily mean that one is affected. Therefore, not all migrants from these countries should be considered HAS. Still, one would expect that accepted AS primarily originates from countries suffering from military conflicts. Among the rejected AS, in contrast, countries affected by conflict should not be over-represented strongly.
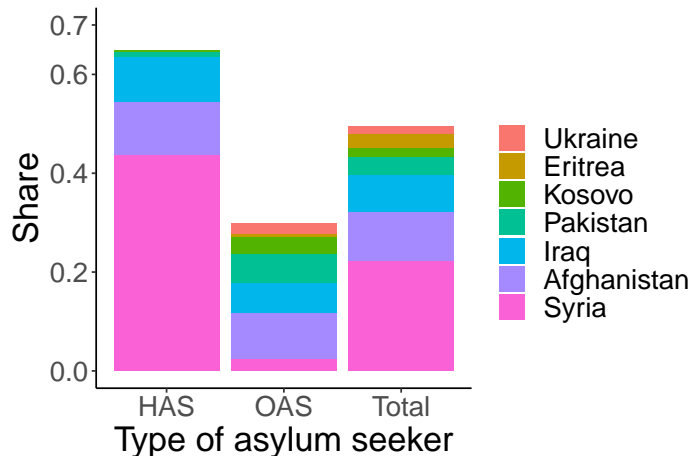
Figure 2.A.4. Countries of Origin of Accepted and Rejected Asylum Seekers
Note: Stacked bar charts illustrate the distributions of countries of origin of accepted and rejected asylum seekers who applied for asylum in Europe between 1.1.2015 and 31.12.2020. Shares do not sum to one because not all countries of origin are included in the figure.

In sum, the data suggests that demographic differences between accepted and rejected AS are moderately large. Although they do originate from different countries, most countries of origin for either group are developing countries. Differences regarding age and gender are small, while differences regarding religious beliefs are moderately large.

Appendix 2.B   More on How to Measure the Acceptance Rate

2.B.1   Distinguishing Humanitarian from Other Asylum Seekers

Figure 2.B.1 is taken from Bansak, Hainmueller, and Hangartner (2016). They use their data to estimate OLS regressions of the form

$$\mathbb{1}[r \text{ accepts } a]_{r,a} \; = \; \alpha + \beta \cdot X_a + \varepsilon_{r,a}$$

where $\mathbb{1}[r \text{ accepts } a]_{r,a}$ indicates whether survey respondent r accepted the immigration of asylum seeker a to his country. X_a is a matrix that contains all potential characteristics of a shown in Table 2.E.1 and Table 2.E.2 as dummy variables (excluding one for each variable as reference category), and β is the vector of corresponding coefficients.

These coefficients are shown with 95% confidence intervals pooled for all countries in Figure 2.B.1. For instance, the figure shows that the average European is ten percentage points less likely to accept Muslim asylum seekers than Christian asylum seekers.

I use the "reason for migrating" characteristic to distinguish HAS from OAS. This variable can take on the following values: "migrated for economic reasons," migrated due to ethnic persecution," "migrated due to religious persecution," and "migrated due to political persecution." I define an asylum seeker as an OAS if he migrated for economic reasons and be a HAS otherwise. Deciding on which form of persecution to use to define HAS is more complicated. HAS might migrate for either form of persecution. However, Figure 2.B.1 shows that for Europeans, it doesn't matter much whether a migrant is persecuted for his political beliefs, ethnicity, or religious affiliation. The estimates are quantitatively very similar and not significantly different from each other. For simplicity, I assume that all HAS are persecuted for their political beliefs.

2.B.2   The (Ir)relevance of Interaction Effects

The estimation of informed attitudes in Section 2.3 focuses on the main effects on characteristics and omits interaction effects between them. This simplification is necessary to reduce the required computing power to make the analysis feasible on ordinary computers. However, excluding interaction effects might bias the results if they are quantitatively large and significant. To assess whether that is the case, I run the following regression:

$$\mathbb{1}[r \text{ accepts } a]_{r,a} \; = \; a + \sum_{j \in J} b_j \cdot \mathbb{1}_{a,j} + \sum_{k \in J \neq j} b_{j,k} \cdot \mathbb{1}_{a,j} \cdot \mathbb{1}_{a,k} + \epsilon_{r,a}, \qquad (2.B.1)$$

where $\mathbb{1}[r \text{ accepts } a]_{r,a}$ is an indicator for whether respondent r accepted asylum seeker a, J is a set containing all potential asylum seeker characteristics (excluding the reference categories) and $\mathbb{1}_{a,j}$ is an indicator for whether asylum seeker a has characteristic j. In contrast to Equation 2.4 estimated in the main part, Equation 2.B.1 also includes all second-order interaction effects between characteristic indicators, excluding the reference categories.

Figure 2.B.2 shows two histograms. The blue histogram shows the distribution of the coefficients on the main effects from Equation 2.B.1. As can be seen, the majority of interaction effects have a coefficient very close to zero, and few interaction
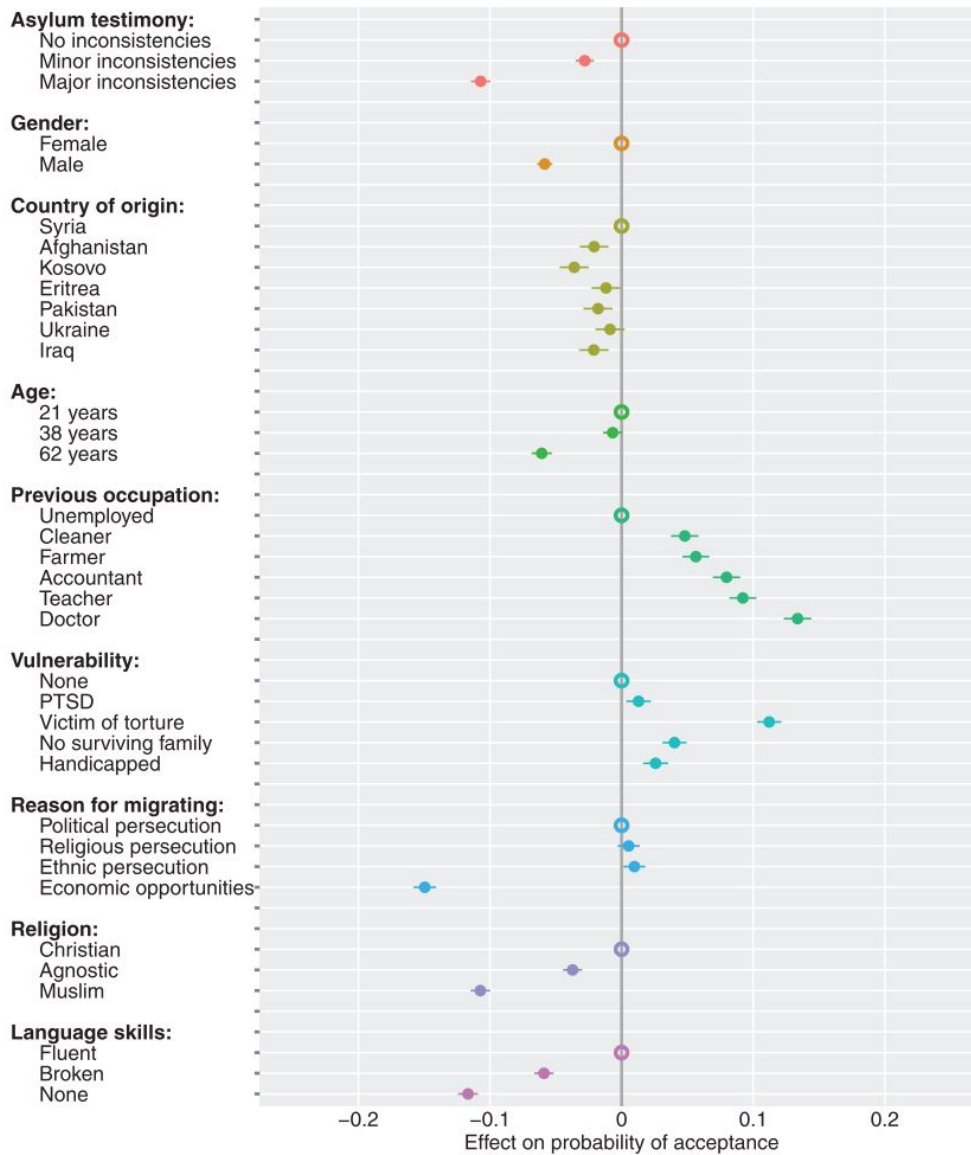
Figure 2.B.1. Effects of Asylum Seeker Characteristics on the Probability to Be Accepted by Europeans

Note: This figure is taken from Bansak, Hainmueller, and Hangartner (2016). Characteristics as shown on the vertical axis. For each variable, one characteristic is omitted as the reference category. The horizontal axis shows the effect that an asylum seeker characteristic has on the probability of being accepted by Europeans relative to the reference characteristic.

effects are large. The distribution of coefficients on main effects looks very different. Relatively few observations are close to zero, and the distribution is much wider than that of the interaction effects. This suggests that interaction effects are unimportant compared to the main effects quantitatively.
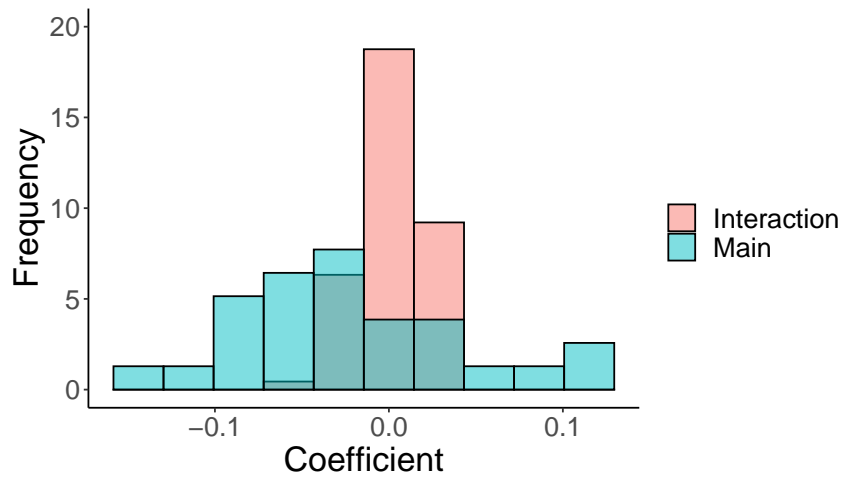
Figure 2.B.2. Histograms of Coefficients of Main- And Interaction Effects
Note: The horizontal axis shows the size of regression coefficients. The blue bars form a histogram of coefficients on all main effects included in Equation 2.B.1. The red bars form a histogram of coefficients on interaction effects included in Equation 2.B.1.
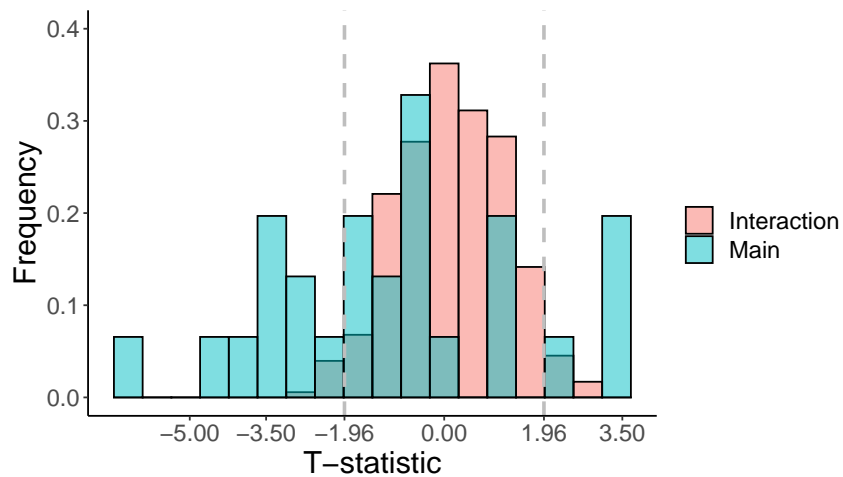
Figure 2.B.3. Histograms of T-Statistics of Main- And Interaction Effects
Note: The horizontal axis shows the size of t-statistics. The blue bars form a histogram of t-statistics of all main effects included in Equation 2.B.1. The red bars form a histogram of t-statistics of all interaction effects included in Equation 2.B.1.

Figure 2.B.3 presents a similar exercise regarding statistical significance. It depicts histograms of t-statistics instead of coefficients. Vertical lines indicate the thresholds for significance at the 5% level. The histogram for the t-statistics of the interaction effects is bell-shaped and centered around its mode at zero. 6% of the interaction effects are significant at the 5% level of significance, slightly more than would be expected due to chance. In contrast, the histogram for t-statistics of main effects does not resemble a normal distribution. Its mode is below zero, and t-statistics

are much more dispersed than those of the interaction effects. Much of the distribution lies outside the vertical lines at 1.96 and -1.96. Exactly half of all main effects are significant at the 5% level.

Hence, the main effects are not only larger than interaction effects, but they are also more frequently statistically significant. Especially the fact that interaction effects are only slightly more often significant than expected under chance suggests that omitting them from the analysis does not strongly bias the results.

Appendix 2.C   Matching Characteristics of Real Asylum Seekers to the Experimental Data

This section explains how I match the experimental data with the data on real AS characteristics. To infer subjects' attitudes toward real asylum seekers from their attitudes toward hypothetical asylum seeker vignettes elicited by Bansak, Hainmueller, and Hangartner (2016), I weigh their attitudes toward hypothetical vignettes with the frequencies of real asylum seekers with the same characteristics as the vignettes. To this end, I construct the asylum seeker type-vectors t_{HAS}, t_{OAS} and t_{AS} (used in Equation 2.3), which contain the frequencies of characteristics of true asylum seekers. In the following, I consider all characteristics in order and explain how I match vignette descriptions to real data for each variable.

Bansak, Hainmueller, and Hangartner (2016) distinguish between males and females and randomly assign each gender with equal probability to the vignettes. To match data on real asylum seeker characteristics to the hypothetical vignettes, the type vectors need to contain the true frequencies of males and females among asylum seekers. To this end, I employ data from Eurostat described in more detail in Section 2.A. The data contains the shares of males and females among accepted and rejected asylum seekers who applied for asylum in any of the (old) 28 EU countries. Figure 2.A.1 shows these shares among all first-time asylum applicants who applied for asylum between 2015 and 2020. In total, about 62% of all rejected AS that came to Europe between 2015 and 2020 were men. The share of males is a bit higher among the OAS at about 67%. Hence, for instance, t_{OAS} = (..., share males among OAS, ...) = (..., 0.67,...) and t_{AS} = (..., share males among HAS, ...) = (..., 0.62,...).

To estimate the impact of age on the acceptance of natives, Bansak, Hainmueller, and Hangartner (2016) randomly assigned one of three age categories (21 years, 38 years, 62 years) to the vignettes. I compile corresponding data of real asylum seekers from Eurostat again. Eurostat provides the shares of several age groups among asylum seekers who applied in any of the (old) 28 EU countries between 2015 and 2020. Figure 2.A.2 shows these shares for the total of asylum seekers, HAS, and OAS. The age groups (18-34-year-old; 35-64-year-old; 65+ year-old) broadly fit the categories distinguished by Bansak, Hainmueller, and Hangartner (2016). Consequently, I match the age categories from Bansak, Hainmueller, and Hangartner (2016) and

the age ranges from Eurostat in the following way: I match 21-year-old and 18-34-year-old, 35-64-year-old and 38-year-old and finally 65+-year-old and 62-year-old. It is hard to fit in the children in this picture because Bansak, Hainmueller, and Hangartner (2016) did not ask the respondents to rate children. Therefore I focus on attitudes toward adults. Hence, to fill t_{HAS}, t_{OAS}, and t_{HAS}, I calculate the shares of the three age groups from the Eurostat data among the adults for each group.

Bansak, Hainmueller, and Hangartner (2016) also inform subjects about the country of origin of the immigrant. They randomly assign the vignettes to one out of seven potential countries of origin: Syria, Afghanistan, Kosovo, Eritrea, Pakistan, Ukraine, and Iraq. Figure 2.B.1 shows how each of these categories affects the acceptance rate. Syrian origin is the omitted category. Europeans generally have a higher likelihood of accepting Syrian migrants. On the other hand, Europeans do not seem to distinguish between the other countries, which strengthens the idea that it's mainly the special inhumane conditions in Syria at the time and their salience that lead to the greater acceptance of Syrians. Hence, it seems to be most relevant to distinguish between Syrians and non-Syrians.

Figure 2.A.4 shows the distribution of country of origins of HAS, OAS, and all AS. It is based on data from Eurostat. I use this data to calculate relative frequencies for each country of origin distinguished by Bansak, Hainmueller, and Hangartner (2016) for AS, HAS, and OAS separately. While few HAS originated from countries not shown to subjects by Bansak, Hainmueller, and Hangartner (2016), about 65% of OAS did. Because Europeans do not strongly distinguish between countries of origin other than Syria, I assign to all AS who migrated from a country that is not in the Bansak, Hainmueller, and Hangartner (2016) data the mean of the coefficients of all non-Syrian countries.

Bansak, Hainmueller, and Hangartner (2016) randomly assigned one of six previous occupations to the vignette: unemployed, cleaner, farmer, accountant, teacher, and doctor. Eurostat does not provide information on the occupation of asylum seekers. The best data I am aware of comes from Schröder, Zok, and Faulbaum (2018). He presents results from a survey among 1,880 asylum seekers in Germany, which has been by far the most frequent destination country for asylum seekers during the refugee crisis. I take frequencies of the AS's former occupations from figure 1 in the paper and match these categories manually to the categories used by Bansak, Hainmueller, and Hangartner (2016). Both papers contain an Unemployment-category which I match. I consider the "Cleaner" and "Farmer" categories from Bansak, Hainmueller, and Hangartner (2016) as examples for blue-collar workers, and I assume that most blue-collar occupations would have similar effects on European acceptance rates (the effects of the cleaners and workers dummy are very similar). I consider the following occupations from the Schröder, Zok, and Faulbaum (2018) blue-collar and occupations with similar skill levels: craftsmanship, Support activities (industrial, factory, and warehouse work), Transport (logistics and traffic), Agriculture, forestry, fishing, Engineering (IT, electrical professions, technical professions), con-

struction, street trading, Motor vehicle sector, soldier housework, retirement, others (given that blue-color is by far the most frequent occupation and only 2.4% are in that category).

I consider the categories accountant a typical white-collar occupation and match it to the following white-collar occupations from Schröder, Zok, and Faulbaum (2018): Office, banks, insurance, Public service, Hotel, and restaurant industry, church, public safety, services, Small-scale self-employment, arts. Moreover, I match the teacher category from Bansak, Hainmueller, and Hangartner (2016) with the "Teaching professions"-category from Schröder, Zok, and Faulbaum (2018). Lastly, I match the doctor category from Bansak, Hainmueller, and Hangartner (2016) with the categories Medical professions, chemical professions, laboratory and Management positions, and company management because these occupations are prestigious and require a high skill level.

To estimate the real share of asylum seekers belonging to the larger Bansak, Hainmueller, and Hangartner (2016) categories, I sum up the relative frequencies of their respective finer sub-categories used by Schröder, Zok, and Faulbaum (2018). Information on these sub-categories is provided in figure 1 in Schröder, Zok, and Faulbaum (2018). This procedure enables classifying about 84% of asylum seekers (I cannot classify students). As Schröder, Zok, and Faulbaum (2018) do not distinguish between HAS and OAS, I use the same weights for all types of AS.

The variable religion is captured by Bansak, Hainmueller, and Hangartner (2016) by a categorical variable that can take on one of three values; Muslim, Christian, and agnostic. Because Eurostat does not provide data on religious affiliation, I take data from Siegert (2020). He provides information on the religious affiliation of a sample of those who applied for asylum in Germany between 2013 and 2016 by country of origin but does not distinguish between rejected and accepted asylum seekers. However, HAS and OAS might differ regarding their religious affiliations. To derive different distributions for the two groups, I use the fact that HAS, OAS, and AS differ mainly regarding their country of origin. Moreover, the religious affiliation of asylum seekers varies strongly by country of origin. Hence, I estimate the share of a religious belief rel $\in$ {Muslim, Christian, Agnostic} among the group $g \in$ {AS, HAS, OAS} as

$$
\begin{aligned}
share_{rel,g} = &\frac{\#\text{Syrians}_g}{||g||} \cdot \frac{\#\text{rel Syrian AS}}{\#\text{Syrian AS}} + \frac{\#\text{Afghans}_g}{||g||} \cdot \frac{\#\text{rel Afghans AS}}{\#\text{Afghan AS}_g} + \\
&\frac{\#\text{Iraqi}_g}{||g||} \cdot \frac{\#\text{rel Iraqi AS}}{\#\text{Iraqi AS}} + \frac{\#\text{Eritreans}_g}{||g||} \cdot \frac{\#\text{rel Eritrean AS}}{\#\text{Eritreans AS}} + \\
&\frac{\#\text{AS with Other Nationalities}_g}{||g||} \cdot \frac{\#\text{rel AS with Other Nationalities}}{\#\text{AS with Other Nationalities}}.
\end{aligned}
$$

Data on the national group sizes and the total number of AS are taken from Eurostat. The shares of the religious denominations by nationality of the four main countries of origin are taken from figure 5 in Siegert (2020). To calculate the

share of the religious beliefs among AS from other countries I combine this data with data on the relative size of the countries of origin in the IAB-BAMF-SOEP-dataset taken from Brücker, Rother, and Schupp (2018). In particular I calculate $x := \dfrac{\text{\# AS with Other Nationalities and religion rel}}{\text{\# AS with Other Nationalities}}$ from the following equation:

$$\frac{\#relAS}{\#\text{ Total AS}} = \frac{\#\text{Syrians}_g}{\#\text{ Total AS}} \cdot \frac{\#\text{rel Syrian AS}}{\#\text{Syrian AS}} + \frac{\#\text{Afghans}_g}{\#\text{ Total AS}} \cdot \frac{\#\text{rel Afghans AS}}{\#\text{Afghan AS}_g}$$

$$\frac{\#\text{Iraqi}_g}{\#\text{ Total AS}} \cdot \frac{\#\text{rel Iraqi AS}}{\#\text{Iraqi AS}} + \frac{\#\text{Eritreans}_g}{\#\text{ Total AS}} \cdot \frac{\#\text{rel Eritrean AS}}{\#\text{Eritreans AS}} +$$

$$\frac{\#\text{AS with Other Nationalities}}{\#\text{ Total AS}} \cdot x.$$

This procedure makes it possible to classify the vats majority of accepted and rejected AS (94.11% and 96.76% respectively) as either Christian, Muslim or Agnostic. In order to preserve the upper-bound property of my estimates for the acceptance rates I assume that all non-classified AS are Christians, given that they are the group most preferred by Europeans.

Another proxy for the skill level is the ability to speak the language of the host country. As can be seen in Figure 2.B.1, Europeans indeed strongly favor those that speak the host country's language fluently over those that do not or are only broken. Again, data on that variable is rather sparse. To my knowledge, the best available evidence comes from Brücker, Rother, and Schupp (2018), which estimates that over 90% of AS had no knowledge of German when they arrived. This puts more than 90% into the worst category as measured by the preferences of the Europeans. Again, this evidence does not distinguish between HAS and OAS. I assume that the remaining 10% spoke the host country's language fluently to ensure that my estimates constitute upper bounds for the true acceptance rate. Regarding this variable, I do not distinguish between accepted and rejected AS.

Bansak, Hainmueller, and Hangartner (2016) describe asylum seekers' vulnerability through five categories. They distinguish between asylum seekers who have post-traumatic stress disorder, have been tortured, have no surviving family, or are handicapped. Compiling objective data on the frequencies of these vulnerabilities among asylum seekers is near impossible because of the confusing circumstances in many countries of origin and flight routes. In the absence of objective measures, I use data on subjective assessments of HAS reported by Schröder, Zok, and Faulbaum (2018). They report results from interviews conducted among 2,021 refugees who migrated to Germany between 2017 and 2018 regarding their vulnerability. Crucially, they were asked about their vulnerability in a way that makes their responses comparable to the categories distinguished by Bansak, Hainmueller, and Hangartner (2016).

The main drawback of this type of data is that refugees are likely to exaggerate their vulnerability to generate more sympathy for them or because they fear that

the interviews might be used to determine whether their applications are justified. Hence, taking the data at face value likely exaggerate the vulnerability of asylums seekers which biases acceptance rates upwards because Europeans are more likely to accept vulnerable asylums seekers as apparent from Figure 2.B.1. Therefore, using this data preserves the fact that estimates are an upper bound for the true acceptance rates.

Schröder, Zok, and Faulbaum (2018) ask asylum seekers whether they made war experiences, were attacked by the military, had disappeared relatives, experienced violence in connection with their flight, experienced torture, had been in other life-threatening situations, had been in camp detention, or hostage detention or incommunicado detention, witnessed torture, killings or sexual violence, whether significant others died violently, whether they experienced other seriously stressful events, whether they were sexually assaulted or raped and finally whether they made none of these experiences. I take the shares of refugees who state that they had made these experiences from figure 2 in their paper.

To associate the vignette data of Bansak, Hainmueller, and Hangartner (2016) with the data on real characteristics of Schröder, Zok, and Faulbaum (2018) I match the following categories: the "None" category from Bansak, Hainmueller, and Hangartner (2016) with the "None of that" category from Schröder, Zok, and Faulbaum (2018), the no surviving family" category from Bansak, Hainmueller, and Hangartner (2016) with the "death of significant others" and "disappeared relatives" categories from Schröder, Zok, and Faulbaum (2018)[6], and the "torture" category from Bansak, Hainmueller, and Hangartner (2016) with the "experience of torture" category from Schröder, Zok, and Faulbaum (2018). Moreover, I assume that every asylum seeker who did report making any of the traumatising experiences elicited in Schröder, Zok, and Faulbaum (2018) has post traumatic stress disorder. Regarding the "handicapped" category used by Bansak, Hainmueller, and Hangartner (2016) I use the fact that 0.2924% of refugees interviewed by Schröder, Zok, and Faulbaum (2018) reported visiting a doctor due to a handicap.

OAS are asylum seekers that asylum ministries classify as having not migrated due to a threat of violence. Hence, I assume that the vulnerability of OAS does not differ notably from the vulnerability of ordinary citizens. Hence, I assume that all OAS have no vulnerability. Certainly, some OAS are disabled or have no surviving family left. However, the vulnerabilities were presented in the context of asylum seekers. Hence, survey respondents likely interpreted these as special vulnerabilities that resulted from war or persecution in the asylum seekers' homeland, which OAS do not have by definition.

---

6. Asylum seekers might have surviving family members even though some significant others died and some of their relatives disappeared. Hence, the category used by Bansak, Hainmueller, and Hangartner (2016) expresses a more severe vulnerability which preserves the upper bound property of the acceptance rate estimates.

The last variables to be discussed are the reason for migrating and the consistency of the asylum testimony. Both variables are closely related to whether an AS is a HAS or an OAS. I assume that all OAS migrated due to economic opportunities, which is the only non-humanitarian reason given by Bansak, Hainmueller, and Hangartner (2016) and probably the most important non-humanitarian reason for AS to come to Europe. Note that the coefficients for all other reasons for migrating are very similar. Without good data to rely on which of these forms of persecution (political, religious, ethnic) is most relevant (such data is probably hard to generate given that these categories are not exclusive and will often induce each other), I take the mean of all three coefficients and add it to the acceptance rate of HAS. Lastly, I assume that all asylum testimonies of HAS have no consistency. This is likely because inconsistencies should not arise if an AS is truly a HAS. After all, such an AS should have a consistent testimony by simply stating what has happened to him. On the other hand, OAS could not hand in a testimony with any chance of being accepted that is consistent and true. I, therefore, assume that all testimonies of OAS have major inconsistencies.

Appendix 2.D   Additional Figures

ATTRIBUTE	APPLICANT 1	APPLICANT 2
Age	21 Years	62 Years
Language Skills	Speaks broken English	Speaks fluent English
Previous Occupation	Unemployed	Teacher
Religion	Christian	Muslim
Consistency of Asylum Testimony	Minor inconsistencies	Major inconsistencies
Vulnerability	Post-traumatic stress disorder (PTSD)	No surviving family members
Origin	Iraq	Pakistan
Reason for Migrating	Seeking better economic opportunities	Persecution for ethnicity
Gender	Male	Male

Figure 2.D.1. Two Examples for the Vignettes Used By Bansak, Hainmueller, and Hangartner (2016)
Note: This figure is taken from Bansak, Hainmueller, and Hangartner (2016). It shows vignettes for two hypothetical asylum seekers as shown to the subjects in Bansak, Hainmueller, and Hangartner (2016). The leftmost column shows variables. The other two columns show the characteristics that the two asylum seekers have regarding these variables.
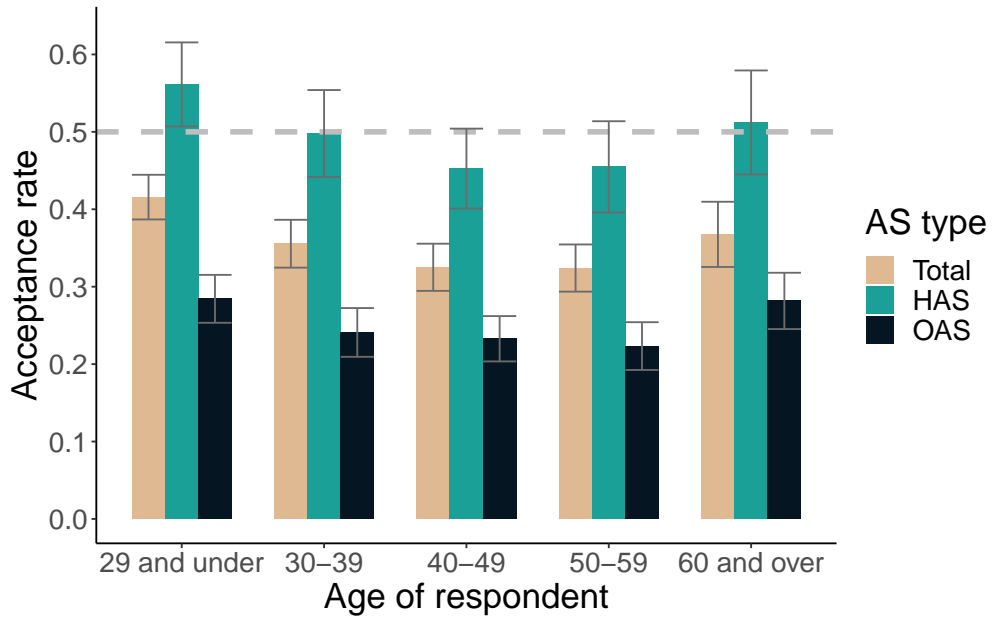
Figure 2.D.2. Acceptance Rates for the Average Asylum Seeker by Type and Age Group
Note: The vertical axis depicts acceptance rates for the average total of asylum seekers, humanitarian asylum seekers, and other asylum seekers, respectively, by age group of the survey respondent. I also depict 95% confidence intervals. Standard errors are clustered at the spectator level. Data is pooled across Europe.
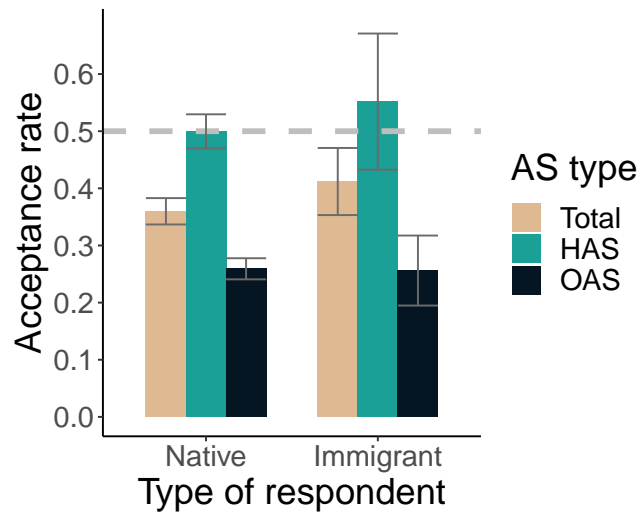
Figure 2.D.3. Acceptance Rates for the Average Asylum Seeker by Type and Immigration Background
Note: The vertical axis depicts acceptance rates for the average total of asylum seekers, humanitarian asylum seekers, and other asylum seekers, respectively, separately for survey respondents who were born in their country of residence and those that were not. I also depict 95% confidence intervals. Standard errors are clustered at the spectator level. Data is pooled across Europe.
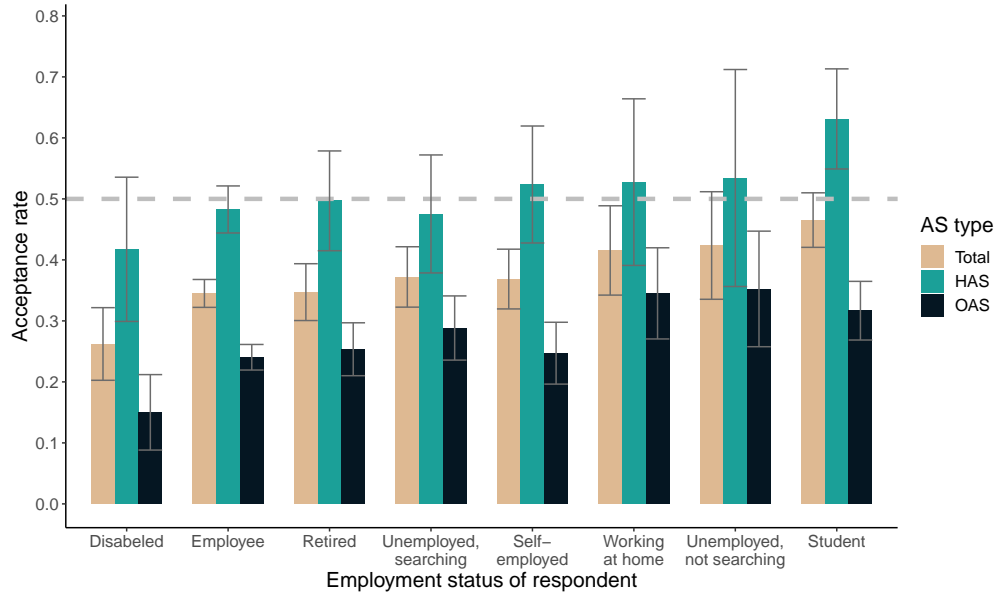
Figure 2.D.4. Acceptance Rates for the Average Asylum Seeker by Type and Occupation
Note: The vertical axis depicts acceptance rates for the average total of asylum seekers, humanitarian asylum seekers, and other asylum seekers, respectively, by occupation category of the survey respondent. I also depict 95% confidence intervals for both groups and all countries. Standard errors are clustered at the spectator level. Data is pooled across Europe.
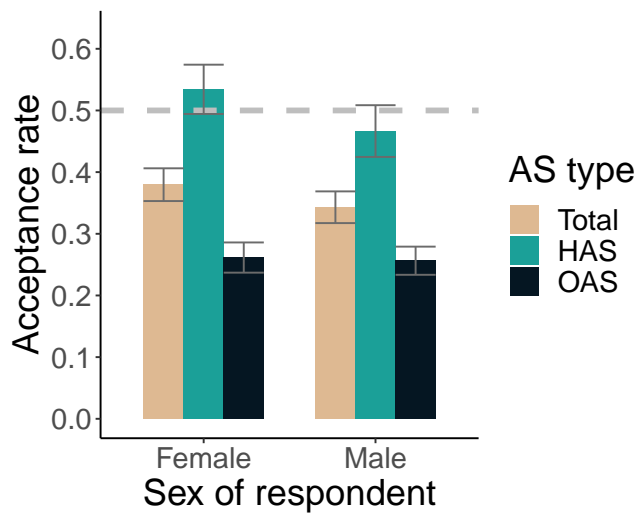
Figure 2.D.5. Acceptance Rates for the Average Asylum Seeker by Type and Sex
Note: The vertical axis depicts acceptance rates for the average total of asylum seekers, humanitarian asylum seekers, and other asylum seekers, respectively, by sex of the survey respondent. I also depict 95% confidence intervals. Standard errors are clustered at the spectator level.

Appendix 2.E   Additional Tables

Table 2.E.1. Possible Characteristics of Asylum Seeker Vignettes (I)

	Testimony consistency	**Gender**	**Country of Origin**	**Age**	**Previous occupation**
Value 1	No inconsistencies	Male	Syria	21	Unemployed
Value 2	Minor inconsistencies	Female	Afghanistan	38	Cleaner
Value 3	Major inconsistencies		Kosovo	62	Farmer
Value 4			Eritrea		Accountant
Value 5			Pakistan		Teacher
Value 6			Ukraine		Doctor
Value 7			Iraq		

Table 2.E.2. Possible Characteristics of Asylum Seeker Vignettes (II)

	Vulnerability	**Reason for migrating**	**Religion**	**Language skills**
Value 1	None	Political persecution	Christian	Fluent
Value 2	PTSD	Religious persecution	Agnostic	Broken
Value 3	Victim of Torture	Ethnic persecution	Muslim	None
Value 4	No surviving family	Economic opportunities		
Value 5	Handicapped			

References

Alesina, Alberto, Armando Miano, and Stefanie Stantcheva. 2023. "Immigration and Redistribution." *Review of Economic Studies* 90 (1): 1–39. DOI: 10.1093/restud/rdac011. (Visited on 03/27/2023). [101, 104]

Andreadis, Ioannis, and Yannis Stavrakakis. 2017. "European Populist Parties in Government: How Well are Voters Represented? Evidence from Greece." en. *Swiss Political Science Review* 23 (4): 485–508. DOI: 10.1111/spsr.12255. (Visited on 03/19/2023). [103]

Babka von Gostomski, Christian, Axel Böhm, Herbert Brücker, Tanja Fendel, Martin Friedrich, Marco Giesselmann, Elke Holst, Yuliya Kosyakova, Martin Kroh, Elisabeth Liebau, David Richter, Agnese Romiti, Nina Rother, Diana Schacht, Jana A. Scheible, Paul Schmelzer, Jürgen Schupp, Manuel Siegert, Steffen Sirries, Parvati Trübswetter, and Ehsan Vallizadeh. 2016. "IAB-BAMF-SOEP-Befragung von Geflüchteten: Überblick und erste Ergebnisse." ger. Research Report 14/2016. IAB-Forschungsbericht. URL: https://www.econstor.eu/handle/10419/149 618 (visited on 04/02/2023). [106]

Bakker, Ryan, Seth Jolly, and Jonathan Polk. 2022. "Analyzing the Cross-National Comparability of Party Positions on the Socio-Cultural and EU Dimensions in Europe." en. *Political Science Research and Methods* 10 (2): 408–18. DOI: 10.1017/psrm.2020.26. (Visited on 03/19/2023). [110]

Bansak, Kirk, Jens Hainmueller, and Dominik Hangartner. 2016. "How Economic, Humanitarian, and Religious Concerns Shape European Attitudes toward Asylum Seekers." en. *Science* 354 (6309): 217–22. DOI: 10.1126/science.aag2147. (Visited on 03/19/2023). [102, 104–107, 109, 119, 121, 123–129]

Barrera, Oscar, Sergei Guriev, Emeric Henry, and Ekaterina Zhuravskaya. 2020. "Facts, Alternative Facts, and Fact Checking in Times of Post-truth Politics." en. *Journal of Public Economics* 182 (2): 104123. DOI: 10.1016/j.jpubeco.2019.104123. (Visited on 03/19/2023). [104]

Blinder, Scott. 2015. "Imagined Immigration: The Impact of Different Meanings of 'Immigrants' in Public Opinion and Policy Debates in Britain." en. *Political Studies* 63 (1): 80–100. DOI: 10.1111/1467-9248.12053. (Visited on 03/19/2023). [103]

Brücker, Herbert, Nina Rother, and Jürgen Schupp, editors. 2018. *IAB-BAMF-SOEP-Befragung von Geflüchteten 2016: Studiendesign, Feldergebnisse sowie Analysen zu Schulischer wie Beruflicher Qualifikation, Sprachkenntnissen sowie Kognitiven Potenzialen.* de. Vol. 30, Forschungsbericht / Bundesamt für Migration und Flüchtlinge (BAMF) Forschungszentrum Migration, Integration und Asyl (FZ). Nürnberg: Bundesamt für Migration und Flüchtlinge (BAMF) Forschungszentrum Migration, Integration und Asyl (FZ). [106, 126]

Brücker, Herbert, Nina Rother, Jürgen Schupp, Christian Babka von Gostomski, Axel Böhm, Tanja Fendel, Martin Friedrich, Marco Giesselmann, Yuliya Kosyakova, Martin Kroh, Simon Kühne, Elisabeth Liebau, David Richter, Agnese Romiti, Diana Schacht, Jana A. Scheible, Paul Schmelzer, Manuel Siegert, Steffen Sirries, Parvati Trübswetter, and Ehsan Vallizadeh. 2016. "Forced Migration, Arrival in Germany, and First Steps toward Integration." eng. *DIW Economic Bulletin* 6 (48): 541–56. URL: https://www.econstor.eu/handle/10419/148611 (visited on 04/02/2023). Publisher: Berlin: Deutsches Institut für Wirtschaftsforschung (DIW). [106]

Ceobanu, Alin M., and Xavier Escandell. 2010. "Comparative Analyses of Public Attitudes Toward Immigrants and Immigration Using Multinational Survey Data: A Review of Theories and Research." *Annual Review of Sociology* 36 (1): 309–28. DOI: 10.1146/annurev.soc.012809.102 651. (Visited on 04/02/2023). _eprint: https://doi.org/10.1146/annurev.soc.012809.102651. [103]

Dalton, Russell J. 2017. "Party Representation across Multiple Issue Dimensions." en. *Party Politics* 23 (6): 609–22. DOI: 10.1177/1354068815614515. (Visited on 03/19/2023). [103]

Dustmann, Christian, Francesco Fasani, Tommaso Frattini, Luigi Minale, and Uta Schönberg. 2017. "On the Economics and Politics of Refugee Migration." *Economic Policy* 32 (91): 497–550. DOI: 10.1093/epolic/eix008. (Visited on 04/02/2023). [102, 103]

Esipova, Neli, Julie Ray, Anita Pugliese, Dato Tsabutashvili, Frank Laczko, and Marzia Rango. 2015. "How the World Views Migration." en. Working paper. Geneva: International Organization for Migration, 68. URL: https://publications.iom.int/books/how-world-views-migration (visited on 04/02/2023). [101]

Facchini, Giovanni, Yotam Margalit, and Hiroyuki Nakata. 2022. "Countering Public Opposition to Immigration: The Impact of Information Campaigns." en. *European Economic Review* 141 (1): 103959. DOI: 10.1016/j.euroecorev.2021.103959. (Visited on 04/02/2023). [104]

Grigorieff, Alexis, Christopher Roth, and Diego Ubfal. 2020. "Does Information Change Attitudes Toward Immigrants?" en. *Demography* 57 (3): 1117–43. DOI: 10.1007/s13524-020-00882-8. (Visited on 03/19/2023). [101, 104]

Günther, Laurenz. 2022. "Political Representation Gaps in Europe." en. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.4230288. (Visited on 04/12/2023). [103, 110]

Hainmueller, Jens, Dominik Hangartner, and Teppei Yamamoto. 2015. "Validating Vignette and Conjoint Survey Experiments against Real-World Behavior." en. *Proceedings of the National Academy of Sciences* 112 (8): 2395–400. DOI: 10.1073/pnas.1416587112. (Visited on 03/19/2023). [107]

Hainmueller, Jens, and Daniel J. Hopkins. 2014. "Public Attitudes Toward Immigration." en. *Annual Review of Political Science* 17 (1): 225–49. DOI: 10.1146/annurev-polisci-102512-194818. (Visited on 03/19/2023). [103]

Hix, Simon, Eric Kaufmann, and Thomas J. Leeper. 2021. "Pricing Immigration." en. *Journal of Experimental Political Science* 8 (1): 63–74. DOI: 10.1017/XPS.2020.6. (Visited on 03/19/2023). [103]

Hopkins, Daniel J., John Sides, and Jack Citrin. 2019. "The Muted Consequences of Correct Information about Immigration." en. *Journal of Politics* 81 (1): 315–20. DOI: 10.1086/699914. (Visited on 03/19/2023). [104]

Jacobsen, Jannes, Martin Kroh, Simon Kühne, Jana A. Scheible, Rainer Siegers, and Manuel Siegert. 2019. *Supplementary of the IAB-BAMF-SOEP Survey of Refugees in Germany (M5) 2017.* eng. SOEP Survey Papers. Berlin: Deutsches Institut für Wirtschaftsforschung (DIW). URL: https://hdl.handle.net/10419/194098 (visited on 04/02/2023). OCLC: 1189736273. [106]

Kaufmann, Eric. 2017. "Levels or Changes?: Ethnic Context, Immigration and the UK Independence Party Vote." en. *Electoral Studies* 48 (8): 57–69. DOI: 10.1016/j.electstud.2017.05.002. (Visited on 03/19/2023). [103]

Kaufmann, Eric. 2018. "Why Culture Is More Important than Skills: Understanding British Public Opinion on Immigration." *British Politics and Policy at LSE*, Publisher: London School of Economics and Political Science. [109]

Kaufmann, Eric. 2019. "Can Narratives of White Identity Reduce Opposition to Immigration and Support for Hard Brexit? A Survey Experiment." en. *Political Studies* 67 (1): 31–46. DOI: 10.1177/0032321717740489. (Visited on 03/19/2023). [104]

Kübler, Marie, and Armin Schäfer. 2022. "Closing the Gap? The Populist Radical Right and Opinion Congruence between Citizens and MPs." en. *Electoral Studies* 80 (12): 102527. DOI: 10.1016/j.electstud.2022.102527. (Visited on 04/02/2023). [103, 110]

Kustov, Alexander, Dillon Laaker, and Cassidy Reller. 2021. "The Stability of Immigration Attitudes: Evidence and Implications." *Journal of Politics* 83 (4): 1478–94. DOI: 10.1086/715061. (Visited on 03/27/2023). Publisher: The University of Chicago Press. [103]

Norris, Pippa, and Ronald Inglehart. 2019. *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. en. Cambridge University Press. Google-Books-ID: I8aGDwAAQBAJ. [115]

Schröder, Helmut, Klaus Zok, and Frank Faulbaum. 2018. "Gesundheit von Geflüchteten in Deutschland — Ergebnisse einer Befragung von Schutzsuchenden aus Syrien, Irak und Afghanistan." *Wido Monit* 15 (1): 1–20. [106, 124–127]

Siegert, Manuel. 2020. *Die Religionszugehörigkeit, Religiöse Praxis und Soziale Einbindung von Geflüchteten*. de. Vol. 2-2020, BAMF-Kurzanalyse. Nürnberg: Bundesamt für Migration und Flüchtlinge (BAMF) Forschungszentrum Migration, Integration und Asyl (FZ). [106, 125]

Skinner, Gideon, and Glenn Gottfried. 2017. "Global Views on Immigration and the Refugee Crisis." Working paper. Ipsos. [101]

Stecker, Christian, and Markus Tausendpfund. 2016. "Multidimensional Government-Citizen Congruence and Satisfaction with Democracy." en. *European Journal of Political Research* 55 (3): 492–511. DOI: 10.1111/1475-6765.12147. (Visited on 04/02/2023). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-6765.12147. [110]

Verkuyten, Maykel, Hadi Ghazi Altabatabaei, and Wybren Nooitgedagt. 2018. "Supporting the Accommodation of Voluntary and Involuntary Migrants: Humanitarian and Host Society Considerations." en. *Social Psychological and Personality Science* 9 (3): 267–74. DOI: 10.1177/1948550617737600. (Visited on 03/19/2023). [102]

Verkuyten, Maykel, Kieran Mepham, and Mathijs Kros. 2018. "Public Attitudes Towards Support for Migrants: The Importance of Perceived Voluntary and Involuntary Migration." en. *Ethnic and Racial Studies* 41 (5): 901–18. DOI: 10.1080/01419870.2017.1367021. (Visited on 03/19/2023). [102]

Wright, Matthew, Morris Levy, and Jack Citrin. 2016. "Public Attitudes Toward Immigration Policy Across the Legal/Illegal Divide: The Role of Categorical and Attribute-Based Decision-Making." en. *Political Behavior* 38 (1): 229–53. DOI: 10.1007/s11109-015-9311-y. (Visited on 03/19/2023). [102]

Chapter 3

Inherited Inequality and the Dilemma of Meritocracy*

Joint with Timo Freyer

3.1 Introduction

In a meritocratic society, inequality is considered to be just if it reflects factors within but not outside individuals' control. However, individuals are often not responsible for their outcomes themselves but benefit differentially from the efforts of others. For example, a child may be lucky to inherit abundant resources acquired by its parents, while another child is born into less favourable circumstances. Such *inherited inequality*[1] exposes a fundamental tension in the meritocratic logic. On the one hand, individuals are entitled to decide how to spend their earned resources, which includes the right to transfer them to others. On the other hand, if two individuals are not involved in the process that generates inequality between them, such inequality does not reflect their individual achievements. In the parent-child example, if one pair of parents works particularly hard such that their children "can have a better life", they have merited to see their child reap the benefits of their efforts.

1. With *inherited inequality* we refer to inequality between individuals that originates from the actions of others. Hence, we interpret the term "inherited" broadly. Our definition encompasses inequality between children who profit differentially from the actions of their parents, but also inequality between people who benefit to a differential extent from their friends, coworkers, or compatriots.

However one child has not merited more favourable circumstances than the other. By meritocratic standards, inherited inequality is just and unjust at the same time and confronts meritocrats with a dilemma—*the dilemma of meritocracy*.

The dilemma of meritocracy is central to various policy debates. Consider as an example the debate on the estate tax. Here, people who seemingly endorse the same fairness ideal—the meritocratic one—can end up taking diametrically opposed positions. Some contend that bequests are a result of the testator's hard work and usually conclude that it is unfair to redistribute. In this vein, it has been argued that "[s]ince the accumulation of a substantial estate is one of the motivations that drive people to work hard, a death tax on saving is indirectly a tax on work" (Posner, 2014). Other people stress that it was certainly not the heir's efforts that generated the bequest and label inheritances as unmerited income, concluding that it should be heavily taxed. For instance, US investor Warren Buffet is quoted in Obama (2006) saying that "[w]hen you get rid of the estate tax, you're basically handing over command of the country's resources to people who didn't earn it". Hence, the meritocratic fairness ideal is being used as a justification for policies at opposite ends of the political spectrum.

A potential explanation for this disagreement is that people differ in whether they prioritize meritocratic fairness toward the benefactors or the beneficiaries. Different priorities may, in turn, translate into different views on policies and demand for redistribution in the context of inherited inequality. To develop policies that are politically implementable and meet the fairness preferences of citizens, it is therefore necessary to better understand people's attitudes toward economic inequality and, in particular, how they deal with the dilemma of meritocracy.

In pursuit of this aim, this study introduces a stylized theoretical framework that formalizes how individuals evaluate (inherited) unequal distributions and reports results from a survey experiment that puts its predictions to the test. The framework covers situations in which money is distributed between two individuals who each benefit from the effort of an associated worker. An impartial spectator observes this situation and makes a fairness judgment based on his or her fairness ideal. This setup nests the case of noninherited inequality, where a beneficiary and the associated worker are identical and, therefore, being fair toward workers is the same as being fair toward beneficiaries. If beneficiaries and their associated workers are not identical, however, meritocrats need to balance two potentially conflicting fairness views: if the two workers exert different levels of effort, the distribution that is considered fair toward the two workers may be different from the distribution that is considered fair toward the two beneficiaries, who both exert no effort. Given that fairness toward the workers calls for no redistribution whereas fairness toward the beneficiaries demands full equalization, individuals face a dilemma because they infringe meritocratic fairness no matter how they redistribute. Because beneficiaries merit similar but inherit different outcomes, meritocrats may be less willing to accept inherited inequality as compared to noninherited inequality.

The corresponding experiment builds on the impartial spectator paradigm (Konow, 2000; Cappelen, Konow, et al., 2013) and consists of two stages. In the earnings stage, an initial distribution of $10 between two stakeholders is determined. In the first of two treatment dimensions, we vary whether the two stakeholders themselves work on a real-effort task to generate earnings (Noninherited Inequality), or whether they each profit from the work of a real-life friend (Inherited Inequality). In the second treatment dimension, we vary whether workers complete the same fixed number of tasks and the initial distribution is determined by a random draw (Luck), or whether workers choose how many tasks to complete and the initial distribution is proportional to the relative number of completed tasks (Effort). In the redistribution stage, we sample 543 impartial spectators representative of the general US population who can redistribute the $10 between pairs of workers (Noninherited Inequality conditions) or workers' friends (Inherited Inequality conditions). Based on the treatment variation in the earnings stage, we implement a 2×2 within-subjects design in the redistribution stage: spectators make redistribution decisions for each of the four types of situations. For each situation, they observe the initial distribution and workers' relative effort before they determine the final allocation. Spectators are impartial in the sense that they have no stakes in the distribution themselves. Because redistribution is costless, we interpret the final allocation as the allocation they consider fair.

Besides the absence of spectator self-interest, this experimental setting has a number of additional advantages. First, it allows to abstract from other factors that affect distributional preferences and support for redistributive policies, such as efficiency considerations or trust in the government (Almås, Cappelen, and Tungodden, 2020; Stantcheva, 2021). Second, the comparability of redistribution decisions across experimental conditions enables us to isolate how variations in our two dimensions of interest—whether the initial distribution is tied to workers' relative efforts or based on a random draw, and whether beneficiaries are responsible for their outcomes themselves or not—affect which distribution spectators find fair. Finally, while the intergenerational transmission of wealth will be our leading example, the phenomenon that individuals derive advantages from the achievements of others is more widespread. Besides inheriting from family members, people might also profit differentially from friendship ties, coworkers, or their countries' institutional environments. Hence, studying fairness preferences in an abstract setting may yield insights into behavior, policy preferences, and fairness views in a variety of settings that have inherited inequality in general and the dilemma of meritocracy in particular at their core.

Our empirical results are in line with our theoretical framework and yet surprising. Consistent with the existing literature, we find that in the Noninherited Inequality & Luck condition redistribution levels are substantially higher than in Noninherited Inequality & Effort (Cappelen, Falch, and Tungodden, 2020). Spectators equalize about 80% of the initial inequality on average in the Luck case

but only about 5% in the Effort case. Comparing redistribution levels between the two Luck conditions reveals that spectators redistribute in a similar way when beneficiaries profit from the random draw of their friends compared to a random draw of themselves. In the Effort domain, however, spectators indeed redistribute significantly more if inequality is inherited. While in the Noninherited Inequality & Effort condition spectators equalize 5% of the inequality in the initial distribution, this share increases to 8% in Inherited Inequality & Effort.

The key takeaway though is that spectators redistribute a small fraction of the initial inequality in Inherited Inequality & Effort, close to the Noninherited Inequality & Effort benchmark but far away from the Luck benchmark of 80%. In other words, most spectators handle the dilemma of meritocracy by prioritizing fairness toward the benefactors over fairness toward the beneficiaries. This result seems to be a general feature of the US population, as it does not vary much by demographic variables like age, gender, or political ideology. Hence, there appears to be a broad consensus among US citizens that inherited inequality is acceptable as long as it is merited by those who bequest.

We examine potential reasons why spectators tend to handle the dilemma of meritocracy in favor of the benefactors by analyzing open-ended responses in which spectators explain their redistribution decisions. Consistent with their decisions, most spectators state to redistribute based on the workers' (and not their nonworking friends') relative efforts in the Inherited Inequality & Effort condition. Zooming in on spectators who acknowledge the dilemma, i.e. that they infringe meritocratic fairness irrespective of how they redistribute, reveals a more instructive consideration behind redistribution decisions: many of these spectators argue that neither of the two non-working friends is entitled to any payoff anyways, such that fairness toward the workers receives a much larger weight in their decision process. Under the assumption that workers prefer their own friends to receive the earnings they have merited through their efforts, this relative weighting of conflicting fairness judgments calls for the low level of redistribution that we observe in the experiment.

These considerations suggest that spectators observe workers' relative efforts, derive their relative entitlements, and then implement redistribution decisions trying to take into account (in particular the more industrious worker's) preferences over the distribution of payoffs between passive friends. To substantiate that this is a common rationale behind spectator's decisions, we explore how decisions are associated with spectators' (incentivized) beliefs about workers' preferred distributions of the $10 between their own and the other worker's friend. Indeed, spectators who believe that workers prefer distributions that more strongly favor their own friends redistribute less. Despite being neither causal nor conclusive, these observations suggest that spectators prioritize meritocratic fairness toward workers and try to respect workers' distributional preferences.

Due to the within-subjects design employed in the spectator stage, we can relate a given spectator's decisions across the four treatment conditions. Both within the NONINHERITED INEQUALITY and the INHERITED INEQUALITY domain, we use this feature to classify spectators into one of three fairness types that have received the most attention in the literature, and a residual type: egalitarians who prioritize equality and always redistribute, libertarians who prioritize property rights and personal freedom and never redistribute, and meritocrats who prefer distributions that reflect relative efforts. In the NONINHERITED INEQUALITY domain, we can classify all but one spectator into one of the three fairness types. By far the most prevalent fairness type is the meritocratic one (76%), followed by libertarians (21%) and only few egalitarians (3%). Most spectators display similar redistribution patterns in situations with NONINHERITED INEQUALITY and INHERITED INEQUALITY. While we observe some switching between meritocrats and libertarians that is not in line with our theoretical framework, more than 85% of the spectators behave in a way that is consistent. We conclude that our theoretical framework can accommodate spectators' redistribution behavior well.

We also relate our experimental measures of fairness preferences to attitudes toward various redistribution-related policies including income and estate taxation, disability and unemployment insurance, and support for equal opportunity programs. Because redistribution decisions across NONINHERITED INEQUALITY and INHERITED INEQUALITY situations are highly correlated both within the LUCK and the EFFORT domain, we apply a factor analysis to reduce the four behavioral measures elicited in the experiment to two factor variables. One of these factor variables captures variation in redistribution behavior in the LUCK domain while the other one captures variation in redistribution behavior in the EFFORT domain. We find that more redistribution in the experiment is related to more support for redistribution regarding all policies. This suggests that the fairness preferences identified in this experiment are a fundamental preference underlying attitudes towards various policies.

Finally, researchers who seek to relate survey responses to individual fairness preferences may often not have the resources to accommodate a thorough experimental elicitation of these preferences. We validate that unincentivized survey questions included in the post-experimental questionnaire correlate strongly with the experimentally elicited preferences in NONINHERITED INEQUALITY situations. Hence, these survey items may constitute an economical alternative in the presence of organizational constraints.

This paper contributes to a growing literature that explores how contextual and personal factors determine individuals' fairness views and redistributional preferences (Cappelen, Falch, and Tungodden, 2020). With regard to personal factors, it has been studied how redistributional preferences are associated with risk preferences (Gärtner, Mollerstrom, and Seim, 2017), depend on experienced inequality (Roth and Wohlfart, 2018), and respond to information on intergenerational

mobility (Alesina, Stantcheva, and Teso, 2018) or inequality and the tax system (Kuziemko et al., 2015). In terms of contextual factors, it is well documented that many people reject inequality that is based on luck but accept inequality if stakeholders are responsible for their outcomes, for example due to investment decisions (Cappelen, Hole, et al., 2007), effort provision (Cappelen, Sørensen, and Tungodden, 2010; Cappelen and Tungodden, 2017; Andre, 2022; Cappelen, Mollerstrom, et al., 2022; Schaube and Strang, 2022), or risk-taking (Cappelen, Konow, et al., 2013; Mollerstrom, Reme, and Sørensen, 2015). Relative to this literature, our study differs in two key aspects: first, we are primarily interested in situations where individuals are not responsible for their outcomes themselves but profit—potentially to a differential extent—from the actions of others. Second, the situations studied in existing papers usually yield interesting decision problems because individuals face uncertainty regarding decision-relevant aspects of the situation, such as to what extent the initial distribution is based on factors within versus outside individuals' control. In contrast, in our case individuals who endorse a meritocratic fairness ideal face a non-trivial decision problem even if they are perfectly informed about all relevant aspects of the situation; the dilemma originates from the fact that they will infringe meritocratic fairness no matter how they redistribute.

Our results may also help to explain why many people oppose redistributive policies. Several studies show that people's preferences regarding redistributive policies are strongly related to whether they find inequality fair or unfair (Alesina and Angeletos, 2005; Alesina and Giuliano, 2011; Stantcheva, 2021). At the same time, economic inequality is often inherited either directly through bequests or indirectly through differential education, social environments, and parenting (Bowles and Gintis, 2002; Björklund, Roine, and Waldenström, 2012; Chetty, Hendren, and Katz, 2016; Kosse et al., 2020). Hence, our finding that individuals tend to consider inequality as fair if it is based on effort at some stage suggests that people may reject redistributive policies based on fundamental fairness preferences. Faced with two similarly unattractive options, many people might perceive inherited inequality or unequal opportunity as the lesser evil and prioritize rewarding the efforts of those who pass on resources.

While Bowles and Gintis (2002) and Stantcheva (2021) briefly discuss the dilemma of meritocracy and Benabou (2000) and Piketty and Saez (2013) study related issues theoretically, Cohen, Maltz, and Ofek-Shanny (2022) is most closely related to our paper. They employ the impartial spectator design to experimentally study fairness preferences in a setting where inequality between two non-working individuals originates from the decision of a worker who has to pass on all earned money to one of these two individuals. Contrary to our results, they find that impartial spectators redistribute between the non-working subjects in a similar way as between two workers who are randomly assigned unequal initial endowments. A key difference to our design, where workers generate payments for real-life friends, is that in Cohen, Maltz, and Ofek-Shanny (2022) the worker can differentiate be-

tween the two individuals only based on their favorite hobbies, which they had to list beforehand. Because the non-working subjects are otherwise strangers to the worker, spectators may wonder whether the worker would not actually prefer an egalitarian split. Notably, the design of Cohen, Maltz, and Ofek-Shanny (2022) requires workers to pass on all of the money to one individual, precluding an equal split. If spectators indeed try to respect workers' preferences — as our analysis suggests — one would then expect redistribution toward an egalitarian split, which is common in the luck case. Hence, the results in Cohen, Maltz, and Ofek-Shanny (2022) can be well reconciled with ours.

The remainder of the paper is structured as follows: Section 3.2 introduces the theoretical framework to study fairness preferences under inherited inequality in general and the Dilemma of Meritocracy in particular. Section 3.3 details the experimental design, Section 3.4 outlines the empirical strategy, and Section 3.5 reports the results. Finally, Section 3.6 concludes.

3.2 Theoretical Framework

We are primarily interested in situations where individuals are not responsible for their outcomes themselves but profit—potentially to a differential extent—from the efforts of others. In such situations, fairness judgments may not only need to take into account whether inequality reflects differential luck or differential efforts but also balance fairness toward individuals who generated payments and toward individuals who receive these payments. To accommodate these situations, we extend the framework in Cappelen, Konow, et al. (2013) and Almås, Cappelen, and Tungodden (2020) to allow for cases of inherited inequality, in which the person responsible for an outcome is not identical to the person who receives that outcome. We derive behavioral hypotheses in Section 3.4.3, after introducing the experimental design.

3.2.1 Setup

We study distributional preferences in a situation in which a fixed sum of money P is distributed between two individuals ("beneficiaries" B_X and B_Y), who each benefit from the effort of an associated worker (W_X and W_Y). Workers exert effort for their respective beneficiaries because they are interested in their well-being; for example, one may think of workers as parents caring for their respective child. Let $e_{W_i} \geq 0$ denote the effort of worker $i \in \{X, Y\}$ and $e_{B_X} = e_{B_Y} = 0$ the effort of the two beneficiaries, who are entirely passive. After workers have exerted effort, an initial distribution of P between the two beneficiaries is realized, which may depend on effort levels and a random process. This distribution is described by $(s_0, 1 - s_0)$, with s_0 being the initial (relative) share of B_X. Without loss of generality, we assume that B_X is the initially weakly disadvantaged beneficiary, i.e., $s_0 \leq 0.5$.

Consider an impartial spectator who observes this situation and contemplates whether the distribution is fair or should be altered. The spectator is impartial in the sense that he does not receive a material benefit but incurs disutility if he perceives the distribution between the two beneficiaries to be unfair. We assume that the spectator's utility function is given by

$$V(s|\sigma) = -\frac{\alpha}{2} (\underbrace{s - s_W^f(\sigma)}_{\substack{\text{deviation from} \\ \text{what is fair} \\ \text{toward workers}}})^2 - \frac{1-\alpha}{2} (\underbrace{s - s_B^f(\sigma)}_{\substack{\text{deviation from} \\ \text{what is fair} \\ \text{toward beneficiaries}}})^2. \qquad (3.1)$$

In that expression, σ encodes information about the situation. The spectator's fairness judgments in situation σ are expressed by the relative shares $s_W^f(\sigma)$ and $s_B^f(\sigma)$, which describe the distributions $(s_L^f(\sigma), 1 - s_L^f(\sigma))$, $L \in \{W, B\}$, that the spectator considers fair toward the workers and beneficiaries, respectively. Quadratic loss functions capture the disutility from distributions that deviate from what is considered fair, and $\alpha \in [0, 1]$ governs how the spectator balances fairness toward workers and beneficiaries. Solving the corresponding maximization problem yields the distribution the spectator finds fair overall, given by

$$s^r(\sigma) = \alpha s_W^f(\sigma) + (1 - \alpha) s_B^f(\sigma). \qquad (3.2)$$

Under the given functional form assumptions, the spectator's preferred distribution is a linear combination of the distribution considered fair toward the workers and the distribution considered fair toward the beneficiaries, with weights α and $1 - \alpha$, respectively.

3.2.2 Fairness Types, Fairness Judgments, and the Dilemma of Meritocracy

Let us turn to the question of how spectators make fairness judgments. We follow the literature by assuming that spectators endorse either an egalitarian (E), libertarian (L), or meritocratic (M) fairness type τ.

Egalitarians ($\tau = E$):. An egalitarian is convinced that total resources should be distributed equally in any case. Hence, the distribution perceived fair toward workers as well as beneficiaries is given by $s_W^f(\sigma) = s_B^f(\sigma) = s^f(\sigma) = \frac{1}{2}$. Because perceived fair shares coincide, egalitarians do not encounter a conflict in the case of inherited inequality, and the preferred distribution is $s^r(\sigma) = \frac{1}{2}$.

Libertarians ($\tau = L$):. A libertarian does not value equality but advocates the opposing standpoint that one should not intervene in the allocation process and therefore accepts the initial allocation. The perceived fair distributions are given by $s_W^f(\sigma) = s_B^f(\sigma) = s^f(\sigma) = s_0$ and the overall preferred distribution is $s^r(\sigma) = s_0$.

Meritocrats ($\tau = M$):. In between, meritocrats think that distributions should reflect individual merits: $s_L^f(\sigma) = \frac{e_{L_X}}{e_{L_X} + e_{L_Y}}$ if $e_{L_X} + e_{L_Y} > 0$ and $s_L^f(\sigma) = \frac{1}{2}$ if $e_{L_X} + e_{L_Y} = 0$, with $L \in \{W, B\}$. Hence, in the case of inherited inequality, meritocrats may face a dilemma: because beneficiaries do not exert any effort but their associated workers may exert different levels of effort ($e_{W_X} \neq e_{W_Y}$), it follows that $s_B^f = \frac{1}{2}$ but usually $s_W^f = e_{W_X}/(e_{W_X} + e_{W_Y}) \neq \frac{1}{2}$ — merit judgments conflict! As a consequence, meritocrats need to balance fairness toward workers and beneficiaries, and the overall perceived fair share is given by

$$s^r(\sigma) \;=\; \alpha \, \frac{e_{W_X}}{e_{W_X} + e_{W_Y}} + (1 - \alpha) \, \frac{1}{2}. \tag{3.3}$$

We denominate this phenomenon the *Dilemma of Meritocracy*. If one worker chose to exert higher effort for the sake of his beneficiary than the other, this pulls the meritocrat toward a distribution between beneficiaries that reflects these differences in effort. Conversely, both beneficiaries are passive and none merited more resources than the other, which pulls the meritocrat toward an egalitarian distribution. The weighting parameter α that governs how this dilemma is handled may be interpreted as the relative importance of the workers' and the beneficiaries' perspectives in the meritocrat's overall fairness judgment.

3.2.3 Noninherited Inequality

Our framework nests the case of noninherited inequality studied in existing research, where each worker is identical to his associated beneficiary, $W_i \equiv B_i$. This implies that $e_{W_i} = e_{B_i}$ and fairness judgments toward workers and beneficiaries coincide for all fairness types: $s_W^f = s_B^f = s^f$. The spectator's utility function collapses to $V(s|\sigma) = -(s - s^f(\sigma))^2$, and the solution is simply $s^r(\sigma) = s^f(\sigma)$, such that one reobtains the formulation used in Cappelen, Konow, et al. (2013) and Almås, Cappelen, and Tungodden (2020).

3.3 Experimental Design

Our experiment builds on the impartial spectator paradigm (Konow, 2000; Cappelen, Konow, et al., 2013) and consists of two stages. In the earnings stage, an initial (pre-redistribution) allocation of \$10 between two stakeholders is determined. In the redistribution stage, impartial spectators may redistribute the \$10 between the two stakeholders to determine the final (post-redistribution) allocation. We are primarily interested in spectators' redistribution decisions; the earnings stage is used to incentivize these decisions.

3.3.1 The Earnings Stage

In the earnings stage, we implement four treatment conditions in a between-subjects design. In all conditions, subjects work on a real-effort task in which they have to reposition sliders into the middle position (Gill and Prowse, 2012). Each task has a fixed duration of 30 seconds and requires repositioning five sliders, which is easy to achieve. Hence, completing tasks is solely a matter of effort and time, but not ability. After workers have completed their participation, they are divided into pairs of two. Treatments differ in two dimensions. One dimension varies whether the initial distribution of the $10 is determined by a random draw ("Luck") or reflects the relative number of completed tasks ("Effort"). The other dimension varies whether the $10 is distributed between a pair of workers themselves ("Noninherited Inequality") or whether each worker designates a real-life friend and the $10 is distributed between the two friends of a pair of workers ("Inherited Inequality"). Working with real-life friends has organizational advantages over, for example, the stricter requirement that workers designate a beneficiary among their family members. At the same time, friendship ties capture two central aspects of relationships between benefactors and beneficiaries that may be prerequisites for the dilemma of meritocracy: there is a meaningful relationship between workers and their friends, and workers are more altruistic toward their own friend than toward the friend of the other worker (Gächter, Starmer, and Tufano, 2015).

The 2x2 variation in the earnings stage results in the following four conditions which are summarized in Table 3.1:

- **Noninherited Inequality & Luck:** Workers complete exactly 20 tasks. $10 are distributed between the two workers of a pair. The initial distribution is determined by a random draw. Each distribution is equally likely.

- **Noninherited Inequality & Effort:** Workers choose to complete between 0 and 40 tasks. $10 are distributed between the two workers of a pair. The initial distribution corresponds to the relative number of completed tasks.

- **Inherited Inequality & Luck:** Workers complete exactly 20 tasks. Each worker chooses a real-life friend, and $10 is distributed between the workers' friends. The initial distribution is determined by a random draw. Each distribution is equally likely.

- **Inherited Inequality & Effort:** Workers choose to complete between 0 and 40 tasks. Each worker chooses a real-life friend, and $10 is distributed between the workers' friends. The initial distribution corresponds to the relative number of completed tasks.

Before they start working, workers know whether they generate earnings for themselves or a real-life friend and how the initial allocation is determined. They also know that another person's decision may affect their (or their friend's) payoff,

Table 3.1. Features of Treatment Arms

Treatment	$10 distr. betw.	# Tasks completed	Initial allocation
NONINHERITED INEQ. & LUCK	Workers	$e_x = e_y = 20$	$s_0 \sim U[0, 1]$
NONINHERITED INEQ. & EFFORT	Workers	$e_x, e_y \in [0, 40]$	$s_0 = e_x/(e_x + e_y)$
INHERITED INEQ. & LUCK	Workers' friends	$e_x = e_y = 20$	$s_0 \sim U[0, 1]$
INHERITED INEQ. & EFFORT	Workers' friends	$e_x, e_y \in [0, 40]$	$s_0 = e_x/(e_x + e_y)$

Notes: e_x and e_y denote the number of tasks by worker X and Y, respectively. $U[\cdot]$ denotes the uniform distribution and s_0 denotes the share of the $10 allocated to stakeholder X according to the initial distribution. The share of the $10 allocated to stakeholder Y according to the initial distribution always equals $1 - s_0$.

but not how and why. Workers (and their friends) never observe the initial allocation or spectators' decisions. Friends are entirely passive.

Workers make a final decision at the end of the earnings stage. We ask workers in the NONINHERITED INEQUALITY conditions how they would distribute additional $10 between themselves and the worker they are matched to if they could freely decide. Likewise, we ask workers in the INHERITED INEQUALITY conditions how they would distribute $10 between their own friend and the friend of the worker they are matched to. Workers are incentivized to report their preferences truthfully, as we would randomly draw one worker and implement his or her preference. We will later refer to these decisions as dictator decisions.

3.3.2   The Redistribution Stage

In the redistribution stage, unrelated subjects ("impartial spectators") can redistribute the $10 between pairs of workers or workers' friends. Based on the four conditions from the earnings stage, we implement a 2x2 within-subjects design in the redistribution stage. Before they make a redistribution decision, spectators learn whether $10 is distributed between workers or passive friends, whether the initial allocation was determined by a random draw or according to the relative number of completed tasks, and the initial allocation. They make their decision by entering the final distribution in the form of relative shares of the two workers (in the NONINHERITED INEQUALITY conditions) or friends (in the INHERITED INEQUALITY conditions) in a table that also contains condensed information about the situation. Figure 3.C.1 shows a screenshot of the decision screen in the INHERITED INEQUALITY & EFFORT condition; the other decision screens had the same structure. To focus on the fairness aspect of the redistribution problem, we abstract from a potential fairness-efficiency tradeoff (Almås, Cappelen, and Tungodden, 2020) by making redistribution costless.

Similar to recent studies that use the impartial spectator design (Schaube and Strang, 2022) we employ a variant of the strategy method (Kube and Traxler, 2011). For each spectator, we construct a set of six initial allocations that consists

of one initial allocation from a randomly drawn situation that has occurred in the earnings stage and five hypothetical initial allocations that are constant across all spectators. The hypothetical initial allocations were ($0.00, $10.00), ($1.00, $9.00), ($2.20, $7.80), ($3.00, $7.00), and ($3.80, $6.20).[2] These initial allocations yield a block of six situations within each of the four conditions – 24 situations in total – for which we ask spectators to make redistribution decisions.

Spectators make redistribution decisions for all situations within a block before they proceed to the next one. After each block, they are prompted to briefly describe the reasoning behind their decisions. We randomize the order of blocks as well as the order of situations within each block between subjects. Spectators know that some situations are hypothetical and that we randomly select one spectator for each pair of workers (friends), whose decision for the relevant situation is implemented. Because spectators do not know whether a decision is potentially relevant or not, all decisions are probabilistically incentivized.

After the spectators completed the redistribution part, we ask them qualitatively to what extent they find luck-based and effort-based inequality between two individuals fair. Because it may be too expensive or time-consuming to elicit incentivized experimental measures of fairness preferences in some surveys, it is useful to know whether such short nonincentivized survey measures can be employed as substitutes. Section 3.A provides evidence that they can.

Thereafter, we elicit their beliefs about workers' dictator decisions. Separately for workers in the NONINHERITED INEQUALITY and INHERITED INEQUALITY conditions, we ask spectators to guess how much workers on average kept for themselves or gave to their own friends, respectively. Spectators receive a bonus of $0.20 for each guess with less than $0.20 distance to the actual value, such that guesses are incentivized as well. Finally, spectators complete a brief questionnaire on their general attitudes toward inequality, their assessment of various policies related to inequality and redistribution, and additional demographics.

3.3.3 Procedures

3.3.3.1 Workers and Friends

The earnings stage was conducted online in March 2022 and implemented using oTree (Chen, Schonger, and Wickens, 2016). Workers were recruited from the BonnEconLab subject pool via Hroot (Bock, Baetge, and Nicklisch, 2014). The invitation mail informed potential participants that some of them would be able to generate a payment for a real-life friend. In the confirmation email, workers in the INHERITED INEQUALITY conditions received a link that they had to pass on to a friend. Via that

---

2. If the initial allocation in the randomly drawn situation was identical to one of the hypothetical initial allocations, the respective hypothetical initial allocation was replaced by a "backup" allocation. This case applied for 52 spectators.

link, friends had to give us their bank details. On the next day, the corresponding workers received another email with a participation link only if a friend had given us his or her bank details before, such that we could ensure to be able to make all payments that were generated in the study. Workers in the Noninherited Inequality conditions were informed in the confirmation email that they were not among those participants that could generate a payment for a friend and received an email with a participation link on the next day as well. All workers could start immediately when they received the participation link and had time to conclude their participation until the end of the day.

In the earnings stage itself, workers had to enter their own bank details before they received condition-specific instructions and entered the work stage. Workers in the Effort conditions could choose how many tasks to complete, whereas workers in the Luck conditions had to complete exactly 20 tasks.[3] After the work stage, workers had to make their respective dictator decision to conclude their participation.

In total, 43 workers completed their participation in the earnings stage, 21 in the Noninherited Inequality conditions and 22 in the Inherited Inequality conditions. In the Nonhereditary Inequality conditions, each worker received a fixed payment of $3, and $10 was distributed between two workers each. In the Inherited Inequality conditions, each worker received a fixed payment of $5, each friend received a fixed payment of $3, and $10 was distributed between two friends each. In addition, one among all workers' dictator decisions was randomly selected and implemented as announced during the study. Payoffs were presented in the form of experimental currency during the earnings stage but eventually made in euros via bank transfer.

3.3.3.2 Spectators

The redistribution stage was conducted online in late April 2022 and implemented using oTree as well. We recruited a sample of 552 adult US citizens via the survey provider Prolific, which has been shown to provide higher data quality than comparable companies (Palan and Schitter, 2018; Peer et al., 2021). In addition to incentivizing redistribution decisions, we took several measures to further promote quality responses, including two attention checks, control questions for each block of redistribution decisions, and graphical instructions that are arguably more engaging than large blocks of text instructions. Details and data quality checks are presented in Section 3.B, which also provides evidence that spectators recognized and understood the differences between treatments.

---

3. Workers could at most attempt 60 tasks until the work stage was automatically concluded. One worker in the Luck conditions did not manage to complete 20 tasks with 60 attempts and did not generate a payment, as was announced beforehand.

Spectators were recruited in two waves within the same week.[4] The first and second wave contained 75 and 477 spectators, respectively. Because participants from the first wave were not excluded from participating in the second wave, nine spectators participated twice. We only include the first observation from these participants, such that we end up with a sample of 543 spectators. The median completion time in the first wave was 21 minutes and subjects earned a base rate of £3.03 plus bonus payments. The median completion time in the second wave was slightly longer at 25 minutes and participants earned a base rate of £2.55 plus bonus payments. For the second wave, Prolific recruited a sample representative of the US adult population aged 18 or older regarding the joint distribution of age, sex, and ethnicity. This was impossible for the first wave due to the low number of participants. Yet, as shown in Table 3.D.1, our total spectator sample is representative of the adult US population in terms of age, gender, and ethnicity. In contrast, our sample over-represents the well-educated and underrepresents the top quartile of the income distribution, which is common for survey samples (Stantcheva, 2023). The study was preregistered at the AER RCT Registry (RCT ID: AEARCTR-0009186). The instructions for the spectator session and the pre-analysis plan can be accessed here: https://doi.org/10.1257/rct.9186.

3.4 Empirical Analysis

3.4.1 Main Variables

Independent Variables. Our main independent variables are the indicators II_σ ($= 1$ if situation σ features inherited inequality) and E_σ ($= 1$ if the initial allocation in situation σ is based on effort). Both indicators together describe the treatment condition situation σ was embedded in. Further, we define the initial extent of inequality $\Delta_\sigma = 0.5 - s_0$, which allows us to investigate whether redistribution decisions depend on how much inequality is present in the initial allocation.

Dependent Variables. Observing that a spectator implements ($4, $6) as the final allocation indicates very different redistributional preferences if the initial allocation was ($2, $8) instead of ($4, $6). In the former case, the spectator reduces inequality while in the latter inequality is left constant. To differentiate between such cases, our analysis needs to take into account that the initial allocation varies across situa-

---

4. The two-wave procedure mainly served to test for technical issues. Indeed, during the first wave, we recognized that for some of the spectators one hypothetical initial allocation was always replaced by the backup allocation due to a bug, which we fixed immediately. Because there is nothing inherently special about our preselected hypothetical initial allocations this is not a big issue, though, and the respective decisions/observations are treated like all other decisions and as described in Section 3.4.2.

tions.[5] Hence, as pre-registered, we define as our main outcome variable the extent of redistribution implemented by spectator i in situation σ,

$$\theta_{i,\sigma} = \frac{s_i^r - s_0}{0.5 - s_0}. \tag{3.4}$$

The extent of redistribution describes the fraction of inequality in the initial situation that is equalized by spectator i's redistribution decision. $\theta_{i,\sigma} = 1$ indicates that spectator i completely equalizes payoffs in situation σ while $\theta_{i,\sigma} = 0$ means that i accepts the initial allocation. For some analyses we use the average of spectator i's redistribution decisions within a given condition, which we refer to as the average extent of redistribution, $\bar{\theta}_{i,c}$, $c \in \{\text{NI-L, NI-E, II-L, II-E}\}$.

3.4.2 Exclusion Criteria and Restricted Sample

To ensure high data quality, we remove some observations from our main sample as preregistered. First, we drop spectators who fail both attention checks. Second, if a spectator rushes unreasonably fast through the instructions for a given block of redistribution decisions, we drop the decisions of that spectator for the corresponding condition. Third, we only include observations for situations that all spectators encountered because these are constant across spectators and admit a clean comparison. Hence, the main sample does not include observations based on a true scenario (except if that scenario coincides with a hypothetical one) or the backup scenario.

Based on the main sample, we further construct a restricted sample that disregards observations that cannot be reconciled with the fairness ideals prevalent in the literature, which was preregistered as well. First, we drop observations which imply $\theta_{i,\sigma} < 0$ (the spectator redistributes money from the already disadvantaged beneficiary to the already advantaged beneficiary) or $\theta_{i,\sigma} > 1$ (the spectator redistributes more to the initially disadvantaged beneficiary than what would lead to a 50/50 split). While such decisions should not prematurely be characterized as "noise" or "irrational", we cannot explain these decisions within our framework and our hypotheses do not pertain to such behavior. Second, we completely drop a spectator from the restricted sample if we disregard three or more decisions of that spectator within any of the four conditions, either because the spectator rushed or because too many decisions imply $\theta_{i,\sigma} \notin [0,1]$.

---

5. This is different from existing studies on fairness preferences in the context of noninherited inequality, where usually one of the two workers receives all of the money in the initial distribution (see e.g. (Cappelen and Tungodden, 2017; Almås, Cappelen, and Tungodden, 2020; Cappelen, Mollerstrom, et al., 2022; Schaube and Strang, 2022)). In that case, it suffices to normalize that the first worker is the initially disadvantaged one (or vice versa) and consider how much that worker receives after redistribution.

Table 3.2. Predicted Extent of Inequality (θ, Share) by Condition and Fairness Type

Condition	Egalitarians	Libertarians	Meritocrats
NONINHERITED INEQ. & LUCK	1	0	1
NONINHERITED INEQ. & EFFORT	1	0	0
INHERITED INEQ. & LUCK	1	0	1
INHERITED INEQ. & EFFORT	1	0	$1 - \alpha$

Starting with 543 spectators and 13,032 decision observations, we end up with 543 spectators and 10,236 decision observations in the main sample and 437 spectators and 8,399 observations in the restricted sample. Unless indicated differently, the results presented in the paper are based on the restricted sample. However, results do not differ notably if we consider the main sample or all of the 13,032 observations for which our main outcome measure is defined, that is, where the initial allocation is not 50/50.

3.4.3 Behavioral Predictions & Preregistered Hypotheses

The theoretical framework outlined in Section 3.2 makes nuanced individual-level predictions about what kinds of behavioral patterns we should observe across the four treatment conditions, given a subjects' fairness type: egalitarians always prefer equal distributions, libertarians always go with the initial distribution, and meritocrats prefer distributions that reflect relative effort. Given that $e_{W_X}/(e_{W_X} + e_{W_Y})$ equals 1/2 in the LUCK conditions and s_0 in the EFFORT conditions, the expression for the perceived fair share (Equation 3.2) collapses to numbers for each of the three fairness types. Plugging these numbers into the definition of the extent of redistribution (Equation 3.4) yields predictions on the extent of redistribution spectators with different fairness types implement in the different conditions. These predictions are summarized in Table 3.2.

Assuming that all types are present in our sample, these predictions imply that the four conditions should be ordered in terms of the average extent of redistribution as follows: $\bar{\theta}_{NI-L} = \bar{\theta}_{II-L} \geq \bar{\theta}_{II-E} \geq \bar{\theta}_{NI-E}$, with at least one of the inequalities being strict. Based on the individual-level predictions and this expected ordering, we derive the following four (preregistered) aggregate-level predictions that we will formally test using ordinary least squares (OLS) regressions and clustering standard errors on the spectator-level:

Hypothesis 3.1. *Spectators redistribute less if inequality is based on effort instead of luck.*

Because this hypothesis should hold both in the noninherited inequality domain (H1a) and — weakly — in the inherited inequality domain (H1b), we will test it separately within both domains. Formally, we estimate the following (regression) equation:

$$\theta_{i,\sigma} \ = \ \beta + \beta_E \cdot E_\sigma + \delta \cdot \Delta_\sigma + \varepsilon_{i,\sigma}. \tag{3.5}$$

We preregistered to test $H_0 : \beta_E = 0$ against $H_1 : \beta_E \neq 0$ and interpret $\beta_E < 0$ and the rejection of H_0 as evidence in favour of Hypothesis 1.

Hypothesis 3.2. *Spectators redistribute more if inequality is inherited.*

Pooling the data from the Luck and Effort conditions, we estimate

$$\theta_{i,\sigma} \ = \ \beta + \beta_{II} \cdot II_\sigma + \delta \cdot \Delta_\sigma + \varepsilon_{i,\sigma}, \tag{3.6}$$

and test $H_0 : \beta_{II} = 0$ against $H_1 : \beta_{II} \neq 0$ as preregistered, interpreting $\beta_{II} > 0$ and the rejection of H_0 as evidence in favour of Hypothesis 2.

Hypothesis 3.3. *The higher extent of redistribution in the case of inherited inequality is driven by situations in which inequality is based on effort.*

To formally test whether the fact that inequality is inherited indeed only matters if the initial allocation is based on effort, we consider the following difference-in-difference-like regression equation:

$$\theta_{i,\sigma} \ = \ \beta + \beta_E \cdot E_\sigma + \beta_{II} \cdot II_\sigma + \beta_{E,II} \cdot E_\sigma \cdot II_\sigma + \delta \cdot \Delta_\sigma + \varepsilon_{i,\sigma}. \tag{3.7}$$

In accordance with our pre-analysis plan, we test $H_0^a : \beta_{II} = 0$ against $H_1^a : \beta_{II} \neq 0$ and $H_0^b : \beta_{E,II} = 0$ against $H_1^b : \beta_{E,II} \neq 0$. We interpret the results as evidence in favour of Hypothesis 3 if we find $\beta_{E,II} > 0$ and reject H_0^b but not H_0^a.

Hypothesis 3.4. *The higher extent of redistribution in the case of inherited inequality, driven by situations in which inequality is based on effort, is driven by meritocrats.*

Due to the within-subjects design, we can relate individual redistribution patterns across conditions. We will classify spectators into the three fairness types (and a residual type) based on their decisions in the Noninherited Inequality conditions (details follow later) and estimate

$$
\begin{aligned}
\theta_{i,\sigma} \ = \quad & \beta^E & + \ \beta^L L_i & \quad + \ \beta^M M_i & \quad + \ \beta^{NC} NC_i \\
& + \beta_E^E E_\sigma & + \ \beta_E^L E_\sigma L_i & \quad + \ \beta_E^M E_\sigma M_i & \quad + \ \beta_E^{NC} E_\sigma NC_i \\
& + \beta_{II}^E II_\sigma & + \ \beta_{II}^L II_\sigma L_i & \quad + \ \beta_{II}^M II_\sigma M_i & \quad + \ \beta_{II}^{NC} II_\sigma NC_i \\
& \\
& + \beta_{E,II}^E E_\sigma II_\sigma & + \ \beta_{E,II}^L E_\sigma II_\sigma L_i & + \ \beta_{E,II}^M E_\sigma II_\sigma M_i & + \ \beta_{E,II}^{NC} E_\sigma II_\sigma NC_i \\
& & & & + \ \delta \Delta_\sigma + \varepsilon_{i,\sigma}.
\end{aligned}
\tag{3.8}
$$

Here, egalitarians are the baseline type and L_i (libertarian), M_i (meritocrat), and NC_i (non-classified) are indicators that equal one if spectator i is classified into the corresponding fairness type. As preregistered, we test $H_0^a : \beta_{E,II}^M = 0$ against $H_1^a : \beta_{E,II}^M \neq 0$ and $H_0^b : \beta_{E,II}^M = \beta_{E,II}^L$ against $H_1^b : \beta_{E,II}^M \neq \beta_{E,II}^L$ and interpret the results as evidence in favour of the hypothesis if $\beta_{E,II}^M > 0$, $\beta_{E,II}^M > \beta_{E,II}^L$, and we reject both H_0^a and H_0^b.

3.5 Results

First, we compare the average extent of redistribution between treatment conditions, displayed in Figure 3.1. Averages are taken over all decisions of all subjects in the restricted sample. Comparing redistribution levels between NONINHERITED
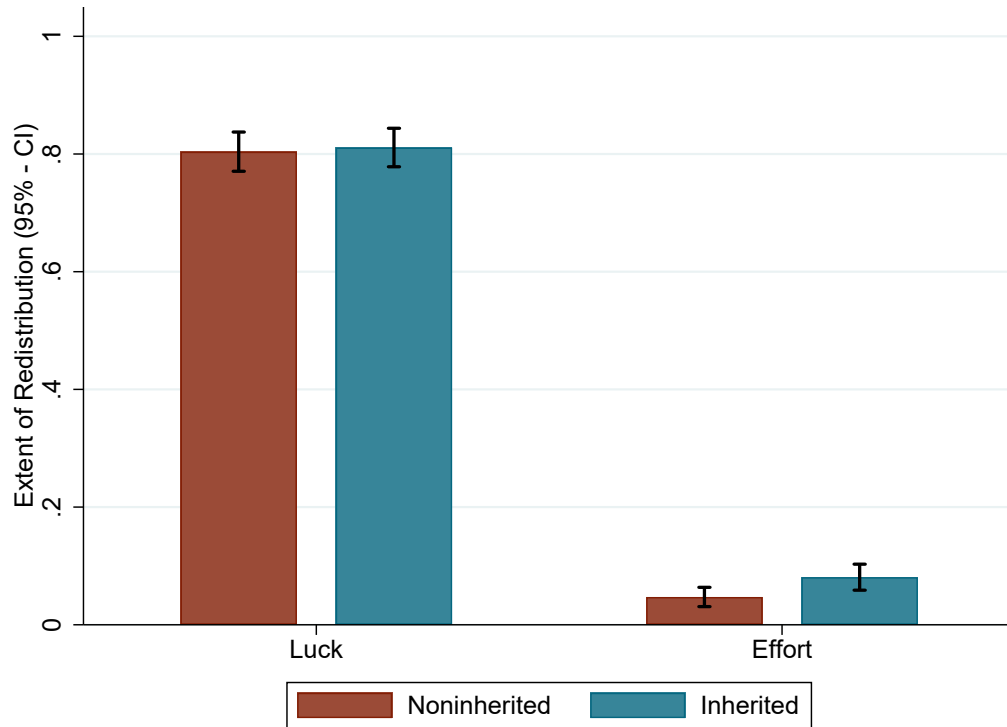
Figure 3.1. Average Extent of Redistribution $\bar{\theta}_{i,c}$ by Treatment Condition

Note: This figure displays the average extent of redistribution $\bar{\theta}_{i,c}$ by treatment condition, together with $95 - \%$ confidence intervals. Averages are taken over all decisions of all subjects in the restricted sample. Confidence intervals are based on standard errors clustered on the spectator level.

INEQUALITY & LUCK and NONINHERITED INEQUALITY & EFFORT, we replicate what many studies have documented before: under noninherited inequality, where workers' actions determine their own earnings and spectators do not need to balance potentially conflicting fairness ideals, they redistribute much less if distributions reflect differential effort than if they are based on a random draw. While they, on average, equalize about 80% of the inequality in the initial distribution in the LUCK case, they equalize only about 5% in the EFFORT case. These numbers suggest that many spectators in our sample subscribe to the meritocratic idea that resource distributions should reflect individual effort and achievement.

Consistent with our theoretical considerations from Section 3.2, a comparison of redistribution levels between NONINHERITED INEQUALITY & LUCK and INHERITED INEQUALITY & LUCK shows that it makes no difference whether inequality is inherited or not in the LUCK domain: the difference is insignificant and small both

in absolute and relative terms.[6] This indicates that in the Luck domain, given that in either case the initial distribution is not tied to relative effort, it does not matter whether the money goes to the workers themselves or is inherited by their passive friends.

To judge how spectators deal with the dilemma of meritocracy, we examine how the average extent of redistribution in Inherited Inequality & Effort compares to the Noninherited Inequality & Luck and Noninherited Inequality & Effort benchmarks. As displayed in Figure 3.1, the fraction of inequality that is equalized in Inherited Inequality & Effort (8%) is significantly higher than the share that is equalized in Noninherited Inequality & Effort (5%).[7] However, the key takeaway is that the average extent of redistribution in Inherited Inequality & Effort is much closer to the Noninherited Inequality & Effort benchmark than to the Noninherited Inequality & Luck benchmark (80%). This is consistent with our theoretical considerations from Section 3.2, but given that any magnitude between the two benchmarks would have been similarly consistent, this result may almost be considered a corner solution. Speaking in model terms, the data suggest that spectators "have a high α": they prioritize fairness toward the workers—whose effort is reflected in the initial distribution—and accept that in the Inherited Inequality case the beneficiaries end up with different shares even though one did not "merit" more than the other. Overall, these results suggest that spectators treat the dilemma of meritocracy by prioritizing fairness toward the workers over fairness toward the friends.

3.5.1 The Aggregate Level: Testing the Hypotheses

To test the hypotheses from Section 3.4.3, we estimate the corresponding preregistered regression equations using OLS regressions. All reported equations control for the initial extent of inequality in a given situation (Δ_σ), and standard errors are always clustered on the spectator level. The results are reported in Table 3.3. The titles below the column numbers indicate which hypothesis is referred to.

The estimates in columns (1) and (2) indicate that, both in the case of Noninherited Inequality and Inherited Inequality, spectators redistribute significantly less if the initial distribution is based on effort rather than luck. The differences in the average extent of redistribution amount to 76%p (Noninherited Inequality) and 73%p (Inherited Inequality), respectively.

We further observe that the initial extent of inequality (Δ_σ) has a weakly significant but small effect on the fraction of inequality spectators equalize. The estimates

---

6. $d = 0.007$ and $p = 0.62$ in an OLS regression of the form $\theta_{i,\sigma} = \beta + \beta_{II} \cdot II_\sigma + \varepsilon_{i,\sigma}$, using only observations from the Luck domain and clustering standard errors on the spectator level.

7. $d = 0.034$ and $p < 0.001$ in an OLS regression of the form $\theta_{i,\sigma} = \beta + \beta_{II} \cdot II_\sigma + \varepsilon_{i,\sigma}$, using only observations from the Effort domain and clustering standard errors on the spectator level.

Table 3.3. Treatment Effects on the Extent of Redistribution

	Dependent Variable: Extent of Redistribution ($\theta_{i,c}$, Share)					
	Restricted Sample				Main Sample	Full Sample
	(1)	(2)	(3)	(4)	(5)	(6)
	H1a	H1b	H2	H3	H3	H3
Effort (E_σ)	-0.757***	-0.730***		-0.757***	-0.747***	-0.741***
	(0.019)	(0.019)		(0.019)	(0.020)	(0.020)
Inherited (II_σ)			0.022**	0.007	0.021	0.017
			(0.009)	(0.014)	(0.015)	(0.016)
Effort (E_σ) × Inherited (II_σ)				0.027	0.022	0.042**
				(0.016)	(0.019)	(0.021)
Initial Inequality (Δ_σ)	0.031*	0.035*	0.024	0.033**	0.079***	0.054
	(0.018)	(0.019)	(0.015)	(0.015)	(0.019)	(0.042)
Constant	0.795***	0.801***	0.421***	0.794***	0.784***	0.789***
	(0.018)	(0.018)	(0.011)	(0.018)	(0.019)	(0.024)
Included Treatments	NI-L & NI-E	II-L & II-E	All	All	All	All
Clusters	437	437	437	437	543	543
Observations	4203	4196	8399	8399	10236	12448
R^2	0.620	0.575	0.001	0.598	0.488	0.364

Note: This table reports results from OLS regressions of the extent of redistribution implemented by spectator i in situation σ on treatment indicators, controlling for the initial extent of inequality in situation σ. Columns (1) and (2) correspond to Equation 3.5 and estimate the difference between redistribution in the Effort versus Luck case, once in the Noninherited Inequality and once in the Inherited Inequality domain. Column (3) corresponds to Equation 3.6 and estimates the difference between redistribution if inequality is inherited versus noninherited, pooling Effort and Luck situations. Columns (4) - (6) correspond to Equation 3.7 and interact both treatment dimensions using observations from all treatment conditions. For information on the composition of the different subsamples, see Section 3.4.2. Standard errors (in parentheses) are clustered on the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

show that the extent of redistribution is $3-4\%p$ higher on average if the initial extent of inequality is one unit larger. Given that the variable is only defined over the interval from 0 (a 50/50 split) to 0.5 (one stakeholder receives everything), the effect is more tangibly described by saying that, for example, going from a 30/70 split to a 20/80 split increases the average extent of redistribution by $0.3-0.4\%p$. Consistent with this result, Figure 3.C.2 in the appendix shows that the average extent of redistribution varies little by initial allocation in either treatment condition. Overall, these observations yield strong support for Hypothesis 3.1:

Result 3.1. *In both the Noninherited Inequality and the Inherited Inequality domain, spectators redistribute considerably less on average if inequality is based on effort instead of luck.*

Moving to the regression equation in column (3), which makes use of all observations in the restricted sample, we see that spectators redistribute significantly more if inequality is inherited. Consistent with Hypothesis 3.2, the average extent of redistribution is 2.2%p higher if the money is distributed between passive friends instead of the workers themselves. Yet, in contrast to the magnitude of the difference in redistribution levels between Effort and Luck situations, the effect is almost negligible. We summarize these observations in the following result:

Result 3.2. *Spectators redistribute significantly more if inequality is inherited. However, the magnitude of the effect is small.*

The remaining columns, (4)-(6), test for an interaction effect: does the fact that payoffs are inherited matter more if the initial distribution is based on workers' relative effort levels instead of a random draw? Whereas the difference in average redistribution levels between Inherited Inequality and Noninherited Inequality situations is less than 1%p if the initial distribution is determined by luck, this difference is about five times as large $(0.007 + 0.027)$ if the initial distribution is proportional to workers' relative effort. The interaction effect is still small, however, and just short of reaching statistical significance. The numbers and qualitative patterns are very similar if the same equation is estimated on the main sample (column (5)), which includes observations that cannot be reconciled with commonly considered fairness ideals, i.e., $\theta_{i,\sigma} \notin [0, 1]$. Similarly, results change little if we consider the full sample (column (6)), which includes situations based on true scenarios and from blocks where spectators rushed through the instructions, albeit the interaction effect is statistically significant here. Relative to our main regression equation in column (4) the share of variance explained drops sharply in columns (5) and (6), which indicates that our sample restrictions successfully reduce the amount of noise in the data. Overall, we interpret these observations as (partial) support in favour of Hypothesis 3.3:

Result 3.3. *The higher extent of redistribution in the case of inherited inequality is, if anything, driven by situations in which inequality is based on effort.*

3.5.2 The Individual Level: Redistribution Patterns & Fairness Types

Our within-subjects setup in the redistribution stage has the advantage that we can relate a given spectator's redistribution decisions across the four different conditions. In this subsection, we use this feature to detect common redistribution patterns. As a first step, we use subjects' decisions in the two Noninherited Inequality conditions to classify them into one of three fairness types discussed in Section 3.2.2: egalitarians (*E*), libertarians (*L*), and meritocrats (*M*). We define a spectator's fairness type in situations of noninherited inequality, $\tau_{i,NI}$, as follows:

$$
\tau_{i,NI} = \begin{cases}
E & \text{if } \bar{\theta}_{i,NI-L} \geq 0.5 \text{ and } \bar{\theta}_{i,NI-E} \geq 0.5 \\
M & \text{if } \bar{\theta}_{i,NI-L} \geq 0.5 \text{ and } \bar{\theta}_{i,NI-E} < 0.5 \\
L & \text{if } \bar{\theta}_{i,NI-L} < 0.5 \text{ and } \bar{\theta}_{i,NI-E} < 0.5 \\
NC & \text{else,}
\end{cases}
\tag{3.9}
$$

where *NC* describes a residual type of "Nonclassifieds".

Figure 3.2 plots the distribution of spectators in the $\bar{\theta}_{i,NI-L} \times \bar{\theta}_{i,NI-E}$ space. The horizontal axis indicates the average extent of redistribution in the NONINHERITED INEQUALITY & LUCK condition. Similarly, the vertical axis measures the average extent of redistribution in NONINHERITED INEQUALITY & EFFORT. Hence, each circle in Figure 3.2 represents the redistribution behavior of a spectator in the NONINHERITED INEQUALITY domain, and circle size is proportional to the number of spectators at the corresponding position.

Figure 3.2. Classification into Fairness Types - NONINHERITED INEQUALITY
Note: Circles correspond to subjects in the spectator role of the experiment. The horizontal axis describes the share of inequality that the individual equalized on average in the NONINHERITED INEQUALITY & LUCK condition. The vertical axis describes the share of inequality that the individual equalized on average in the NONINHERITED INEQUALITY & EFFORT condition. Circle size is proportional to the number of spectators at the corresponding position. Subjects were classified according to the label names in the four quadrants, and colors indicate the respective classes.

Two aspects of the plot attract particular attention. First, the majority of spectators (76%) fall into the bottom right quarter and are, therefore, classified as meritocrats. A much smaller fraction of spectators (21%) are classified as libertarians, and only a few (3%) are classified as egalitarians. Only a single spectator in the restricted

sample remains unclassified. Second, spectators in general behave very consistently: most of them make either perfectly meritocratic (59%), libertarian (10%), or egalitarian (3%) decisions.

As a second step, in analogy to the noninherited inequality classification, we define a spectator's redistribution pattern in situations with inherited inequality, $\tau_{i,II}$:

$$
\tau_{i,II} = \begin{cases}
E & \text{if } \bar{\theta}_{i,II-L} \geq 0.5 \text{ and } \bar{\theta}_{i,II-E} \geq 0.5 \\
M & \text{if } \bar{\theta}_{i,II-L} \geq 0.5 \text{ and } \bar{\theta}_{i,II-E} < 0.5 \\
L & \text{if } \bar{\theta}_{i,II-L} < 0.5 \text{ and } \bar{\theta}_{i,II-E} < 0.5 \\
NC & \text{else.}
\end{cases}
\tag{3.10}
$$

Figure 3.3 shows, in the familiar fashion, where spectators are positioned in the $\bar{\theta}_{i,II-L} \times \bar{\theta}_{i,II-E}$ space. To relate spectators' redistribution patterns across situations
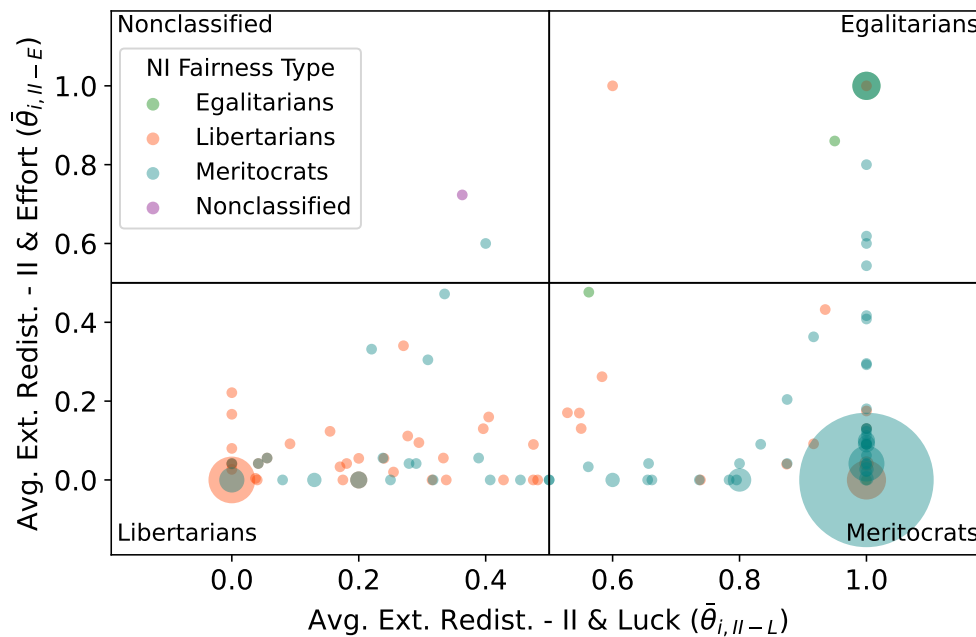
Figure 3.3. Classification by Redistribution Patterns - Inherited Inequality
Note: Circles correspond to subjects in the spectator role of the experiment. The horizontal axis describes the share of inequality that the individual equalized on average in the Inherited Inequality & Luck condition. The vertical axis describes the share of inequality that the individual equalized on average in the Inherited Inequality & Effort condition. Circle size is proportional to the number of spectators at the corresponding position. Subjec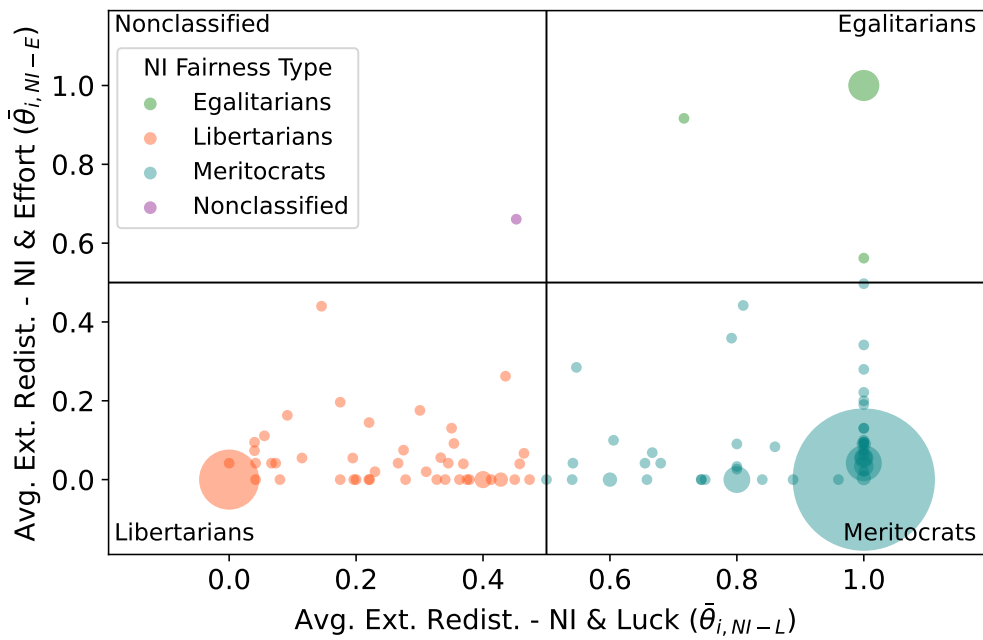ts were classified according to the labels in the four quadrants. Colors indicate how spectators were classified in the Noninherited Inequality situations.

with noninherited and inherited inequality, spectators' noninherited inequality fairness type is indicated by the color of the corresponding circle. Recall from Section 3.2 that we would not expect subjects who were classified as egalitarians and libertarians to display differential redistribution patterns if inequality is inherited. Hence,

we should observe that green dots ($\tau_{i,NI} = E$) are situated in the upper right quarter of the figure, and that orange dots ($\tau_{i,NI} = L$) are situated in the lower left quarter. For meritocrats (teal circles), the theoretical prediction is vague: depending on α—how they weigh fairness toward workers versus beneficiaries—they should either behave meritocratically ($\alpha > 0.5$, lower right quarter) or in an egalitarian way ($\alpha < 0.5$, upper right quarter).

The figure shows that, just like before, many spectators behave very consistently and are either placed on a corner or on an edge. Most spectators "remain in their quarter", that is, display similar redistribution patterns in situations featuring inherited and noninherited inequality. Focusing on those spectators who have been classified as meritocrats under noninherited inequality, we see that only a few switch to an egalitarian redistribution pattern when inequality is inherited. This indicates that most of them prioritize fairness toward the workers ($\alpha > 0.5$). In contrast to our expectations, we observe some switching between meritocrats and libertarians.

These observations are quantified in the moving matrix displayed in Figure 3.4, which shows the distribution of two-dimensional redistribution patterns in a more condensed way. The position on the vertical axis describes spectators' fairness type
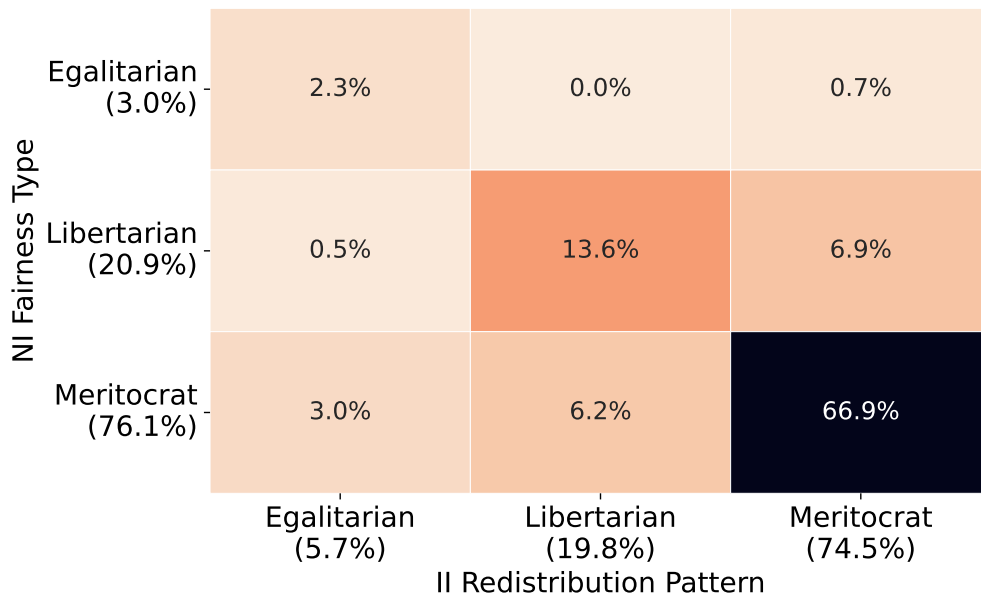
Figure 3.4. Two-Dimensional Redistribution Patterns
Note: This moving matrix displays the distribution of spectators over two-dimensional redistribution patterns. Fairness types under noninherited inequality are shown on the vertical axis. Redistribution patterns under inherited inequality are shown on the horizontal axis.

under noninherited inequality, and the position on the horizontal axis describes their redistribution pattern under inherited inequality.[8] Marginal distributions are

8. The figure disregards two spectators who are nonclassified in at least one dimension.

reported with the axis labels. The figure shows that most spectators are "on the diagonal", that is, they display the same redistribution pattern under both inherited and noninherited inequality. Only 3% of all spectators in the restricted sample switch from meritocratic to egalitarian, meaning that they prioritize fairness toward beneficiaries ($\alpha < 0.5$ in the theoretical framework). Between 6% and 7% of spectators each switch from meritocratic to libertarian or vice versa, which is not consistent with our theoretical framework and suggests that this may be more than just noise. Besides that, there are only very few "inconsistent" spectators. Overall, more than 85% of spectators are classified in a way that is consistent with our theoretical framework, which—together with the observation that spectators make very consistent observations *within* each condition—indicates that the framework explains spectators' behavior well.

As shown theoretically in Section 3.2, the fact that the money is distributed between passive stakeholders who differentially profit from their friends' effort in the INHERITED INEQUALITY conditions should only matter for meritocrats, and only if the initial distribution reflects relative effort. To formally test whether this is the case, we estimate regression Equation 3.8 using OLS and clustering standard errors on the spectator level. We are particularly interested in the triple interaction of the INHERITED INEQUALITY and EFFORT indicators (II_σ and E_σ) with spectators' (noninherited inequality) fairness type. The results are displayed in Table 3.4, in which some coefficients are suppressed for increased readability.[9]

The estimates in column (1), which corresponds to Equation 3.8 and uses egalitarians as the reference fairness type, show that the triple interaction effect amounts to 24.3%p and is significant for meritocrats. This indicates that, relative to egalitarians, the fact that inequality is inherited nudges meritocrats more strongly to redistribute more if inequality is based on effort instead of luck. As the triple interaction effect for meritocrats is also significantly higher than that for libertarians (Wald test, $p < 0.0001$), the data formally yields strong support for Hypothesis 3.4.

Result 3.4. *The fact that inheritance increases the extent of redistribution more strongly if inequality is based on effort instead of luck is driven by meritocrats.*

Considering columns (2) - (4), where Equation 3.7 is estimated separately for the three fairness types, it becomes apparent that the data do not perfectly fit the story behind Hypothesis 3.4, though. While the interaction effect of INHERITED INEQUALITY and EFFORT amounts to almost 10%p for meritocrats and is highly significant, in the LUCK domain they redistribute on average about 6%p less if inequality is inherited, which is a significant difference as well. Conversely, libertarians redistribute on average about 27%p more if inequality is inherited in the LUCK domain,

---

9. For a regression table that reports the same regression equations but does not omit coefficients, please refer to Table 3.D.2 in Section 3.D.

Table 3.4. Treatment Effects on the Extent of Redistribution by Fairness Type

	Dependent Variable: Extent of Redistribution ($\theta_{i,c}$, Share)			
	(1) Pooled	(2) Egalitarians	(3) Meritocrats	(4) Libertarians
EFFORT (E_σ)	-0.025	-0.025	-0.960***	-0.109***
	(0.036)	(0.038)	(0.006)	(0.018)
INHERITED (II_σ)	-0.018	-0.017	-0.059***	0.268***
	(0.031)	(0.032)	(0.012)	(0.042)
EFFORT (E_σ) × INHERITED (II_σ)	-0.144	-0.144	0.099***	-0.232***
	(0.103)	(0.108)	(0.015)	(0.044)
EFFORT (E_σ) × INHERITED (II_σ) × Meritocrat	0.243**			
	(0.104)			
EFFORT (E_σ) × INHERITED (II_σ) × Libertarian	-0.088			
	(0.112)			
Initial Inequality (Δ_σ)	0.031**	-0.052	-0.004	0.175***
	(0.014)	(0.101)	(0.012)	(0.045)
Constant	0.977***	1.001***	0.977***	0.084***
	(0.015)	(0.036)	(0.006)	(0.019)
Clusters	437	13	332	91
Observations	8399	249	6403	1731
R^2	0.817	0.106	0.864	0.228

Note: This table reports results from OLS regressions of the extent of redistribution implemented by spectator i in situation σ on treatment indicators and spectator i's fairness type, controlling for the initial extent of inequality in situation σ. Results are based on observations in the restricted sample. Column (1) corresponds to Equation 3.8. Columns (2) - (4) correspond to Equation 3.7 but are estimated on subsets of spectators who share the corresponding fairness type. Standard errors (in parentheses) are clustered on the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

while the interaction effect largely offsets this difference ($-23\%p$) for the EFFORT domain, and both coefficients are highly significant again.

3.5.3 Potential Channels

3.5.3.1 Spectators' Explanations for Their Redistribution Decisions

Why do spectators redistribute so little when they face the dilemma of meritocracy? To develop an understanding of how people reason about the dilemma and to generate hypotheses for potential channels, we analyze the open-ended explanations subjects gave for their redistribution decisions. Most spectators use the opportunity to write open-ended explanations after each decision block. For all open-ended explanation fields, more than 98% of spectators make an entry. Figure 3.B.3 in Sec-

tion 3.B shows that responses correspond well to treatment arms and fairness types. Hence, open-ended responses seem to provide useful information.

To get an overview of how spectators explain their decisions, we sort all mentioned explanations by hand into categories. Table 3.D.3 shows the complete list of categories and gives examples of the kind of explanations they encompass. Most spectators state specific rationales for their behavior. Yet, 49 spectators do not explain their decisions or use explanations like "I just tried to be fair", which cannot be assigned to a meaningful category. Consequently, our analysis excludes these spectators and is based on the remaining 388 subjects, who comprise about 89% of the spectators in the restricted sample.

Figure 3.5 depicts the frequencies with which explanations for redistribution decisions in Inherited Inequality & Effort are given by the explanation category. The plurality of spectators mentions that they implemented final allocations proportional to relative efforts without specifying whether that refers to the efforts of the workers or the (non-existent) efforts of the friends. Of those who specify this, most refer to the workers' efforts and few to the friends' efforts, which is consistent with our results for the redistribution decisions. The three corresponding categories contain nearly 82% of all explanations. Hence, relative effort levels appear to be the main theme behind redistribution decisions.
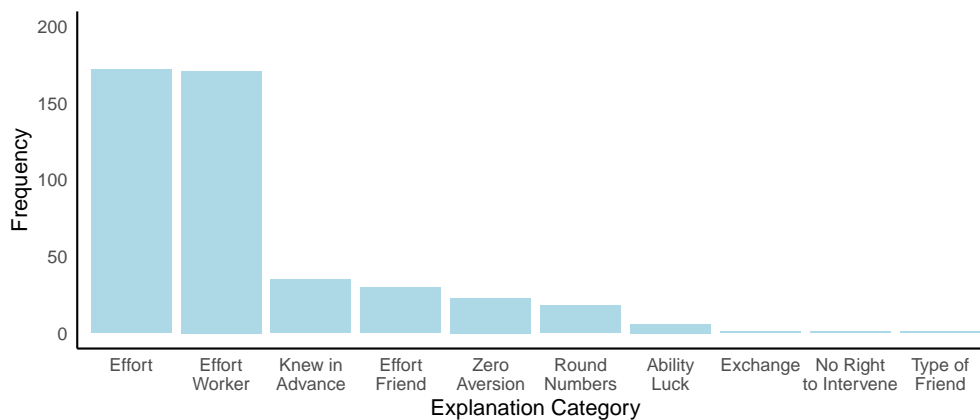
Figure 3.5. **Spectators' Explanations for their Decisions in Inherited Inequality & Effort**
Note: This figure displays the frequency of explanations spectators gave for their redistribution decisions in Inherited Inequality & Effort by explanation category. Results are based on up to three arguments made by the 388 spectators from the restricted sample who gave specific explanations for their behavior. We included up to three arguments per spectator.

Alternative explanations are much less frequently mentioned by spectators. For instance, it is conceivable that a worker's effort changes the spectators' belief about what kind of person the respective friend is. However, only a single spectator mentions this as relevant to his decision. Similarly, only one spectator mentions being influenced by the thought that workers and their friends might exchange money after the experiment. Slightly more frequently mentioned explanation categories in-

clude that subjects "Knew in Advance" and agreed to the rules of the study, such that redistribution would mean an unfair ex-post rule adjustment[10]; an aversion to giving people zero or very little money; a preference for round numbers; the idea that some people might have been less able to perform the task due to bad luck; and the belief that one must not intervene in the affairs of others. Figure 3.B.5, Figure 3.B.4 and Figure 3.B.6 in the appendix show similar results for the other three treatment conditions. Consistent with our other results, most spectators in each condition argue that earnings should be based on effort but not on luck.

Why do most spectators base their decisions on the relative efforts of the workers rather than on the relative efforts of the friends? To examine this question, we focus on the explanations of spectators in INHERITED INEQUALITY & EFFORT who acknowledge the dilemma of meritocracy, because they consciously think about fairness toward the workers versus fairness toward the friends. We consider a spectator to acknowledge the dilemma of meritocracy if he provides arguments for and against redistribution based on the meritocratic fairness ideal in his explanation. Due to this strong selection requirement, this includes only 25 spectators who provide 34 arguments collectively.

Figure 3.6 shows the frequencies of explanation categories spectators use to rationalize their decisions. About 82% of all explanations belong to two categories: explanations in the "Worker Entitled" category argue that the workers are entitled to the fruits of their labor. Conversely, explanations in the "Friend Not Entitled" category state that, in contrast to workers, friends are not entitled to the bonus payment because they did not earn it through effort. Both explanation categories refer to the same asymmetry between workers and friends: workers work for the bonus while friends do not. In the view of most spectators who mentioned the dilemma of meritocracy, this makes the entitlement of workers stronger than the entitlement of friends. This can explain why most spectators prefer to be fair toward the workers rather than toward their friends.

Again, alternative explanations are mentioned much less frequently. About 6% of the respondents mention that priority should be given to friends precisely because they did not work and are therefore blameless for the initial distribution. Another 6% view a worker and his friends as one team and argue that resources that were earned by the team should remain within the team. One respondent expects the friend to return some of his earnings to his associated worker and another respondent argues that a friend who is not worked for is not worth the work.

Hence, most spectators seem to believe workers earned the right to distribute a monetary amount that is proportional to their relative effort levels. While spectators might at the same time find it unfair that some passive friends receive less than

---

10. As described in Section 3.3 workers were informed that their (or their friend's) payoff could be affected by the decision of a third person, and spectators knew that. Spectators who refer to this issue apparently still consider altering the initial distribution an unfair rule adjustment.
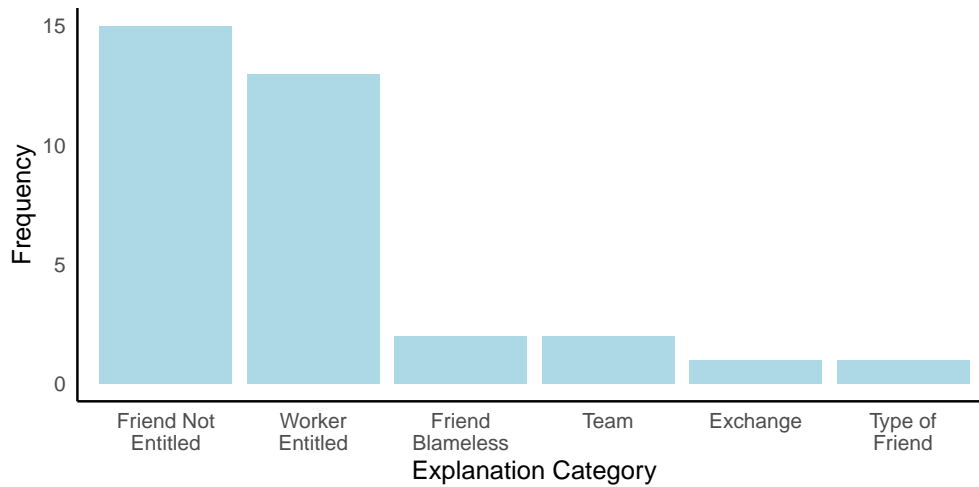
Figure 3.6. Spectators' Explanations for Resolving the Dilemma of Meritocracy
Note: This figure displays the frequency of explanations spectators gave for resolving the dilemma of meritocracy in the way they did by explanation category. Results are based on up to three arguments made by 25 spectators from the restricted sample who mentioned the dilemma of meritocracy in their explanations.

others even though neither of them worked themselves, the former consideration might be perceived as more important. These considerations suggest that in the EFFORT conditions (meritocratic) spectators' redistribution decisions should depend on their belief about workers' preferred distributions. For example, a spectator might equalize the distribution between passive friends based on the belief that workers prefer a 50/50 split. Conversely, a spectator who believes that workers only care about their own friends might not redistribute to respect workers' preferences.

3.5.3.2   Redistribution Decisions and Spectators' Beliefs about Workers' Preferences

To pursue this potential explanation, we make use of spectators' beliefs about how workers would distribute money in a dictator game between a) themselves and another worker and b) their own friend and the friend of another worker, elicited subsequent to the redistribution blocks.[11] If spectators indeed make merit judgments based on workers' relative effort and then try to respect their distributional preferences (in particular: those of the more industrious worker), we should observe that these beliefs are associated with the average extent of redistribution implemented by spectators. We should further observe that these are stronger in the EFFORT conditions and driven by meritocrats.

To test these predictions, we proceed in two steps. First, we regress subjects' average extent of redistribution in a given condition on the corresponding belief about

11. Histograms of these beliefs and the individual-level differences in these beliefs are shown in Figure 3.C.3 and Figure 3.C.4 in Section 3.C.

workers' preferred distribution. To make estimates comparable across conditions, we standardize both the dependent variable (across spectators but within conditions) as well as the independent variable into z-scores. Formally, we estimate the following regression equation using OLS:

$$std(\bar{\theta}_{i,c}) = \alpha + \beta_{c,k} \cdot std(\mu_{i,k}) + \varepsilon_{i,c,k}. \tag{3.11}$$

As usual, $\bar{\theta}_{i,c}$ is the average extent of redistribution implemented by spectator i in condition $c \in$ {NI-L, NI-E, II-L, II-E}. $\mu_{i,k}$ describes the belief of spectator i about workers' preferred distributions in case k, with k indicating which dictator decision is used: for $c \in$ {NI-L, NI-E} we use spectators' beliefs about workers' preferred distribution between themselves and the other worker, and for $c \in$ {II-L, II-E} we use spectators' beliefs about workers' preferred distribution between their own friend and the friend of the other worker. The coefficients from these regressions are displayed in Figure 3.7.
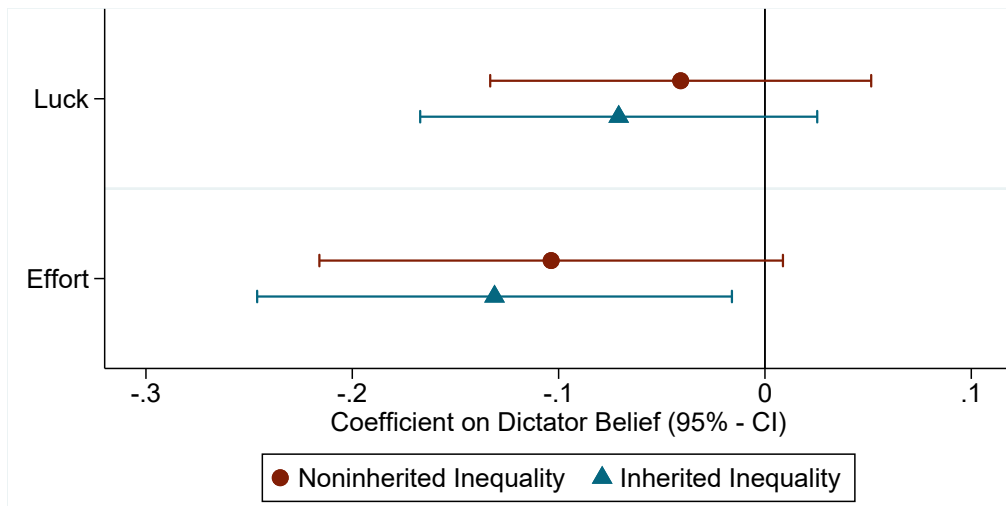
Figure 3.7. Association between Beliefs about Workers' Preferences and Redistribution Decisions

Note: This figure displays coefficients on spectators' beliefs about workers' preferred distributions, obtained from separate regressions of redistribution levels (standardized across spectators but within conditions) on the corresponding standardized beliefs (see Equation 3.11). The corresponding regression results are reported in Table 3.D.4 in Section 3.D.

In NONINHERITED INEQUALITY & LUCK, an increase of one standard deviation (SD) in the belief about the share of the $10 workers on average keep for themselves is associated with a 0.04 SD reduction in the average extent of redistribution ($p = 0.39$). With a 1 SD increase in the same belief being associated with a 0.10 decrease in the average extent of redistribution, the estimate for the NONINHERITED INEQUALITY & EFFORT conditions is more than twice as large and weakly significant ($p = 0.07$). In the INHERITED INEQUALITY domain, the pattern is very similar

but estimated coefficients a bit larger in terms of absolute value. In Inherited Inequality & Luck, a 1 SD increase in the belief about the share of the \$10 workers on average give to their own friends is associated with a 0.07 SD decrease in the average extent of redistribution ($p = 0.15$). Again, with a 1 SD increase in the belief being associated with a 0.13 SD decrease in the average extent of redistribution, the same estimate for the Inherited Inequality & Effort condition is about twice as large and statistically significant ($p = 0.03$). These patterns indicate that spectators' beliefs about workers' preferred distributions are, in particular in the Effort case, indeed associated with their redistribution decisions in the expected way.

As a second step, we test the more nuanced prediction that these associations are most pronounced for spectators classified as meritocrats in the Noninherited Inequality domain. We estimate the same regression equation as before, but separately for the three fairness types and, to increase comparability of effects across types, standardizing the belief (redistribution) variable not across all spectators (and within a given condition), but across spectators of a given type (and within a given condition). The results for the Effort domain, reported in Table 3.5, are mixed.[12]

Table 3.5. Association between Beliefs and Redistribution Decisions by Fairness Type

	Dependent Variable: Average Extent of Redistribution ($\bar{\theta}_{i,c}$, Z-score)					
	Noninherited Inequality			Inherited Inequality		
	(1) Egalitarians	(2) Meritocrats	(3) Libertarians	(4) Egalitarians	(5) Meritocrats	(6) Libertarians
Guess Self/Other	0.244* (0.134)	0.043 (0.045)	-0.089 (0.089)			
Guess Own/Other's Friend				-0.246 (0.291)	-0.115 (0.075)	0.036 (0.136)
Observations	13	332	91	13	332	91
R^2	0.060	0.002	0.008	0.060	0.013	0.001

Note: This table reports results from OLS regressions of spectators' average extent of redistribution in the two Effort conditions, standardized across spectators of a given (Noninherited Inequality) fairness type and within experimental conditions, on their beliefs about workers' preferred distributions, standardized across spectators of the same fairness type. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

While our sample includes too few egalitarians to consider the corresponding estimates reliable (columns (1) and (4)), the estimates for meritocrats (columns (2) and (5)) and libertarians (columns (3) and (6)) are insignificant. Focusing on meritocrats, we observe that in the Noninherited Inequality & Effort condition,

---

12. For completeness, a similar regression table reporting the results for the Luck domain can be found here: Table 3.D.5.

the association goes in the wrong direction ($p = 0.34$). In the Inherited Inequality & Effort condition, a 1 SD increase in the belief about the share workers on average keep for their own friends is associated with a 0.12 SD decrease in the average extent of redistribution among meritocrats. This effect, however, does not reach statistical significance ($p = 0.13$).

Overall, our observations on the relation of spectators' beliefs about workers' preferences and their redistribution decisions suggest that spectators making merit judgments and then seeking to respect (the more diligent) workers' preferences may be a part of what is behind our results. However, the associations documented in the first step seem to be driven to some extent by differentially distributed beliefs across different fairness types, and this potential explanation requires a more thorough investigation.[13]

3.5.4 Heterogeneity between Demographic Groups

The previous analysis has shown that most people do not redistribute in the Inherited Inequality & Effort treatment. To investigate whether this result masks heterogeneity between sociodemographic groups, we construct binary sample splits along a variety of dimensions and test whether spectators on different sides of these sample splits make different redistribution decisions. We consider the following sociodemographic characteristics: age, voting frequency (below vs. above median); sex (female vs. male); education (college degree vs. no college degree); income (below vs. above $68,000$); wealth (below vs. above $124,000$); party identification (republican vs. democrat); perceived social class (above vs. below middle class); and economic ideology (state- vs. market-oriented).[14] Because we have not preregistered any hypotheses regarding heterogeneity, we rely on the main sample for this exercise.

For the different sample splits, Figure 3.8 displays subgroup averages (with equal weights) of spectators' average extent of redistribution in Inherited Inequality & Effort. Heterogeneity is most pronounced along the wealth dimension. This is consistent with the notion that inherited inequality can be considered just from the perspective of those who bequest but unjust from the perspective of those who inherit — the key idea behind the dilemma of meritocracy. High-wealth individuals might be more likely to take the benefactors' perspective while for low-wealth in-

---

13. The average beliefs about the share workers on average keep for themselves (when they distribute between themselves and the worker they are matched to) are $4.98 (Egalitarians), $6.14 (Meritocrats), and $6.35 (Libertarians). The average beliefs about the share workers on average give to their own friends (when they distribute between their own friend and the friend of the worker they are matched to) are $5.20 (Egalitarians), $6.13 (Meritocrats), and $6.22 (Libertarians).

14. When spectators reported their political affiliation, perceived social class, and economic ideology, they could select a middle option; when we consider these sociodemographic dimensions, we drop spectators who selected this middle option.
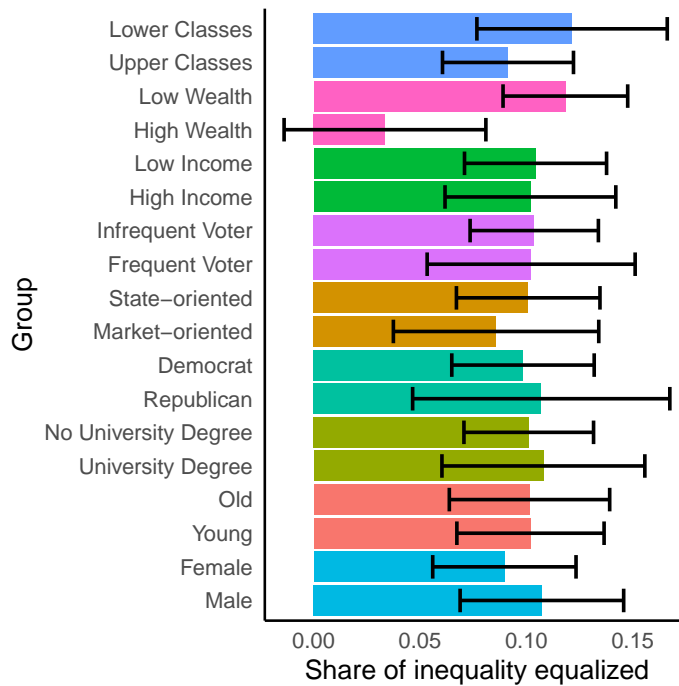
Figure 3.8. Average Equalization in Condition Inherited Inequality & Effort by Demographic Group

Note: Shares of inequality equalized for a group are calculated by averaging over the average extent of redistribution in the Inherited Inequality & Effort condition for all spectators in the main sample who belong to the group. 95% confidence intervals around the averages based on standard errors of the mean.

dividuals the beneficiaries' perspective might be more salient. Similarly, those from the upper classes tend to redistribute less than those from the lower classes.[15]

Yet, there is not much heterogeneity overall; in particular, Democrats and Republicans redistribute to a similar extent on average, and no subgroup equalizes more than $12 of the initial inequality on average. As shown in Figure 3.C.5, Figure 3.C.6 and Figure 3.C.7 in Section 3.C, the patterns in Noninherited Inequality & Effort closely resemble those in Inherited Inequality & Effort displayed here, and heterogeneity in the two Luck conditions is even less pronounced.

To test formally whether there is heterogeneity in the treatment effects across any of the binary splits in the Inherited Inequality & Effort condition, we run the following OLS regression:

---

15. A potential explanation for heterogeneity along the wealth/socio-economic status dimension could be that individuals take perspectives, endorse fairness ideals, and form beliefs in a self-serving way (Konow, 2000; Rodriguez-Lara and Moreno-Garrido, 2012; Deffains, Espinosa, and Thöni, 2016; Cassar and Klein, 2019; Valero, 2022).

$$\theta_{i,\sigma} = \alpha + \alpha^D D_i + \alpha_E E_\sigma + \alpha_E^D E_\sigma D_i + \beta II_\sigma + \beta^D II_\sigma D_i \qquad (3.12)$$
$$+\beta_E E_\sigma II_\sigma + \beta_E^D E_\sigma II_\sigma D_i + \delta \Delta_\sigma + \epsilon_{i,\sigma}$$

where D_i indicates whether spectator i belongs to a certain sociodemographic subgroup. We cluster standard errors on the spectator level. Figure 3.C.8 in the appendix plots estimates for β^D and β_E^D by demographic variable, which describe the differences across the sample split in a) the effect of inequality being inherited in the luck domain and b) the "difference-in-differences" effect of inequality being inherited in the effort versus luck domain. Table 3.D.6 and Table 3.D.7 in Section 3.D also report estimated coefficients on other variables. Few estimates for β^D and β_E^D are significant before controlling for multiple hypothesis testing, and after applying the Benjamini-Hochberg procedure none of the coefficients differs significantly from zero. Hence, resolving the dilemma of meritocracy in favor of those who bequest is common across sociodemographic groups.

To explore whether the distribution of redistribution patterns differs by socioeconomic characteristics, we calculate for each demographic subgroup the distribution over the two-dimensional redistribution patterns $(\tau_{NI}, \tau_{II}) \in \{$(Egalitarian, Egalitarian), (Libertarian, Libertarian), (Meritocrat, Meritocrat), (Meritocrat, Egalitarian)$\}$, which are consistent with our theoretical framework, and a residual type which encompasses all remaining spectators. Figure 3.C.9 in Section 3.C shows the resulting distribution of redistribution patterns by demographic subgroups. There is no notable variation between demographic subgroups. In each subgroup, most spectators can be classified into one of the four main patterns, and in each subgroup more than half of all spectators display a meritocratic redistribution pattern in both dimensions. Using Fisher's exact test, we do not detect any significant differences in the distribution between any two subgroups of the same demographic variable.

3.5.5 External Validity

As a next step, we investigate to what extent our experimental measures of redistributional preferences are associated with preferences over real-world policies elicited in the post-experimental questionnaire. Because spectators' average extent of redistribution is highly correlated both within the Luck and Effort domain ($\rho_{\bar{\theta}_{i,NI-L},\bar{\theta}_{i,II-L}} = 0.64$ and $\rho_{\bar{\theta}_{i,NI-E},\bar{\theta}_{i,II-E}} = 0.60$), we apply a factor analysis on the four variables that capture an individual's tendency to redistribute in the four conditions, retaining two factors (eigenvalues equal to 1.11 and 0.91; -0.21 for the third factor). $\bar{\theta}_{i,NI-L}$ and $\bar{\theta}_{i,II-L}$ load heavily on the first factor (0.73 in both cases) but not the second one (0.02 and 0.03). Conversely $\bar{\theta}_{i,NI-E}$ and $\bar{\theta}_{i,II-E}$ load heavily on the second factor (0.69 in both cases) but not the first one (0.02 and 0.04). Hence, we conclude that the first factor captures an individual's preference for redistribution if inequality is based on luck ("Redistribution (Luck)"), while the second factor captures the preference for redistribution if inequality is the result of differential effort ("Redistribution (Effort)").

In the questionnaire, we elicited preferences regarding six inequality-related policies. First, we asked spectators to indicate their preferred maximum marginal income and estate tax rates on scales from $0\% - 100\%$. Second, we used 7-point Likert scales to elicit their support for disability insurance, unemployment insurance, and equal opportunity programs, with options ranging from "[the policy] should be significantly reduced" to "significantly extended". Finally, we asked to what extent spectators find intergenerational transmission fair, eliciting responses by means of a 6-point Likert scale from "clearly unfair" to "clearly fair". To facilitate the analysis, we reverse-coded the last variable such that higher values always indicate stronger support for redistribution. Further, we standardized all policy variables and the two factor variables.

Figure 3.9 displays coefficients from OLS regressions of the policy variables on the two factor variables. Without exception, the estimated coefficients are positive,
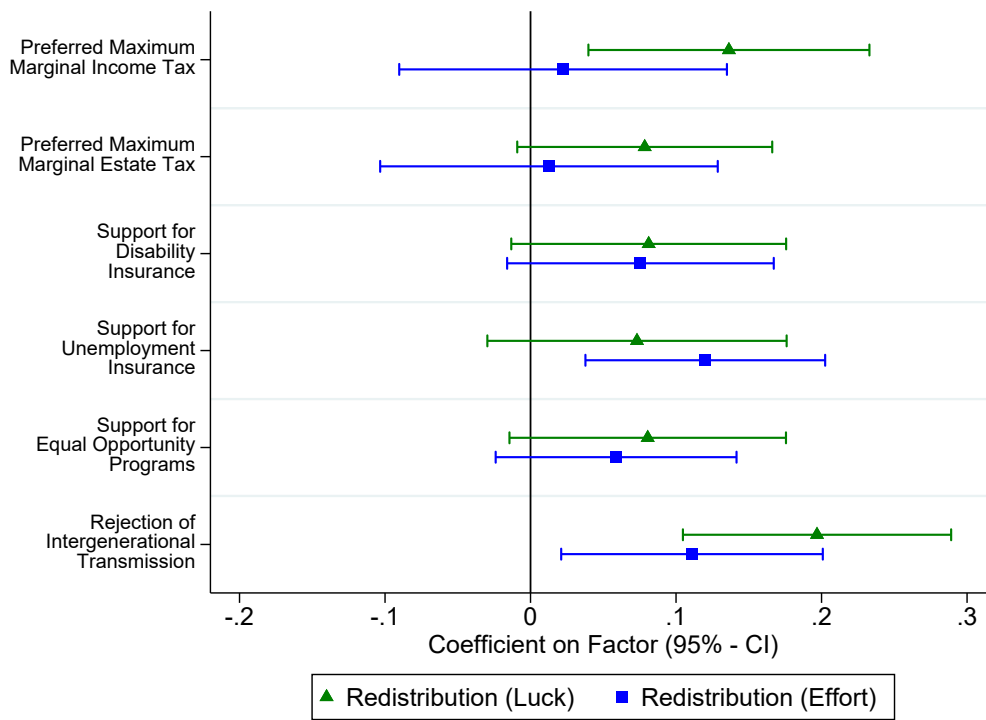
Figure 3.9. Association between Experimental Measures and Policy Preferences
Note: This figure plots coefficients from OLS regressions of spectators' (standardized) policy preferences on (standardized) factor variables based on the average extent of redistribution in the four treatment conditions. 95% confidence intervals are based on robust standard errors. The corresponding regressions are reported in Table 3.D.8. Results are based on the main sample.

indicating that more redistribution in the impartial spectator experiment is associated with stronger support for redistributive policies. A 1SD increase in one of the factor variables is often associated with an increase in support for the respective policy by about 0.1SD. Given that recent research has shown that preferences over

real-world (redistributive) policies are strongly influenced by factors other than inequality preferences such as views on government efficiency (Stantcheva, 2021), it is perhaps unsurprising that the associations are not too strong. However, for all policy variables, at least one of the two factor variables is significant at the 10%-level. In sum, the results suggest that the experimental measures capture meaningful information about individuals' fairness preferences, and that these preferences are associated with preferences over real-world (redistributive) policies.

3.6 Conclusion

Human beings tend to be more altruistic toward their family members, friends, and compatriots than toward non-relatives, strangers and foreigners (Bernhard, Fischbacher, and Fehr, 2006; Cappelen, Enke, and Tungodden, 2022). In many instances the underlying relationships are accidental; for example, we do not choose to which parents or in which country we are born. In meritocratic societies where inequality is accepted if it is based on factors within individuals' control but rejected if it is based on factors outside individuals' control, this creates a fundamental dilemma: unequal outcomes between individuals who differentially profit from other people's efforts are at the same time within the benefactors' control (and therefore just) but outside the beneficiaries' control (and therefore unjust). This paper studied US citizens' fairness preferences in situations with such inherited inequality and how they deal with this dilemma.

Our results show that most US citizens prioritize the benefactors' efforts and accept inherited inequality, which can help to explain why many people accept high levels of inequality and unequal starting positions within and across societies. It is not that they find it fair that some people have better opportunities than others; rather, they weigh this concern against another—in their view stronger—fairness argument. For example, creating equal opportunities among children requires preventing parents from channeling extra resources to their children, even if they themselves earned them fairly. When meritocrats have to decide whether to accept unequal opportunities or prevent families or friends from endowing their loved ones with extra endowments, our results suggest that they choose the former.

Since we find that individuals clearly prioritize rewarding the benefactors' efforts over equalizing payoffs between the non-working beneficiaries when facing the dilemma of meritocracy, a natural avenue for future research is to explore how much the decision environment has to be tweaked for spectators to redistribute more. Our setup is ideally suited to do so because it admits controlled variation in a variety of dimensions.

One potentially relevant dimension is the relationship between benefactor and beneficiary, which varies between outside-the-lab contexts. For example, people usually bequest their resources to their children, and the parent-child relationship is

usually stronger than the relationship between friends (Cappelen, Enke, and Tungodden, 2022). In light of our finding that spectators tend to redistribute less if they think that workers tend to prioritize their own friends more strongly, it seems unlikely that the results would differ if we had used family ties instead of friendships, where redistribution levels are already low. Instead, redistribution in the friends-case likely poses an upper bound to redistribution in the family case. Still, spectators might view kinship differently from friendships because people can choose their friends but not their kin. To examine this possibility, researchers could combine our experimental design with a subject sample containing pairs of relatives.

The size of the stakes involved constitutes a second dimension that might be relevant for fairness judgments. High stakes may not only induce individuals to make considerate decisions but, in the context of redistribution, also call into play different motivations such as taking into account individuals' needs (Konow, 2001). Further, employing high stakes may also enable researchers to study preferences over more nuanced (e.g., progressive) redistribution schemes. While the correlation between spectators' behavior in our experiment and their policy preferences indicates that a lot can be learned also from small-stakes settings, it might be worthwhile to study how the stake size affects the relevance of different fairness motives and overall fairness judgments.

Third, our Effort and Luck treatments make it very clear that the initial distribution is either exclusively determined by workers' relative efforts or by luck, whereas resource distributions are usually determined by a combination of the two that is hard to disentangle. Recent research has documented in the context of non-inherited inequality that if inequality is based on both effort and luck, this affects redistribution behavior in a non-trivial way. For example, spectators prioritize rewarding effort when the relative contribution of effort and luck can be decomposed (Cappelen and Tungodden, 2017), but uncertainty induces meritocrats to behave in a more egalitarian way (Cappelen, Mollerstrom, et al., 2022). Similarly, uncertainty allows individuals to form biased beliefs about the source of inequality (Konow, 2000; Rodriguez-Lara and Moreno-Garrido, 2012; Deffains, Espinosa, and Thöni, 2016; Cassar and Klein, 2019; Valero, 2022). Hence, it might be interesting to study how uncertainty about the source of inequality affects preferences for redistribution in the context of inherited inequality.

Fourth, individuals may not only inherit differential amounts of resources that can be consumed but also differential opportunities to generate resources themselves. Some papers investigate preferences for redistribution under unequal opportunities, albeit in settings where those unequal opportunities arise exogenously (Eisenkopf, Fischbacher, and Föllmi-Heusi, 2013; Alesina, Stantcheva, and Teso, 2018; Andre, 2022; Schwaiger et al., 2022). Our setup could easily be extended to accommodate the inheritance of unequal opportunities by introducing a second production stage in which the beneficiaries' returns to effort depend on their benefactors' efforts in the first production stage. This would introduce a dilemma similar

to the one studied in this paper because a meritocrat should reject unequal opportunities but welcome that higher effort in the first stage pays off for beneficiaries in the second stage, leading to a very different decision problem for individuals making fairness judgments as compared to those in the papers mentioned above.

Finally, we have provided suggestive evidence for a potential mechanism behind individuals' fairness judgments in the context of inherited inequality. Our observations — and also the results from Cohen, Maltz, and Ofek-Shanny (2022) — are consistent with the idea that individuals determine entitlements based on the benefactors' merits and then try to take into account the benefactors' preferences over resource distributions between potential beneficiaries when making fairness judgments. Devising a causal test of this mechanism seems to be a promising endeavor.

Appendix 3.A   Validation of Survey Items

To test whether short nonincentivized survey measures can be employed as substitutes, we asked spectators to what extent they find luck-based and effort-based inequality between two individuals fair. Responses were elicited by means of 6-point Likert scales ranging from "clearly unfair" to "clearly fair."[16]

   To assess how closely the experimental and survey measures are related, we run OLS regressions with the average extent of redistribution in either the NONINHERITED INEQUALITY & LUCK or the NONINHERITED INEQUALITY & EFFORT condition as the dependent variable and the (standardized) survey measures as the independent variable(s). The results are reported in Table 3.A.1 and indicate that the experimental measures of redistributional preferences are strongly related to the corresponding survey measure, but not related to the non-corresponding survey measure.

Table 3.A.1. Association between Experimental and Survey Measures of Redistributional Preferences

	Dependent Variable: Average Extent of Redistribution ($\bar{\theta}_{i,c}$, Share)					
	$\bar{\theta}_{i,NI-L}$			$\bar{\theta}_{i,NI-E}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Luck Survey Measure	-0.148***		-0.148***	0.004		0.003
	(0.017)		(0.017)	(0.010)		(0.010)
Effort Survey Measure		0.008	0.006		-0.067***	-0.066***
		(0.016)	(0.018)		(0.018)	(0.018)
Constant	0.799***	0.799***	0.799***	0.048***	0.048***	0.048***
	(0.016)	(0.017)	(0.016)	(0.008)	(0.008)	(0.008)
Observations	437	437	437	437	437	437
R^2	0.172	0.000	0.172	0.000	0.147	0.147

Note: This table reports results from OLS regressions of the average extent of redistribution in the NONINHERITED INEQUALITY & LUCK ($\bar{\theta}_{i,NI-L}$) and NONINHERITED INEQUALITY & EFFORT ($\bar{\theta}_{i,NI-E}$) conditions on the respective (standardized) survey measures. Results are based on the main sample. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

   Columns (1)-(3) refer to the average extent of redistribution in the NONINHERITED INEQUALITY & LUCK condition. We observe that a 1SD increase in the luck survey measure is associated with a decrease in the average extent of redistribution by almost 15%p. In contrast, there is no association at all between the experimental

----

16. The survey questions asked spectators to complete the sentences "If one person receives more than another due to having better luck, I find that ..." and "If one person receives more than another due to exerting higher effort, I find that ..." by selecting the option on the Likert scale that corresponded most closely to their view. Figure 3.C.10 in Section 3.C show cumulative distribution functions for the two survey questions.

measure for this condition and the effort survey measure. Conversely, focusing on the NONINHERITED INEQUALITY & EFFORT case in columns (4)-(6), a 1SD increase in the effort survey measure is associated with a $6 - 7\%p$ decrease in the average extent of redistribution, but there is no association between the experimental measure for this condition and the luck survey measure. These observations are corroborated by the fact that at least 15% of the variance in the average extent of redistribution is explained if the regression includes the "right" survey measure, but none of the variance is explained if only the "wrong" survey measure is included as a regressor. Overall, our results suggest that if researchers have to economize on survey content these nonincentivized survey measures constitute decent alternatives to elicit fairness preferences and even allow to differentiate between different sources of inequality.

Appendix 3.B   Data Quality

In this section, we detail how we tried to promote high-quality responses in the spectator survey and report various data quality checks. The data reveal that a) very few spectators fail attention checks, b) the vast majority states that the instructions were comprehensible, c) spectators make few errors on control questions, d) most spectators write detailed and thoughtful responses to open-ended questions, and e) few spectators perceive the survey to have been biased in either political direction.

Attention Checks.   The survey features two attention checks, and participants are informed on the first page that they will be rejected if they fail both of them. In line with Prolific's attention check policy, the first attention check instructs subjects to select prespecified options, and the second attention check is a nonsensical question for which only two options are objectively correct. Attention checks are placed strategically: one is administered right at the start of the survey, and the other one is administered as part of the policy preferences questionnaire and resembles the other questions at first glance. None of the 543 subjects who completed the spectator survey failed both attention checks, such that we do not have to exclude anyone in the main sample to follow our pre-analysis plan. Generally, few spectators failed attention checks at all: among the 543 spectators in the main sample, two failed the first attention check, and 15 failed the second attention check. Considering only the 437 spectators in the restricted sample (see Section 3.4.2), only one failed the first attention check, and 11 failed the second attention check.

Comprehensibility.   We attach great importance to not confronting spectators with walls of text. For example, we introduce them to each condition of the earnings stage and how they can make their redistribution decisions with the help of individual slideshows. Each slideshow displays graphical representations of the different steps in the earnings stage with only minimal text, and spectators can go back and forth

within each slideshow. The slideshow and the combination of visual and text information are designed to make the survey as engaging and easy to digest as possible.

At the end of the survey, we ask spectators how comprehensible they find the instructions. On a 7-point Likert scale, subjects can choose options from "not comprehensible at all" to "perfectly comprehensible". For spectators in the restricted sample, Figure 3.B.1 shows the distribution of the responses (the figure for the main sample looks very similar). We observe that spectators judge the instructions very favorably.
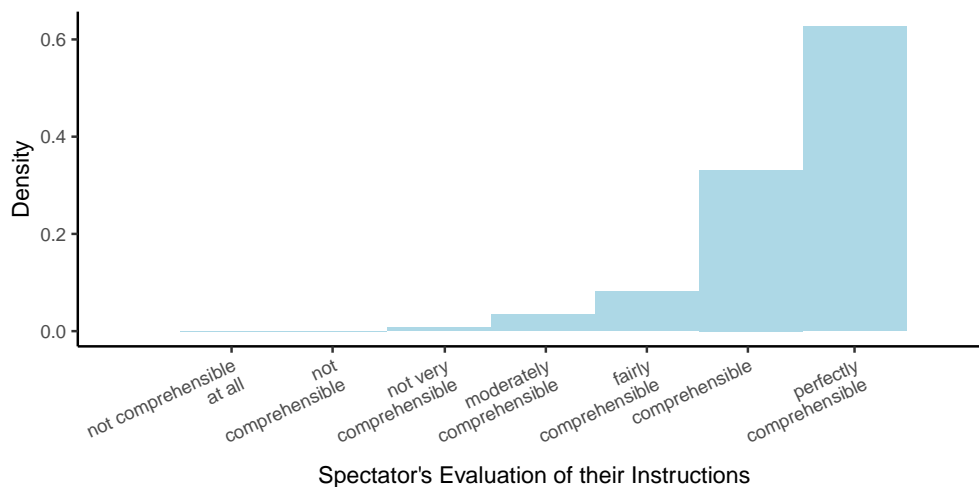

Figure 3.B.1. Spectators' Assessment of the Instructions

Note: Histogram showing how spectators in the restricted sample chose to complete the sentence "Overall, I found the instructions ..." on a 7-point Likert scale from "not comprehensible at all" to "perfectly comprehensible."

The vast majority (58%) say that the instructions were "perfectly comprehensible," and 89% assess the instructions as at least "fairly comprehensible." It is particularly reassuring that less than 1% of the spectators perceive the instructions as "not very comprehensible," and no one chooses the lowest two options.

Control Questions. To check more directly whether spectators understand the instructions, they have to answer two control questions each after they were introduced to a particular type of situation by means of the slideshow. They can proceed to the corresponding block of decisions only if they answered both questions correctly; otherwise, they are referred to the slideshow again. Control questions ask about the most crucial features of the situation: whether workers worked for themselves or friends and whether the initial allocation of the $10 would be based on a random draw or the relative number of completed tasks. In total, each spectator responds to eight control questions. Figure 3.B.2 depicts a histogram of the total number of errors spectators in our sample made. We observe that most spectators made few errors, which indicates that they usually understood the instructions well.
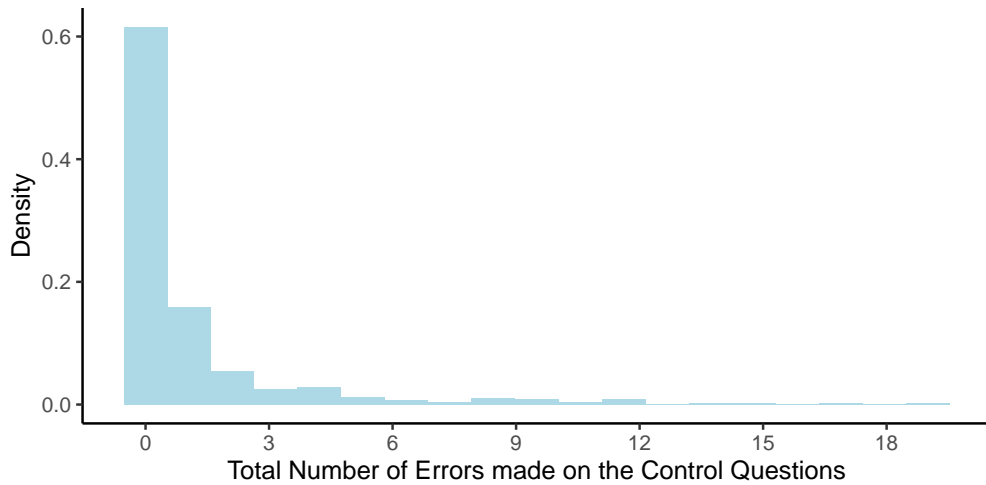
Figure 3.B.2. Control Question Errors
Note: Histogram of the total number of errors that spectators in the restricted sample made when responding to the eight control questions.

About 65% of spectators made no error, and only about 13% made more than two errors in total.

Open-Ended Questions. The spectator survey features several open-ended questions. After spectators have made all redistribution decisions within a particular block, we ask them to describe their considerations regarding these decisions. Further, at the end of the survey, subjects can leave a final comment on the general topic, the instructions, whether they experienced difficulties or anything else they have on their mind. Most open-ended responses are quite detailed and thoughtful. Only one spectator in the restricted sample (four spectators in the main sample) did not write any open-ended response during the study, suggesting that spectators generally put considerable effort into the study.

Figure 3.B.3 summarizes responses in four word clouds, one for each treatment. To generate these word clouds, we remove all numbers from the open-ended responses, transform all words to lowercase and remove punctuation and stop words. Finally, we reduce all words to their base word (stem). The size of words in Figure 3.B.3 indicates the frequency with which that word was used. The term "work" was among the most often used terms in all conditions, consistent with the large share of meritocrats in our sample. In the LUCK conditions, the term "equal" was also used very frequently, while it was nearly absent in the EFFORT conditions. Similarly, the term "friend" belongs to the most commonly used terms in the INHERITED INEQUALITY conditions but is rarely used in the NONINHERITED INEQUALITY treatments. This suggests that subjects understood the conditions and gave thoughtful explanations.

(a) Noninherited & Effort

(b) Inherited & Effort

(c) Noninherited & Luck

(d) Inherited & Luck
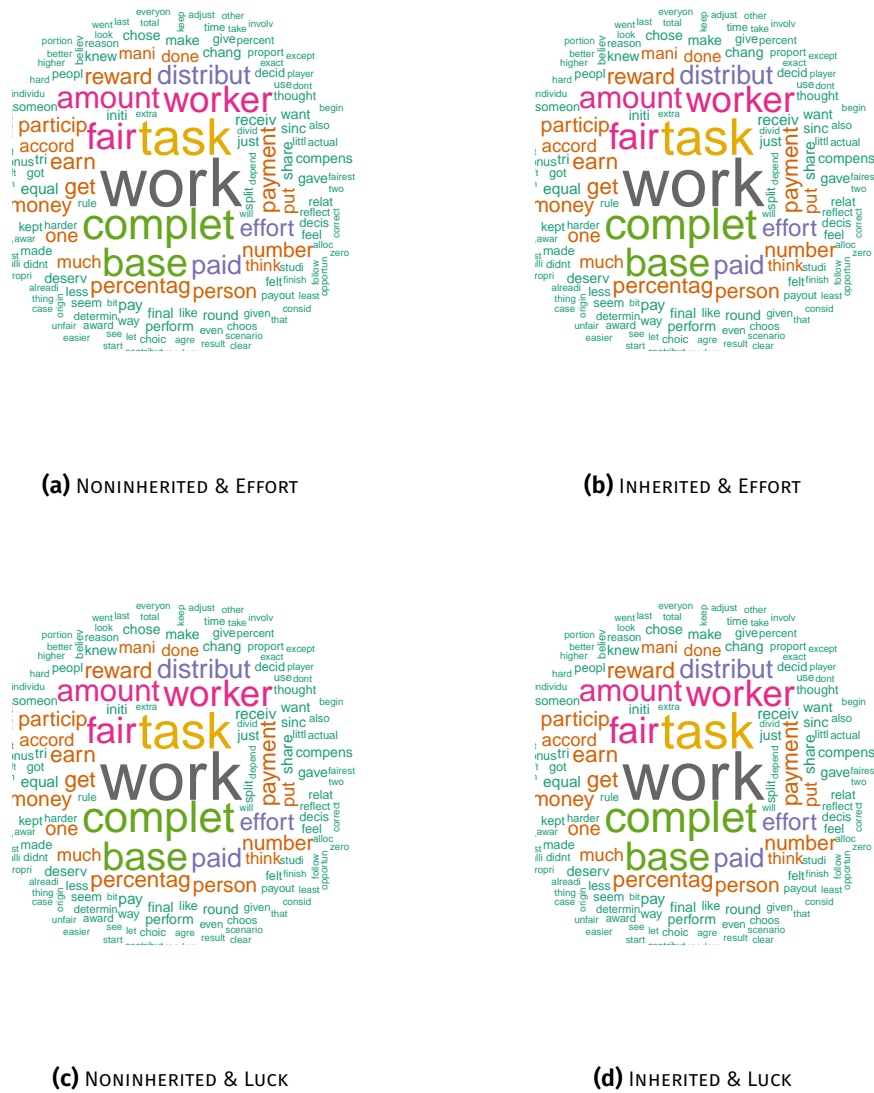
Figure 3.B.3. Word Clouds of Terms Subjects Used to Explain Their Considerations When Making Redistribution Decisions by Treatment Condition

Note: The size of the word relates to the frequency with which the word was chosen by spectators in the respective treatment.

Figure 3.B.4, Figure 3.B.5 and Figure 3.B.6 show the frequencies of explanations that spectators give for their decisions by explanation category.

Table 3.D.3 provides an overview of all categories with definitions and examples. Figure 3.B.4 shows that, consistent with their redistribution decisions, most spectators state to redistribute in the Noninherited Inequality & Effort condition based on the workers' efforts. Figure 3.B.5 reveals that most spectators rationalize

Figure 3.B.4. Spectators' Explanations for their Decisions in Noninherited Inequality & Effort
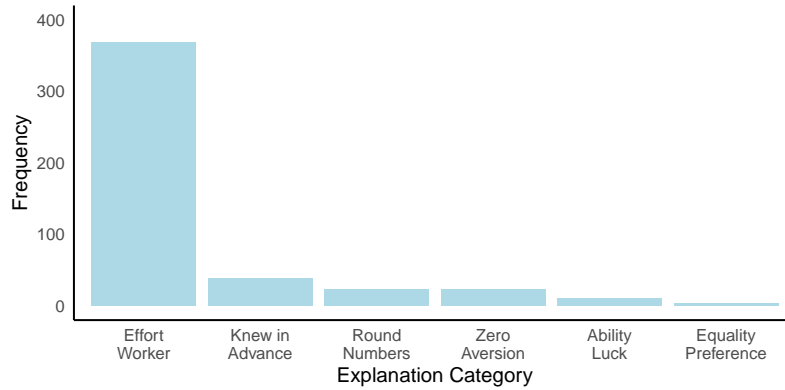*Note:* This figure displays the frequency of explanations spectators gave for their redistribution decisions in Noninherited Inequality & Effort by explanation category. Results are based on up to three arguments made by 432 spectators from the restricted sample. We included up to 3 arguments per spectator.
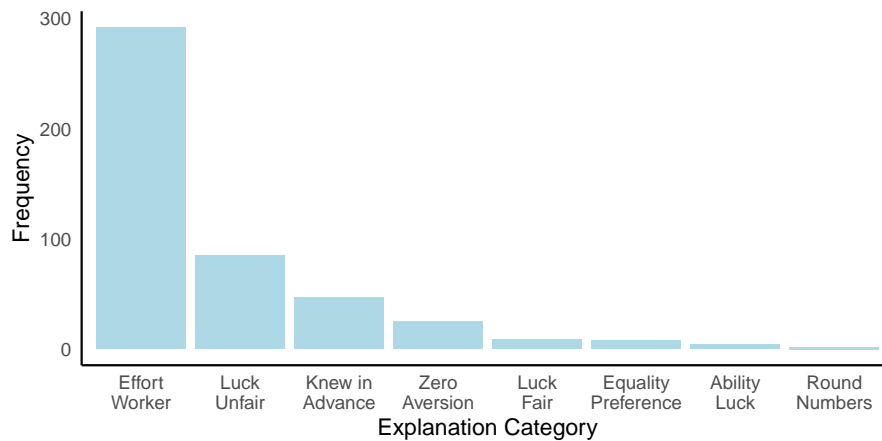
Figure 3.B.5. Spectators' Explanations for their Decisions in Noninherited Inequality & Luck
Note: This figure displays the frequency of explanations spectators gave for their redistribution decisions in Noninherited Inequality & Luck by explanation category. Results are based on up to three arguments made by 435 spectators from the restricted sample. We included up to three arguments per spectator.

their behavior in the Noninherited Inequality & Luck condition with a preference for a distribution based on effort too. However, many also mention that they find distributions based on luck unfair, while a few argue that the random allocation of resources is a fair method of distribution. Similarly, Figure 3.B.6 shows that many spectators justify their behavior in the Inherited Inequality & Luck treatment with arguments based on luck. Moreover, many spectators specifically refer to the effort of the workers or their friends. Hence, the explanations spectators give for their decisions correspond reasonably to the treatment conditions, which suggests that they had a good understanding of the study setup.

Figure 3.B.6. Spectators' Explanations for their Decisions in Inherited Inequality & Luck
Note: This figure displays the frequency of explanations spectators gave for their redistribution decisions in Inherited Inequality & Luck by explanation category. Results are based on up to three arguments made by 432 spectators from the restricted sample. We included up to three arguments per spectator.
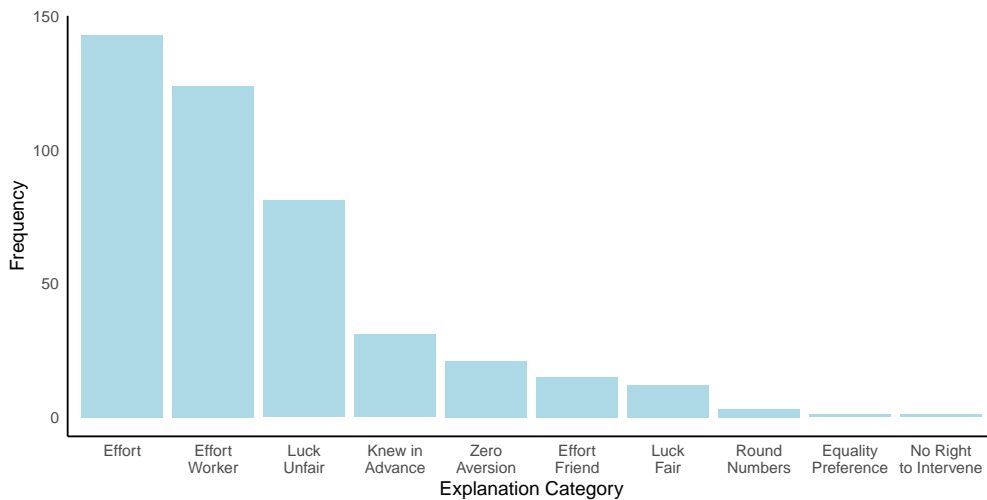
Finally, Figure 3.B.7 shows a word cloud of final comments spectators could make at the end of the survey. Again, to generate this word cloud, we remove all numbers from the open-ended responses, transform all words to lowercase and remove punctuation and stop words. Finally, we stem all words. Most comments are positive. Many spectators mention that they found the study interesting and understandable.

Political Bias. For surveys on highly politicized topics such as redistribution, it may be particularly important to phrase instructions and questions in a neutral way. We tried to keep this caveat in mind when we decided on the formulations used in the survey. Additionally, we ask subjects at the end of the survey whether they have the impression that the survey is biased toward a particular political stance, using a 7-point Likert scale with options from "strong left bias" to "strong right bias." Figure 3.B.8 displays how spectators' responses in the restricted sample are distributed (again, the figure for the main sample looks very similar). Less than 5% of the spectators perceive a strong bias in either direction. About 23% perceive a left-wing bias of any strength, whereas about 6% perceive a right-wing bias of any strength. More than 70% of the spectators in the restricted sample respond with "No or almost no bias," which is remarkable given that the theme of the survey is redistribution.

Figure 3.B.7. Word Cloud of Final Comments
Note: The figure refers to final comments spectators could make at the end of the survey. The size of the word relates to the frequency with which the word was chosen by spectators.
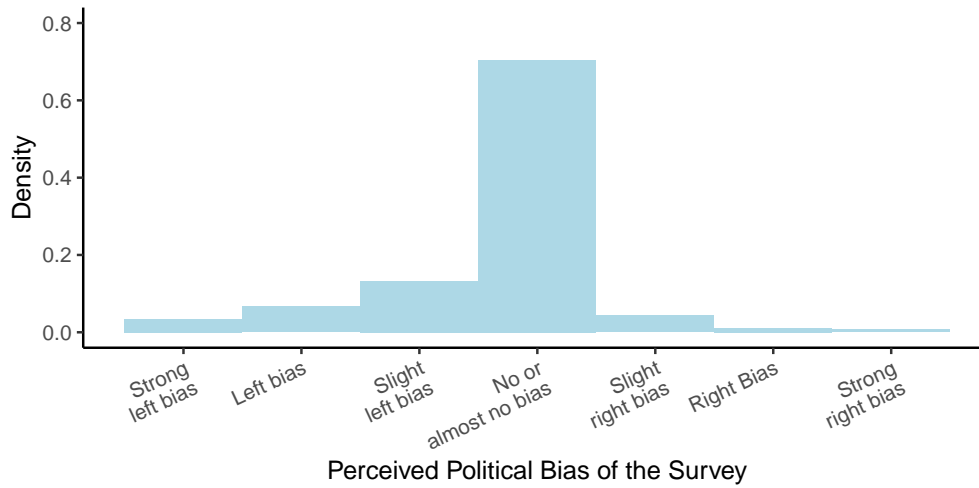
Figure 3.B.8. Spectators' Perception of the Survey's Political Bias
Note: Histogram of how subjects in the restricted sample respond to the question "Do you think this survey was biased toward a certain political stance?", asked at the end of the survey using a 7-point Likert scale from "strong left bias" to "strong right bias".

Appendix 3.C   Additional Figures

Figure 3.C.1. Screenshot of the Decision Screen for Spectator's Redistribution Decisions

Note: This decision screen corresponds to the Inherited Inequality & Merit condition. The decision screens for the other conditions had the same structure.
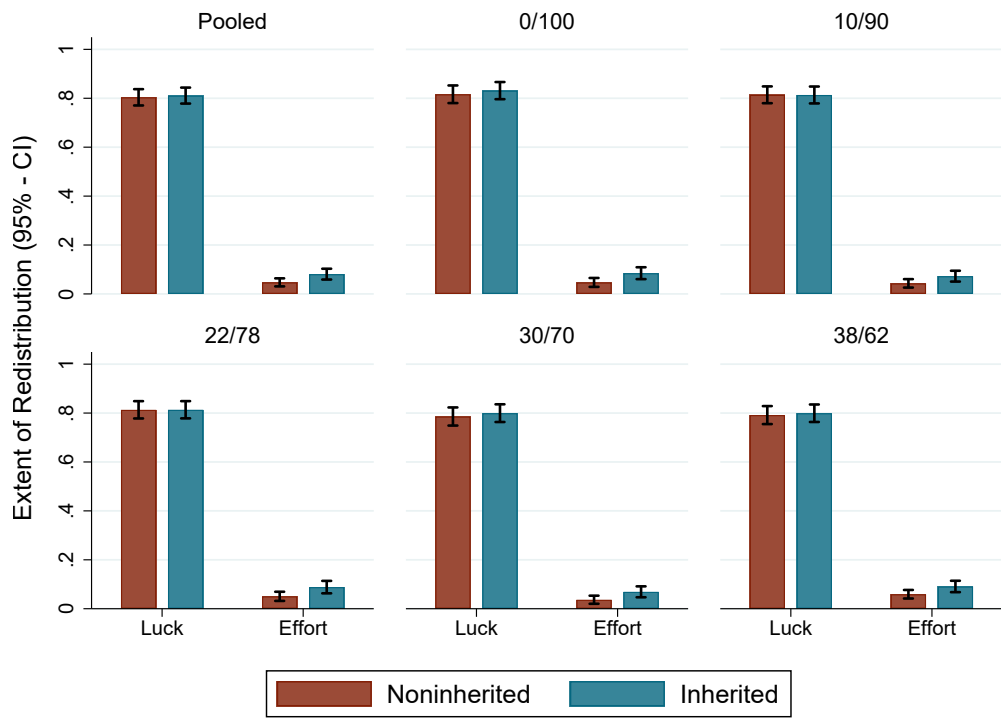
Figure 3.C.2. Average Extent of Redistribution by Treatment Condition and Initial Allocation

Note: This figure displays the average extent of redistribution $\bar{\theta}_{i,c}$ by treatment condition and initial allocation, together with $95-\%$ confidence intervals. The panel in the top left pools observations from all initial allocations, while each of the other panels refers to a different (hypothetical) initial allocation. Averages are taken over all decisions of all subjects in the restricted sample. Confidence intervals are based on standard errors clustered on the spectator level.

(a) Self vs. Other Worker   **(b)** Own Friend vs. Other Worker's Friend
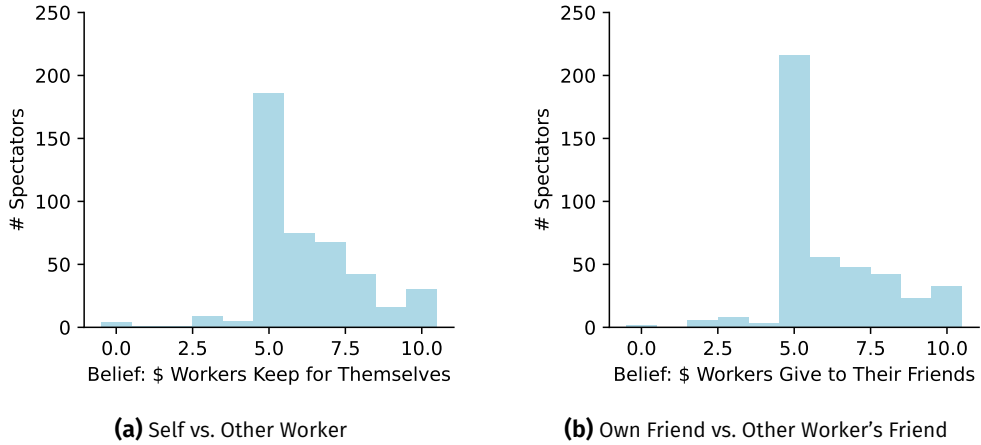
Figure 3.C.3. Spectators' Beliefs about Workers' Preferred Distributions
Note: Panel (a) displays a histogram of spectators' incentivized beliefs about the share of the $10 workers on average keep for themselves when they are asked how they would like to distribute $10 between themselves and the worker they are matched to in the first incentivized dictator decision. Panel (b) displays a histogram of spectators' incentivized beliefs about the share of the $10 workers on average give to their own friends when they are asked how they would like to distribute $10 between their own friend and the friend of the worker they are matched to in the second incentivized dictator decision.
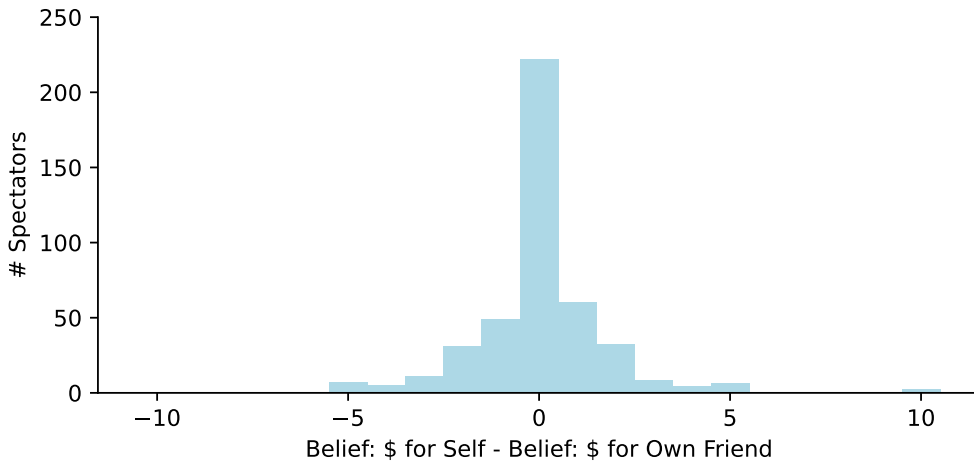
Figure 3.C.4. Differences in Spectators' Beliefs about Workers' Preferred Distributions
Note: This figure displays a histogram of the individual differences in spectators' beliefs about workers' preferred distributions in the dictator decisions for a) themselves vs. the worker they are matched to and b) their own friend vs. the friend of the worker they are matched to. For example, if a spectator indicated a belief that workers on average keep $8 for themselves when they are asked how they would like to distribute $10 between themselves and the worker they are matched to, and that workers on average give $7 to their own friend when they are asked how they would like to distribute $10 between their own friend and the friend of the worker they are matched to, this would yield a difference of $1.
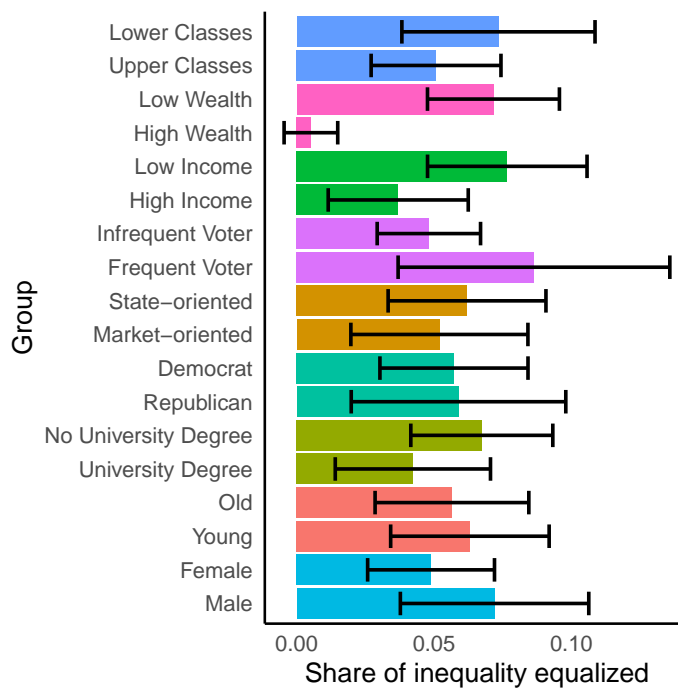
Figure 3.C.5. Average Equalization in Condition Noninherited Inequality & Effort by Demographic Group

Note: Shares of inequality equalized for a group are calculated by averaging over the average extent of redistribution in the Noninherited Inequality & Effort condition for all spectators in the main sample who belong to the group. 95% confidence intervals around the averages based on standard errors of the mean.
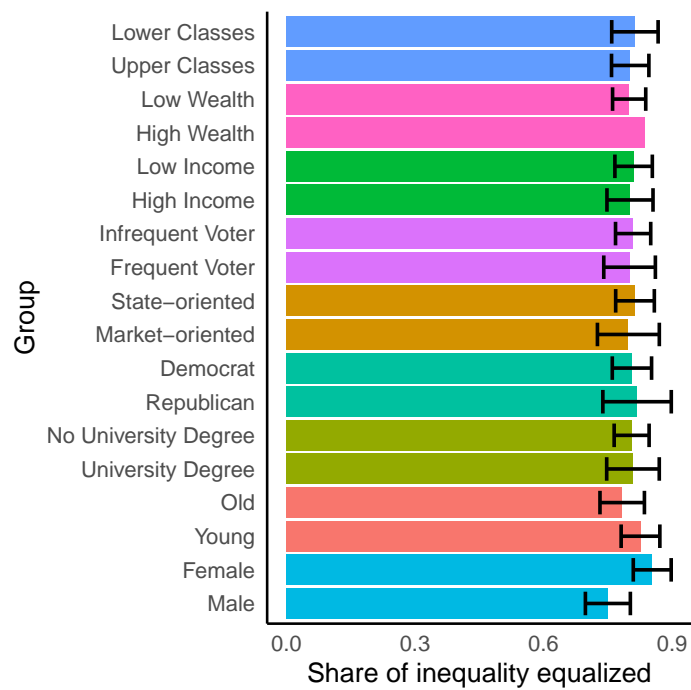
Figure 3.C.6. Average Equalization in Condition Noninherited Inequality & Luck by Demographic Group

Note: Shares of inequality equalized for a group are calculated by averaging over the average extent of redistribution in the Noninherited Inequality & Luck condition for all spectators in the main sample who belong to the group. 95% confidence intervals around the averages based on standard errors of the mean.
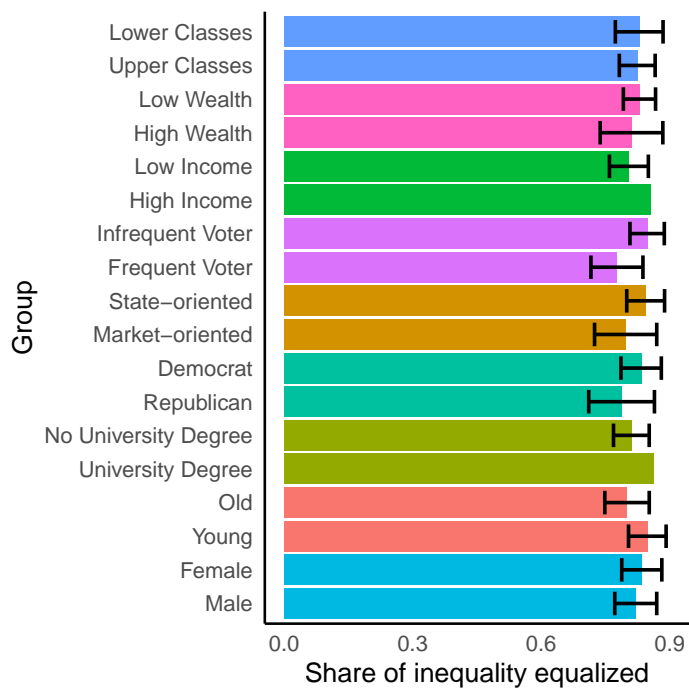
Figure 3.C.7. Average Equalization in Condition Inherited Inequality & Luck by Demographic Group

Note: Shares of inequality equalized for a group are calculated by averaging over the average extent of redistribution in the Inherited Inequality & Luck condition for all spectators in the main sample who belong to the group. 95% confidence intervals around the averages based on standard errors of the mean.
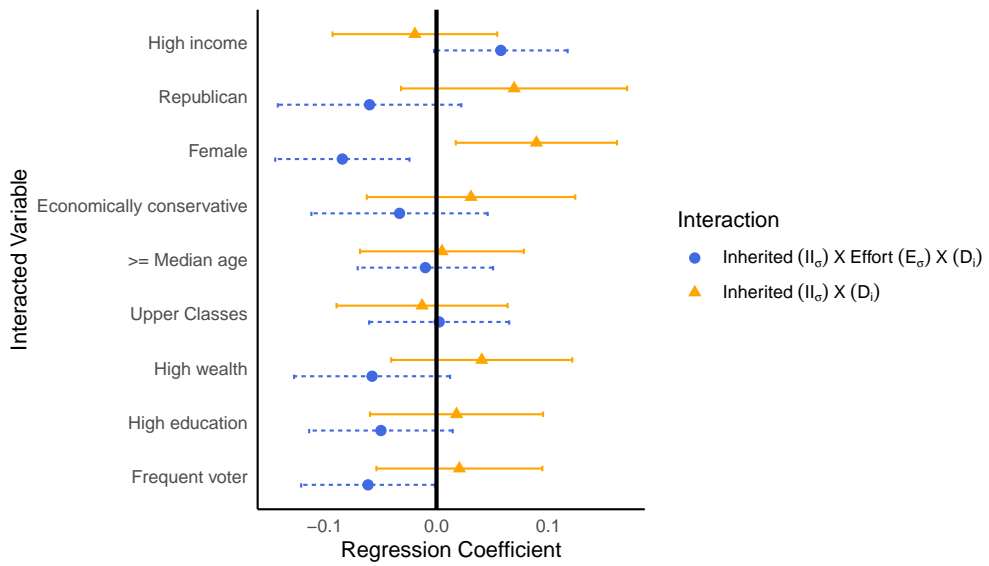
Figure 3.C.8. Heterogeneity in Treatment Effects between Demographic Groups
Note: This figure shows coefficients and 95% confidence intervals. The vertical axis shows demographic variables. These variables were interacted with two other terms in Equation 3.12. The blue points show the coefficient on the interaction term of each demographic variable (D_i) with the indicator for the INHERITED INEQUALITY conditions (II_σ). The orange points visualize the interaction of D_i with INHERITED INEQUALITY and an indicator for the EFFORT conditions (E_σ). Results are based on the main sample.
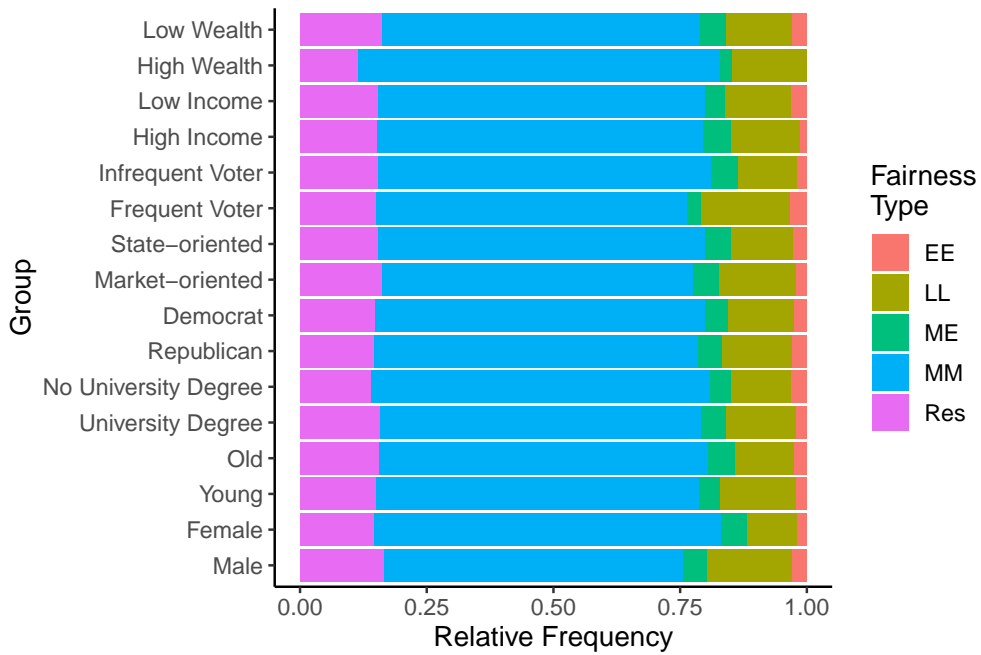
Figure 3.C.9. Distribution of Fairness Types by Demographic Group

Note: The vertical axis depicts demographic subgroups. Colors indicate five fairness types based on redistribution decisions under noninherited and inherited inequality. The horizontal axis shows the relative frequency with which these fairness types appear within the demographic subgroups. The fairness type ME stands for spectators who are classified as meritocrats under noninherited inequality and as egalitarians under inherited inequality. Likewise, EE, LL, and MM stand for egalitarian/egalitarian, liberterian/libertarian, and meritocrat/meritocrat, respectively. All spectators who do not belong to either of these types are summarized in the residual category "Res".

(a) Luck

(b) Effort
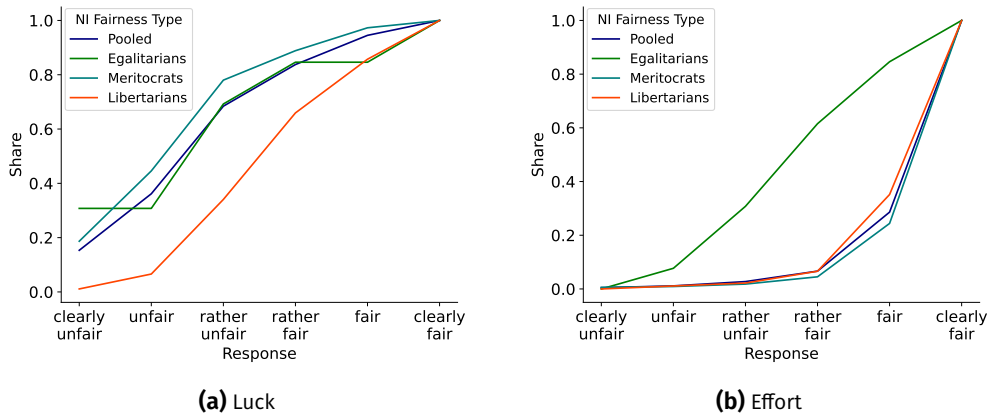
Figure 3.C.10. CDFs of the Responses to the Inequality Acceptance Survey Measures

Note: This figure shows cumulative redistribution functions of spectators' responses to the inequality acceptance survey questions. Panel (a) corresponds to the question "If one person receives more than another due to having better luck, I find that ..." and panel (b) corresponds to the question "If one person receives more than another due to exerting higher effort, I find that ...". Included are the responses of spectators in the restricted sample.

Appendix 3.D   Additional Tables

Table 3.D.1. Descriptives and Representativeness

	Spectator Sample (%)		US Population (%)
	Full/Main Sample	Restricted Sample	
Female	50.6	50.6	50.5
Age Groups			
18-19	1.5	1.6	3.4
20-24	9.9	8.5	8.3
25-29	11.7	9.5	8.6
30-34	9.3	8.8	8.9
35-39	10.8	9.5	8.7
40-44	8.9	9.0	8.3
45-49	6.9	7.2	7.7
50-54	8.4	8.8	8.1
55-59	10.8	11.3	8.2
60-64	9.1	10.6	8.4
65-69	7.1	8.5	7.1
70-74	3.2	3.7	6.0
75-79	2.2	2.5	3.8
80-84	0.4	0.5	2.4
85+	0.0	0.0	2.3
Education Groups			
No High School Diploma	0.4	0.2	10.6
High School Diploma Equivalent	30.4	30.0	45.6
Bachelor's or Associate's Degree	51.7	51.3	30.0
Master's Degree or Higher	17.5	18.5	13.8
Income Groups			
$< \$34,000$	26.7	27.5	25.0
$\$34,000 - \$68,000$	30.0	30.9	25.0
$\$68,000 - \$125,000$	30.0	28.4	25.0
$> \$125,000$	13.3	13.3	25.0
Race			
White	72.6	73.5	75.8
Black	12.6	12.9	13.6
Asian	7.2	6.3	6.1
Mixed	4.0	3.7	2.9
Other	3.6	3.5	1.6
Observations	543	437	

Note: This table reports descriptive statistics for our spectator sample and how they compare to the US general population. The survey company did not provide us with information on a spectator's age in two cases, gender in one case, and ethnicity in 13 cases. Shares in these groups are relative to the sample of spectators for which this information is available. Data for the US population are obtained from the 2021 American Community Survey, S0101 Age and Sex, via the United States Census Bureau (https://data.census.gov/table?tid=ACSST1Y2021.S0101, last accessed: January 9th, 2023; age and gender), the 2021 American Community Survey, S1501 Educational Attainment, via the United States Census Bureau (https://data.census.gov/table?tid=ACSST1Y2021.S1501, last accessed: January 9th, 2023; education groups), the United States Census Bureau QuickFacts table (https://www.census.gov/quickfacts/fact/table/US/PST045221, last accessed: January 16th, 2023; race), and https://dqydj.com/2020-household-income-percentile-calculator/, last accessed: January 9th, 2023; household income groups. Population data on educational attainment is based on citizens aged 25 years or older because for younger citizens the reported education groups did not match those we used in our survey. Likewise, we used the data on household income referenced above because they provided quartile household income group thresholds which we used in our survey.

Table 3.D.2. Treatment Effects on the Extent of Redistribution by Fairness Type

	Dependent Variable: Extent of Redistribution ($\theta_{i,\sigma}$, Share)			
	(1) Pooled	(2) Egalitarians	(3) Meritocrats	(4) Libertarians
Effort (E_σ)	-0.025	-0.025	-0.960***	-0.109***
	(0.036)	(0.038)	(0.006)	(0.018)
Inherited (II_σ)	-0.018	-0.017	-0.059***	0.268***
	(0.031)	(0.032)	(0.012)	(0.042)
Effort (E_σ) × Inherited (II_σ)	-0.144	-0.144	0.099***	-0.232***
	(0.103)	(0.108)	(0.015)	(0.044)
Meritocrat	-0.010			
	(0.015)			
Libertarian	-0.850***			
	(0.023)			
Nonclassified	-0.532***			
	(0.014)			
Effort (E_σ) × Meritocrat	-0.935***			
	(0.036)			
Effort (E_σ) × Libertarian	-0.083**			
	(0.040)			
Effort (E_σ) × Nonclassified	0.234***			
	(0.036)			
Inherited (II_σ) × Meritocrat	-0.042			
	(0.034)			
Inherited (II_σ) × Libertarian	0.286***			
	(0.052)			
Inherited (II_σ) × Nonclassified	-0.071**			
	(0.031)			
Effort (E_σ) × Inherited (II_σ) × Meritocrat	0.243**			
	(0.104)			
Effort (E_σ) × Inherited (II_σ) × Libertarian	-0.088			
	(0.112)			
Effort (E_σ) × Inherited (II_σ) × Nonclassified	0.296***			
	(0.103)			
Initial Inequality (Δ_σ)	0.031**	-0.052	-0.004	0.175***
	(0.014)	(0.101)	(0.012)	(0.045)
Constant	0.977***	1.001***	0.977***	0.084***
	(0.015)	(0.036)	(0.006)	(0.019)
Clusters	437	13	332	91
Observations	8399	249	6403	1731
R^2	0.817	0.106	0.864	0.228

Note: This table reports results from the same regression equations as Table 3.4 but does not omit coefficients. Results are based on observations in the restricted sample. Standard errors (in parentheses) are clustered on the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.D.3. Categories of Explanations That Spectators Give for Their Redistribution Decisions

Category Name	Argument Made by Spectator	Example
Effort	The final distribution should be based on the relative amount of tasks done (The spectator does not mention whether he means the tasks done by the workers or the tasks done by their friends).	The money should be based on the percentage of work each one did.
Effort Workers	The final distribution should be based on the relative amount of tasks done by the workers.	I made the payment based on the amount of work that each worker produced. It made no difference to me where the money ended up going, I just wanted to make sure that payments were made according to the amount of work produced.
Effort Friends	The final distribution should be based on the relative amount of tasks done by the friends.	I think it is fair to split the money evenly between the friends of the participants. They did not do any work.
Knew in Advance	All subjects knew the rules of the experiment in advance and agreed by participating. Changing rules after decisions have been made is unfair.	It was an easy task, and all participants were aware of what they were working towards - it would be unethical to change that agreement after the fact.
Zero Aversion	Every subject should receive something (of the bonus)/should at least receive a certain amount (e.g., $1).	i tried to be fair and also give 10% to those that completed 0
Round Numbers	Spectator has a preference for round numbers.	i prefer even numbers. even percentages.
Ability Luck	Some workers were more able to perform on the task than other workers due to lucky circumstances.	… I did want to move it back closer to an even split a little bit in case one worker had an advantage that made the task easier for them
Equality Preference	Money should always be distributed equally (no specific reasons stated).	No matter how much work I do, I think everyone has the right to about the same amount of money.
Luck Unfair	Outcomes that result from luck are unfair.	Just because your luck ran out on certain examples shouldn't be a cause to distribute that way
Luck Fair	Distributing based on luck is a fair procedure.	A random drawing is about as fair as it gets so I kept the same numbers. The workers just needed to cross their fingers that day.
No Right to Intervene	Spectator has no right to intervene in the affairs of others.	… If the Friend was lucky, why should I change things for them so that I make things fair for everyone within my own sense of justice or fairness. I can't play God. I believe it is contingent upon the person who has been lucky to give off his/her/they/their wealth to others who were less fortunate.
Exchange	The workers should earn what they worked for and the spectator expects the friends to share with their workers after the study.	… I think the people who did the work deserve to get the outcome they expected. Some of them probably selected a friend who would give them the money.

Table 3.D.3. Continued: Categories of Explanations That Spectators Give for Their Redistribution Decisions

Category Name	Argument Made by Spectator	Example
Type of Friend	The worker working for his friend means that the friend is a good person, and a good person should be rewarded.	… If Bill felt like knocking out a lot of tasks for his friend, who am I to take some of that and give it to James' friend when James did not think his friend was worth it?
Friend Not Entitled	The friends did not work for the money. Hence, they are not entitled to receive nay money.	These "friends" should feel lucky to be receiving anything at all. Neither friend is entitled to anything — especially more so for, that which the friend did *not* work for, ze'mself
Worker Entitled	The workers worked for the money. Hence, each worker is entitled to the amount he earned through his work.	The participants worked for and earned their share of the money. Even though the friends had no choice, the participants should receive (for their friend) a payment equivalent to how hard they worked
Friend Blameless	The friends did not work and are therefore not to blame for the distribution, in contrast to the workers. Hence, it is unfair that one friend gets less than another.	I had to make a decision between honoring the initiative of the workers or the making the receipts more equitable. Since the friends were "blameless" (and unconscious?) regarding the amount of labor involved, I elected to honor that side of the exercise with a 50-50 split
Team	Worker and friend are one team. What the team earns should stay with the team.	Even though friends did not work, he is a part of the team regardless and should be paid equally
NA	Comment without any explanation for the spectators' decisions.	Now is the time for the communist revolution! No more can these capitalist pigs turn us against one another! Throw off your chains, comrades, and let us create a world where no one goes hungry and we are truly free to pursue our passions!

Table 3.D.4. Association between Beliefs about Workers Preferences and Average Extent of Redistribution

	Dependent Variable: Average Extent of Redistribution ($\bar{\theta}_{i,c}$, Z-score)			
	Noninherited Inequality		Inherited Inequality	
	(1) Luck	(2) Effort	(3) Luck	(4) Effort
Guess Self/Other	-0.041 (0.047)	-0.104* (0.057)		
Guess Own/Other's Friend			-0.071 (0.049)	-0.131** (0.059)
Observations	437	437	437	437
R^2	0.002	0.011	0.005	0.017

Note: This table reports results from OLS regressions of spectators' average extent of redistribution ($\bar{\theta}_{i,c}$, standardized across spectators but within conditions), on their standardized beliefs about workers' preferred distributions. The coefficients are displayed in Figure 3.7. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.D.5. Association Between Beliefs and Average Extent of Redistribution

	Dependent Variable: Average Extent of Redistribution ($\bar{\theta}_{i,c}$, Z-score)					
	Noninherited Inequality			Inherited Inequality		
	(1) Egalitarians	(2) Meritocrats	(3) Libertarians	(4) Egalitarians	(5) Meritocrats	(6) Libertarians
Guess Self/Other	0.244* (0.134)	0.043 (0.045)	-0.089 (0.089)			
Guess Own/Other's Friend				-0.246 (0.291)	-0.115 (0.075)	0.036 (0.136)
Observations	13	332	91	13	332	91
R^2	0.060	0.002	0.008	0.060	0.013	0.001

Note: In analogy to Table 3.5, this table reports results from OLS regressions of spectators' average extent of redistribution in the two Luck conditions, standardized across spectators of a given (Noninherited Inequality) fairness type and within experimental conditions, on their beliefs about workers' preferred distributions, standardized across spectators of the same fairness type. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.D.6. Heterogeneity in Treatment Effects by Demographic Group (I)

	Dependent Variable: Extent of Redistribution ($\theta_{i,o}$, Share)			
	(1) Social Class D_i=1 if upper	(2) Wealth D_i=1 if high	(3) Income D_i=1 if high	(4) Education D_i=1 if high
Effort (E_o)	-0.742***	-0.728***	-0.735***	-0.767***
	(0.033)	(0.023)	(0.026)	(0.033)
Inherited (II_o)	0.019	0.031*	-0.004	0.055**
	(0.026)	(0.017)	(0.021)	(0.027)
Effort (E_o) × Inherited (II_o)	0.030	0.015	0.031	0.010
	(0.032)	(0.021)	(0.024)	(0.032)
D_i	-0.028	0.037	-0.009	-0.003
	(0.054)	(0.039)	(0.035)	(0.037)
Effort (E_o) × D_i	0.012	-0.104**	-0.030	0.029
	(0.062)	(0.041)	(0.039)	(0.041)
Inherited (II_o) × D_i	0.033	-0.058	0.058*	-0.050
	(0.046)	(0.036)	(0.030)	(0.033)
Effort (E_o) × Inherited (II_o) × D_i	-0.068	0.040	-0.019	0.018
	(0.060)	(0.041)	(0.038)	(0.039)
Initial Inequality (Δ_o)	0.062**	0.079***	0.079***	0.079***
	(0.025)	(0.019)	(0.019)	(0.019)
Constant	0.796***	0.777***	0.788***	0.786***
	(0.029)	(0.022)	(0.024)	(0.033)
Clusters	287	543	543	543
Observations	5435	10236	10236	10236
R^2	0.480	0.490	0.489	0.489

Note: This table shows reports OLS estimates corresponding to Equation 3.12 for the first set of sample splits. Sample sizes vary because for social class the middle category ("Middle Class") is disregarded. Results are based on the main sample. Standard errors (in parentheses) are clustered at the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.D.7. Heterogeneity in Treatment Effects by Demographic Group (II)

	Dependent Variable: Extent of Redistribution ($\theta_{i,\sigma}$, Share)				
	(1) Voting Freq. D_i=1 if high	(2) Econ. Ideology D_i=1 if conserv.	(3) Party Ident. D_i=1 if Rep.	(4) Age D_i=1 if old	(5) Sex D_i=1 if female
EFFORT (E_σ)	-0.762***	-0.750***	-0.747***	-0.735***	-0.680***
	(0.023)	(0.026)	(0.026)	(0.029)	(0.031)
INHERITED (II_σ)	0.039**	0.031	0.029	0.020	0.068***
	(0.019)	(0.021)	(0.022)	(0.023)	(0.022)
EFFORT (E_σ) × INHERITED (II_σ)	0.016	0.006	0.011	0.030	-0.029
	(0.023)	(0.025)	(0.027)	(0.027)	(0.028)
D_i	-0.010	-0.011	0.013	0.032	0.101***
	(0.037)	(0.042)	(0.046)	(0.034)	(0.034)
EFFORT (E_σ) × D_i	0.050	0.000	-0.014	-0.023	-0.125***
	(0.045)	(0.047)	(0.051)	(0.039)	(0.039)
INHERITED (II_σ) × D_i	-0.061**	-0.033	-0.060	0.000	-0.085***
	(0.031)	(0.040)	(0.042)	(0.031)	(0.030)
EFFORT (E_σ) × INHERITED (II_σ) × D_i	0.020	0.031	0.069	-0.015	0.090**
	(0.038)	(0.048)	(0.052)	(0.037)	(0.036)
Initial Inequality (Δ_σ)	0.079***	0.070***	0.072***	0.078***	0.080***
	(0.019)	(0.022)	(0.022)	(0.019)	(0.019)
Constant	0.787***	0.792***	0.783***	0.768***	0.729***
	(0.023)	(0.025)	(0.025)	(0.027)	(0.028)
Clusters	543	417	398	543	542
Observations	10236	7853	7485	10236	10216
R^2	0.489	0.502	0.488	0.489	0.492

Note: This table shows reports OLS estimates corresponding to Equation 3.12 for the second set of sample splits. Sample sizes vary because for economic ideology and party identification the middle categories ("Moderate" and "Neither Republican nor Democrat") are disregarded. Results are based on the main sample. Standard errors (in parentheses) are clustered at the spectator level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.D.8. Association between Experimental Measures and Policy Preferences

	Dependent Variable (Z-score):					
	Preferred Max. Marg. Rate		Support for			Rejection of
	(1) Income Tax	(2) Estate Tax	(3) Disability Insurance	(4) Unempl. Insurance	(5) Equal Opp. Programs	(6) Intergenerational Transmission
Redistribution (Luck)	0.136***	0.078*	0.081*	0.073	0.081*	0.197***
	(0.049)	(0.045)	(0.048)	(0.052)	(0.048)	(0.047)
Redistribution (Effort)	0.022	0.013	0.076	0.120***	0.059	0.111**
	(0.057)	(0.059)	(0.047)	(0.042)	(0.042)	(0.046)
Observations	437	437	437	437	437	437
R^2	0.019	0.006	0.013	0.020	0.010	0.052

Note: This table shows OLS estimates of (standardized) survey-based policy attitudes on (standardized) factor variables based on spectators' average extent of redistribution in the four treatment conditions. The coefficients are plotted in Figure 3.9. Results are based on the main sample. Robust standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

References

Alesina, Alberto, and George-Marios Angeletos. 2005. "Fairness and Redistribution." en. *American Economic Review* 95 (4): 960–80. DOI: 10.1257/0002828054825655. (Visited on 03/19/2023). [142]

Alesina, Alberto, and Paola Giuliano. 2011. "Preferences for Redistribution." en. In *Handbook of Social Economics*. Vol. 1, Elsevier, 93–131. DOI: 10.1016/B978-0-444-53187-2.00004-8. (Visited on 03/19/2023). [142]

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso. 2018. "Intergenerational Mobility and Preferences for Redistribution." en. *American Economic Review* 108 (2): 521–54. DOI: 10.1257/aer.20162015. (Visited on 03/19/2023). [142, 173]

Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden. 2020. "Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" *Journal of Political Economy* 128 (5): 1753–88. DOI: 10.1086/705551. (Visited on 04/09/2023). Publisher: The University of Chicago Press. [139, 143, 145, 147, 151]

Andre, Peter. 2022. "Shallow Meritocracy." en. [142, 173]

Benabou, Roland. 2000. "Meritocracy, Redistribution, and the Size of the Pie." *Meritocracy and economic inequality*, 317–39. Publisher: Princeton University Press Princeton. [142]

Bernhard, Helen, Urs Fischbacher, and Ernst Fehr. 2006. "Parochial Altruism in Humans." en. *Nature* 442 (7105): 912–15. DOI: 10.1038/nature04981. (Visited on 04/09/2023). Number: 7105 Publisher: Nature Publishing Group. [172]

Björklund, Anders, Jesper Roine, and Daniel Waldenström. 2012. "Intergenerational Top Income Mobility in Sweden: Capitalist Dynasties in the Land of Equal Opportunity?" en. *Journal of Public Economics* 96 (5-6): 474–84. DOI: 10.1016/j.jpubeco.2012.02.003. (Visited on 03/19/2023). [142]

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch. 2014. "hroot: Hamburg Registration and Organization Online Tool." en. *European Economic Review* 71 (10): 117–20. DOI: 10.1016/j.euroecorev.2014.07.003. (Visited on 04/09/2023). [148]

Bowles, Samuel, and Herbert Gintis. 2002. "The Inheritance of Inequality." en. *Journal of Economic Perspectives* 16 (3): 3–30. DOI: 10.1257/089533002760278686. (Visited on 04/09/2023). [142]

Cappelen, Alexander W, James Konow, Erik Ø Sørensen, and Bertil Tungodden. 2013. "Just Luck: An Experimental Study of Risk-Taking and Fairness." en. *American Economic Review* 103 (4): 1398–413. DOI: 10.1257/aer.103.4.1398. (Visited on 03/19/2023). [139, 142, 143, 145]

Cappelen, Alexander W, Johanna Mollerstrom, Bjørn-Atle Reme, and Bertil Tungodden. 2022. "A Meritocratic Origin of Egalitarian Behaviour." *Economic Journal* 132 (646): 2101–17. DOI: 10.1093/ej/ueac008. (Visited on 04/09/2023). [142, 151, 173]

Cappelen, Alexander W., Benjamin Enke, and Bertil Tungodden. 2022. "Moral Universalism: Global Evidence." en. SSRN Scholarly Paper. Rochester, NY. DOI: 10.2139/ssrn.4291744. (Visited on 04/09/2023). [172, 173]

Cappelen, Alexander W., Ranveig Falch, and Bertil Tungodden. 2020. "Fair and Unfair Income Inequality." en. In *Handbook of Labor, Human Resources and Population Economics*. Edited by Klaus F. Zimmermann. Cham: Springer International Publishing, 1–25. DOI: 10.1007/978-3-319-57365-6_130-1. (Visited on 03/19/2023). [139, 141]

Cappelen, Alexander W., Astri Drange Hole, Erik Ø Sørensen, and Bertil Tungodden. 2007. "The Pluralism of Fairness Ideals: An Experimental Approach." en. *American Economic Review* 97 (3): 818–27. DOI: 10.1257/aer.97.3.818. (Visited on 04/09/2023). [142]

Cappelen, Alexander W., Erik Ø. Sørensen, and Bertil Tungodden. 2010. "Responsibility for What? Fairness and Individual Responsibility." en. *European Economic Review* 54 (3): 429–41. DOI: 10.1016/j.euroecorev.2009.08.005. (Visited on 03/19/2023). [142]

Cappelen, Alexander W., and Bertil Tungodden. 2017. "Fairness and the Proportionality Principle." en. *Social Choice and Welfare* 49 (3-4): 709–19. DOI: 10.1007/s00355-016-1016-6. (Visited on 03/19/2023). [142, 151, 173]

Cassar, Lea, and Arnd H. Klein. 2019. "A Matter of Perspective: How Failure Shapes Distributive Preferences." *Management Science* 65 (11): 5050–64. DOI: 10.1287/mnsc.2018.3185. (Visited on 04/09/2023). Publisher: INFORMS. [169, 173]

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. "Otree—an Open-Source Platform for Laboratory, Online, and Field Experiments." en. *Journal of Behavioral and Experimental Finance* 9 (3): 88–97. DOI: 10.1016/j.jbef.2015.12.001. (Visited on 03/19/2023). [148]

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." en. *American Economic Review* 106 (4): 855–902. DOI: 10.1257/aer.20150572. (Visited on 03/19/2023). [142]

Cohen, Rachelle, Amnon Maltz, and Yuval Ofek-Shanny. 2022. *Social Preferences Over Taste-Based Inequality*. [142, 143, 174]

Deffains, Bruno, Romain Espinosa, and Christian Thöni. 2016. "Political Self-Serving Bias and Redistribution." en. *Journal of Public Economics* 134 (2): 67–74. DOI: 10.1016/j.jpubeco.2016.01.002. (Visited on 04/09/2023). [169, 173]

Eisenkopf, Gerald, Urs Fischbacher, and Franziska Föllmi-Heusi. 2013. "Unequal Opportunities and Distributive Justice." en. *Journal of Economic Behavior & Organization* 93 (9): 51–61. DOI: 10.1016/j.jebo.2013.07.011. (Visited on 04/09/2023). [173]

Gächter, Simon, Chris Starmer, and Fabio Tufano. 2015. "Measuring the Closeness of Relationships: A Comprehensive Evaluation of the 'Inclusion of the Other in the Self' Scale." en. *PLOS ONE* 10 (6): e0129478. DOI: 10.1371/journal.pone.0129478. (Visited on 04/09/2023). Publisher: Public Library of Science. [146]

Gärtner, Manja, Johanna Mollerstrom, and David Seim. 2017. "Individual Risk Preferences and the Demand for Redistribution." en. *Journal of Public Economics* 153 (9): 49–55. DOI: 10.1016/j.jpubeco.2017.06.009. (Visited on 03/19/2023). [141]

Gill, David, and Victoria Prowse. 2012. "A Structural Analysis of Disappointment Aversion in a Real Effort Competition." en. *American Economic Review* 102 (1): 469–503. DOI: 10.1257/aer.102.1.469. (Visited on 04/09/2023). [146]

Konow, James. 2000. "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions." en. *American Economic Review* 90 (4): 1072–92. DOI: 10.1257/aer.90.4.1072. (Visited on 03/19/2023). [139, 145, 169, 173]

Konow, James. 2001. "Fair and Square: The Four Sides of Distributive Justice." en. *Journal of Economic Behavior & Organization* 46 (2): 137–64. DOI: 10.1016/S0167-2681(01)00194-9. (Visited on 04/09/2023). [173]

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk. 2020. "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy* 128 (2): 434–67. DOI: 10.1086/704386. (Visited on 04/14/2023). Publisher: The University of Chicago Press. [142]

Kube, Sebastian, and Christian Traxler. 2011. "The Interaction of Legal and Social Norm Enforcement." en. *Journal of Public Economic Theory* 13 (5): 639–60. DOI: 10.1111/j.1467-9779.2011.01515.x. (Visited on 03/19/2023). [147]

Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." en. *American Economic Review* 105 (4): 1478–508. DOI: 10.1257/aer.20130360. (Visited on 03/19/2023). [142]

Mollerstrom, Johanna, Bjørn-Atle Reme, and Erik Ø. Sørensen. 2015. "Luck, Choice and Responsibility — An Experimental Study of Fairness Views." en. *Journal of Public Economics* 131 (11): 33–40. DOI: 10.1016/j.jpubeco.2015.08.010. (Visited on 03/19/2023). [142]

Obama, Barack. 2006. *The Audacity of Hope: Thoughts on Reclaiming the American Dream.* Englisch. First Edition. Crown. [138]

Palan, Stefan, and Christian Schitter. 2018. "Prolific.ac—a Subject Pool for Online Experiments." en. *Journal of Behavioral and Experimental Finance* 17 (3): 22–27. DOI: 10.1016/j.jbef.2017.12.004. (Visited on 03/19/2023). [149]

Peer, Eyal, David Rothschild, Andrew Gordon, Zak Evernden, and Ekaterina Damer. 2021. "Data Quality of Platforms and Panels for Online Behavioral Research." en. *Behavior Research Methods* 54 (4): 1643–62. DOI: 10.3758/s13428-021-01694-3. (Visited on 03/19/2023). [149]

Piketty, Thomas, and Emmanuel Saez. 2013. "A Theory of Optimal Inheritance Taxation." en. *Econometrica* 81 (5): 1851–86. DOI: 10.3982/ECTA10712. (Visited on 04/09/2023). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10712. [142]

Posner, Richard A. 2014. *Economic Analysis of Law.* en. Aspen Publishing. Google-Books-ID: o77fDgAAQBAJ. [138]

Rodriguez-Lara, Ismael, and Luis Moreno-Garrido. 2012. "Self-Interest and Fairness: Self-Serving Choices of Justice Principles." en. *Experimental Economics* 15 (1): 158–75. DOI: 10.1007/s10683-011-9295-3. (Visited on 04/09/2023). [169, 173]

Roth, Christopher, and Johannes Wohlfart. 2018. "Experienced Inequality and Preferences for Redistribution." en. *Journal of Public Economics* 167 (11): 251–62. DOI: 10.1016/j.jpubeco.2018.09.012. (Visited on 04/09/2023). [141]

Schaube, Sebastian, and Louis Strang. 2022. "(Not) Everyone Can Be a Winner–The Role of Payoff Interdependence for Redistribution." en. [142, 147, 151]

Schwaiger, Rene, Jürgen Huber, Michael Kirchler, Daniel Kleinlercher, and Utz Weitzel. 2022. "Unequal Opportunities, Social Groups, and Redistribution: Evidence from Germany." en. *Journal of Behavioral and Experimental Economics* 100 (10): 101911. DOI: 10.1016/j.socec.2022.101911. (Visited on 04/09/2023). [173]

Stantcheva, Stefanie. 2021. "Understanding Tax Policy: How do People Reason?" *Quarterly Journal of Economics* 136 (4): 2309–69. DOI: 10.1093/qje/qjab033. (Visited on 04/09/2023). [139, 142, 172]

Stantcheva, Stefanie. 2023. "How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible." *Annual Review of Economics* 15 (1): null. DOI: 10.1146/annurev-economics-091622-010157. (Visited on 03/30/2023). _eprint: https://doi.org/10.1146/annurev-economics-091622-010157. [150]

Valero, Vanessa. 2022. "Redistribution and Beliefs about the Source of Income Inequality." en. *Experimental Economics* 25 (3): 876–901. DOI: 10.1007/s10683-021-09733-8. (Visited on 04/09/2023). [169, 173]