

The Cost of Exclusion: Political Discrimination and Trust Deficits in the European Union

Carlos A. Toruño Paniagua¹ and Santiago Pardo González²

^{1,2}*The World Justice Project*

June 2025

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer tempus sagittis lobortis. Ut ac ligula a dolor faucibus rutrum. Donec accumsan sem nibh, quis ornare lorem volutpat a. Sed nec efficitur diam. Vivamus molestie turpis lorem, vitae semper lacus convallis sodales. Cras tellus leo, ultrices sit amet fringilla et, aliquam commodo tellus. Maecenas varius tincidunt libero, in aliquam orci iaculis eu. Vestibulum consequat ultrices eros id feugiat. Aliquam scelerisque porta lorem, id venenatis justo. Curabitur feugiat ante nec tellus pulvinar elementum. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Sed pellentesque ac dui id accumsan. Nam tellus libero, venenatis nec libero quis, varius rutrum velit. Quisque eu finibus diam. Nunc sit amet massa purus.

1 Introduction

2 Literature Review

3 Data

This study uses microdata from general population surveys collected in all 27 European Union Member States as part of the *EUROVOICES* project, conducted by The World Justice Project.¹ The dataset provides statistically representative coverage across 110 subnational regions throughout the European Union, offering comprehensive geographic representation of the EU population.

¹The World Justice Project *EUROVOICES* was a data collection effort that aimed to evaluate, analyze, and map out national and regional variations in how people in the EU perceive and experience democratic governance, fundamental rights, justice, safety, corruption, transparency, regulatory enforcement, and the business climate (See WJP, 2024).

The World Justice Project produced subnational information following the framework of territorial divisions established by the Nomenclature of Territorial Units for Statistics (NUTS) system. In each covered country, one level of this system of nested territorial divisions was selected (NUTS-1 or NUTS-2), resulting in a total of 110 subnational regions. It should be noted that the level of statistical representation varies between countries.

The surveys were administered to respondents in 110 regions of the 27 EU Member States. Data collection employed mixed methodologies, with face-to-face interviews conducted in ten countries and online polling conducted in 17 countries. Survey respondents in each country-region were selected using probability sampling methods designed to ensure representativeness across key sociodemographic criteria, including age, sex, income level, and degree of urbanization (categorized as urban and rural areas). Detailed information on sample sizes for each country, survey methodology employed, data collection periods, and representation levels can be found in Appendix A.

However, due to missing values across the variables of interest, the effective analytical sample for this study is smaller than the total sample published by The World Justice Project. Table 1 presents descriptive statistics for the key demographic and political variables among survey respondents included in the study *full* sample.

The *EUROVOICES* survey includes a comprehensive discrimination module that measures 11 different grounds for discrimination: sex, age, disability, ethnicity, migration background, socioeconomic status, religion, and political opinion. Participants indicated whether they experienced discrimination or harassment (D/H) based on any of these characteristics during the 12 months before the survey, and identified where such incidents took place. Since each discrimination ground was assessed independently, individual respondents could report multiple types of discriminatory experiences within the year prior to survey administration.

The survey also includes a comprehensive Trust module asking respondents to rate their trust levels in various public institutions as “a lot”, “some”, “a little”, or “no trust”. These institutions include local and national government officials, police officers, prosecutors, public defense attorneys, judges, magistrates, political parties, and Parliament members. We recoded these responses into binary indicators, where “a lot” and “some trust” equal 1, and “a little” and “no trust” equal 0. To construct a comprehensive measure of political trust, we created a Trust in Political Institutions Index through Logistic Principal Component Analysis (Landgraf & Lee, 2020), which aggregates individual trust responses across all measured institutions.

Finally, as displayed in Table 1, the survey not only collects important demographic data but it also asks respondents key political traits such as their

Table 1: Sample Demographics and Political Preferences

	Unique	Mean	SD	Min	Max
Age	79	49.7	16.6	18.0	99.0
Political Ideology	11	5.4	2.3	0.0	10.0
		N	%		
Sex	Female	22,659	51.2		
	Male	21,576	48.8		
Area of Residence	Urban	31,993	72.3		
	Rural	12,242	27.7		
Employment	Employed	24,821	56.1		
	Unemployed/Inactive	19,414	43.9		
Citizenship Status	Citizen	43,224	97.7		
	Foreigner	1,011	2.3		
Marital Status	Married	26,838	60.7		
	Not married	17,397	39.3		
Ethnic Group	Ethnic Majority	26,629	60.2		
	Ethnic Minority	17,606	39.8		
Financial Situation [†]	Constrained	10,103	22.8		
	Unconstrained	34,132	77.2		
Education	Higher Education	14,278	32.3		
	No Higher Education	29,957	67.7		
Interest in Politics [‡]	Interested	21,887	49.5		
	Uninterested	22,348	50.5		
Political Alignment [§]	Incumbent Political Party	7,742	17.5		
	Non-Incumbent Political Party	36,493	82.5		
Political D/H	D/H Experience	7,002	15.8		
	No D/H Experience	37,233	84.2		

[†] Classified as “*constrained*” if respondents mentioned that their income is not enough for basic necessities or clothing and “*unconstrained*” otherwise.

[‡] Individuals had to answer if they were very interested, interested, a little interested, or not at all interested in politics.

[§] Political alignment was captured using the respondents’ vote intention, where they had to pick a political party to vote for if the general elections were happening that weekend.
Source: Eurovoices General Population Poll 2024

interest in politics, vote intention, and it also asks people to rate their political ideology from 0 to 10, where 0 is associated with a far left ideology and 10 is associated with a far right ideology.

4 Discrimination Experiences in the European Union

Data from the World Justice Project *EUROVOICES* household survey reveals that discrimination represents a significant challenge throughout the European Union. In most EU regions, over 25% of respondents reported experiencing some form of discrimination within the previous year. Table 2 displays the prevalence of self-reported discrimination across EU Member States, categorized by the grounds for discrimination.

Political discrimination represents the most frequently reported form of discrimination in 13 of the 27 members, surpassing discrimination based on sex, ethnicity, or migration status. The highest rates are observed in Hungary (34.5%), Czechia (28.7%), Slovakia (26.5%), Austria (25.3%), and Germany (24.4%), while Portugal (2.6%) and Bulgaria (1.5%) show considerably lower prevalence. Across the EU, an average of 14.7% of respondents have encountered political discrimination.

One reason why political grounds for discrimination display such high incidence rates across most countries in our sample is that, unlike other grounds such as sex, ethnicity, or religion, political discrimination affects all demographic groups in our sample. While political opinion may not be the predominant form of discrimination experienced by specific minority or marginalized groups, they become the most common ground when considering the entire population. This pattern reflects the universal nature of political discrimination—it can target individuals regardless of their demographic characteristics.

These findings indicate that political discrimination constitutes a prominent and pervasive issue across numerous European contexts. Notably, this form of discrimination has become more prevalent than traditionally recognized categories of discrimination in a substantial number of countries, suggesting an evolution in the nature of exclusionary experiences faced by individuals.

5 Methodology

The main objective of this study is to examine whether experiencing discrimination or harassment (D/H experiences)² due to political opinions has a causal effect on trust in political institutions. However, a fundamental challenge in

²Throughout the remainder of this document, political discrimination or harassment will be referred to as D/H experiences or events.

Table 2: Experiences of Discrimination in the EU

Country	Political Opinion	Sex	Gender	Ethnicity	Migration Status	Social Status	Religion
Austria	25.3	13.0	9.5	13.1	12.6	17.8	11.6
Belgium	13.2	15.0	8.5	13.0	9.2	18.1	10.7
Bulgaria	1.5	1.8	0.7	2.1	0.6	3.1	1.1
Croatia	11.8	8.2	2.6	5.0	3.1	11.2	8.3
Cyprus	11.0	13.3	6.1	6.4	5.5	12.3	6.7
Czechia	28.7	13.0	10.0	13.9	16.6	18.5	5.4
Denmark	11.5	13.0	7.5	10.8	9.9	14.7	10.1
Estonia	15.5	9.5	3.4	6.7	2.2	10.1	2.8
Finland	13.7	11.4	4.6	4.4	2.8	15.9	5.0
France	10.4	11.6	4.6	7.7	4.4	12.2	6.7
Germany	24.4	11.9	8.0	11.1	10.0	16.0	9.5
Greece	5.2	3.6	0.8	1.5	1.4	2.8	1.0
Hungary	34.5	15.9	16.5	22.4	16.5	27.8	14.3
Ireland	12.7	15.8	7.1	9.1	8.7	14.8	9.6
Italy	13.7	13.7	7.8	8.1	6.3	13.9	8.2
Latvia	6.0	2.5	0.5	7.5	1.9	5.5	1.9
Lithuania	5.3	2.2	0.1	2.9	0.5	4.9	1.0
Malta	13.0	6.8	1.6	4.4	3.4	4.4	4.0
Netherlands	15.2	14.0	9.4	12.5	11.2	15.2	10.6
Poland	6.7	3.0	1.0	1.3	1.0	3.4	2.4
Portugal	2.6	3.5	2.6	4.1	2.8	4.1	3.5
Romania	4.7	3.6	1.8	3.7	1.7	5.3	2.7
Slovakia	26.5	14.1	10.7	12.3	10.1	18.5	11.1
Slovenia	18.1	10.1	5.3	9.2	7.1	17.0	9.0
Spain	19.8	14.6	7.5	9.4	8.4	15.2	8.3
Sweden	12.1	15.3	5.7	10.8	7.2	13.4	7.3

Note: Table displays the percentage of respondents in each country that answered to have had experienced discrimination or harassment for each of the grounds presented to them by the survey.

Source: Eurovoices General Population Poll 2024

causal inference is that assessing the causal effect of these experiences would require observing both potential outcomes in the same individual—those who experienced a D/H event and those who did not. Since we never observe both potential outcomes for any individual, we must infer the effects of political discrimination and harassment by comparing average trust levels between those who experienced such events and those who did not.

The problem is that people who experience political discrimination or harassment might be systematically different from those who do not. For example, individuals who experience such events might be more interested in and informed about politics, or display higher levels of civic participation. If these systematic differences correlate with institutional trust levels, then a naïve comparison between the two groups would be misleading. This is known as the selection problem (Angrist & Pischke, 2009).

Researchers typically turn to randomized controlled experiments—the gold standard for causal inference—to address such questions. However, the nature of our research question makes it both practically difficult and ethically problematic to randomly expose individuals to discrimination or harassment based on their political views. As a result, we adopt quasi-experimental designs to examine the causal effects of political discrimination and harassment on institutional trust [citations needed?].

For this study, we implement the matching methodology suggested by Ho et al. (2007) to preprocess our observational data. Our goal is to balance the distribution of covariates between our treatment group (people who experienced D/H events) and control group (people who have not), thereby replicating the randomization achieved in an experimental study (Stuart, 2010).

While regression methods can also address confounding from measured covariates, using regression on matched samples reduces the dependence of our treatment effect estimates on correct model specification. This approach provides more robust causal inferences by making the groups more comparable before applying statistical models.

Importantly, we use matching as a preprocessing method to improve balance between groups, rather than as an imputation method for estimating missing outcomes, as described by Abadie and Imbens (2006, 2016).

We performed nearest neighbor matching³ based on the propensity score, defined as the probability of experiencing a D/H event conditional on a set of observed covariates (Rubin, 1973; Rosenbaum and Rubin, 1983). We estimate this probability using a logistic regression. Each unit from the treatment group is matched with the control group unit that has the closest propensity score

³The matching estimation was performed using the MatchIT R package (Ho et al., 2011).

(1:1 matching without replacement)⁴. We also establish a region of common support⁵ and discard units that fall outside this region.⁶

The propensity score is estimated using two sets of covariates: demographic variables (geographic region,⁷ sex, age, area of residence, financial situation, education level, employment status, marital status, citizenship status, and ethnic group) and political variables (interest level in politics, political ideology, and political alignment). All covariates and their possible values are listed in Table 1.

Once we have a fully matched study sample, we estimate the effect of having experienced a D/H event on the levels of trust in political institutions through the following specification:

$$\log(Y_{ic}) = \alpha + \tau X_{ic} + \mathbf{Z}'_{ic}\beta + \gamma_c + \varepsilon_{ic} \quad (1)$$

Where:

- Y_{ic} : political trust for individual i in region c
- X_{ic} : D/H event (1 = treated, 0 = control)
- τ : treatment effect (ATT)
- \mathbf{Z}_{ic} : vector of covariates
- β : coefficient vector for covariates
- γ_c : region fixed effects
- ε_{ic} : error term

This specification is estimated using fixed effects OLS regression⁸. The value of τ can be interpreted as the Average Treatment Effect on the Treated (ATT)⁹

⁴Matching with replacement means that each control unit can be reused to be matched with any number of treated units if that individual is the closest neighbor to multiple treatment units. Matching without replacement means that each control unit can only be matched with one single treated unit.

⁵The common support region refers to the range of covariate values (or propensity scores) where both treated and control units have sufficient representation.

⁶All treatment units fell within the common support region, resulting only in exclusion of individuals from the control group, i.e., people who did not experience political discrimination or harassment.

⁷Equivalent to one of the 110 NUTS regions in the study.

⁸The estimation of the fixed effects was performed using the fixest R package (Bergé, 2018).

⁹The Average Treatment Effect (ATE) and the Average Treatment Effect on the Treated

provided that the conditional ignorability (unconfoundedness) assumption holds (Abadie and Cattaneo, 2018; Greifer and Stuart, 2010).

The unconfoundedness assumption requires that treatment assignment be independent of potential outcomes conditional on the set of observed covariates included. However, even when matching achieves balance in observed covariates, unobserved confounders remain a threat to causal identification. Unobserved variables that relate to both experiencing D/H events and trust in political institutions would violate the unconfoundedness assumption and bias our treatment effect estimates. Since conditional ignorability cannot be directly tested, the literature suggests performing sensitivity analyses to assess the plausibility of this assumption and determine how sensitive our estimated effects are to violations of unconfoundedness (Stuart, 2010).

We use the framework proposed by Cinelli and Hazlett (2020) to assess how fragile are our main results to the possibility of unobserved confounding. This framework allows us to estimate the minimum strength of an unobserved confounder that would be required to nullify our estimated treatment effect. We also assess the robustness of our results to violations of the unconfoundedness assumption by estimating the treatment effect under different levels of unobserved confounding.

We use two measures to test the sensitivity of our results: the robustness value, indicating the minimum confounding strength needed to alter our conclusions, and the partial R² of the treatment, showing how strongly confounders must be associated with the treatment to completely override our results (Cinelli and Hazlett, 2020).

Finally, we investigate the mechanisms through which political discrimination and harassment affect trust in political institutions. We explore two potential pathways: political polarization and political alignment.

Regarding political polarization, we hypothesize that individuals who experience D/H events may become more politically polarized, which in turn could lead to lower levels of trust in political institutions. To test this hypothesis, we include a measure of political polarization in our regression model and assess the transmission mechanism using the mediation analysis proposed by Imai et al. (2011) and Imai and Yamamoto (2013).¹⁰

The mediation analysis follows a two-step process. First, we examine how D/H experiences affect political polarization (see Equation 2). Second, we analyze

(ATT) are two foundational metrics in causal inference. ATE represents the average effect of a treatment across the entire population, whereas ATT focuses only on the average effect among individuals who actually received the treatment.

¹⁰We use the mediation R package developed by Tingley et al. (2014) for the practical implementation of this framework.

how both D/H experiences and political polarization jointly influence trust in political institutions (see Equation 3). We measured political polarization at the individual level using a measure of political conformity/dissonance based on the political ideologies of the respondents. This measure is calculated as the square of the difference between the respondent's political ideology and the average political ideology of the country.¹¹

$$M_{ic} = \alpha + \phi X_{ic} + \mathbf{Z}'_{ic}\beta + \gamma_c + \epsilon_{ic} \quad (2)$$

$$Y_{ic} = \alpha + \delta X_{ic} + \rho M_{ic} + \mathbf{Z}'_{ic}\eta + \lambda_c + \mu_{ic} \quad (3)$$

Where:

- Y_{ic} : political trust for individual i in region c
- X_{ic} : D/H event (1 = treated, 0 = control)
- M_{ic} : political dissonance for individual i in region c
- ϕ , δ , and ρ : coefficients of interests
- \mathbf{Z}_{ic} : vector of covariates
- β and η : coefficient vectors for covariates
- γ_c and λ_c : region fixed effects
- ϵ_{ic} and μ_{ic} : error terms

After fitting these models, we use bootstrapping to estimate causal effects under counterfactual scenarios. In each simulation, we compare predicted trust levels between two scenarios: experiencing D/H events versus not experiencing them, while allowing political dissonance to adjust to its expected level under each condition. Averaging across simulations yields the average causal mediation effect (ACME), which captures the indirect effect through polarization, and the average direct effect (ADE), which captures the remaining direct impact.

Regarding political alignment, we hypothesize that there is heterogeneous effects between individuals that are politically aligned with the incumbent party and those who are not. To test this hypothesis, we include a measure of political

¹¹It is important to note that we use the full study sample to estimate the average political ideology of the country, rather than the matched sample. We believe that using the full representative country sample captures the true distance between the respondent's political ideology and the average political ideology of the country.

alignment with the incumbent party in our regression model (see Equation 4) and assess its interaction with the D/H event variable.

$$\log(Y_{ic}) = \alpha + \tau X_{ic} + \mathbf{Z}'_{ic}\beta + \gamma_c + \varepsilon_{ic} \quad (4)$$

Where:

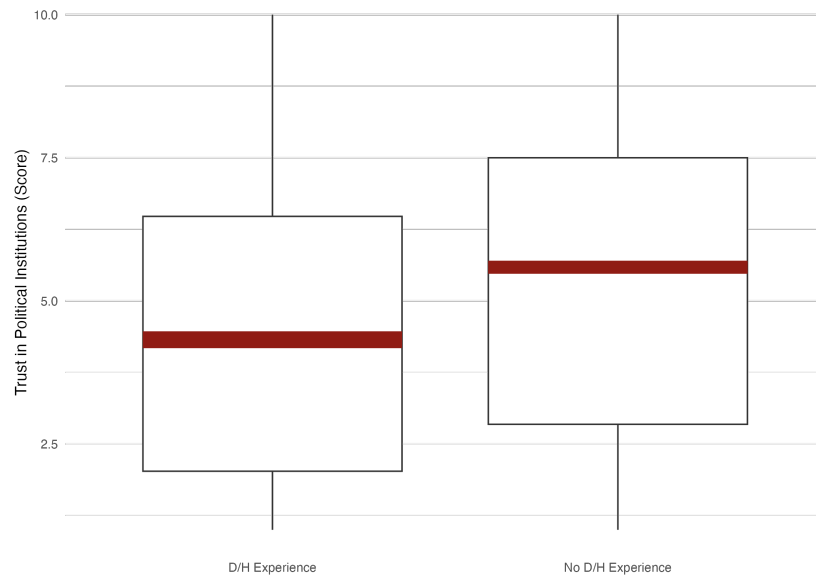
- Y_{ic} : political trust for individual i in region c
- X_{ic} : D/H event (1 = treated, 0 = control)
- τ : treatment effect (ATT)
- \mathbf{Z}_{ic} : vector of covariates
- β : coefficient vector for covariates
- γ_c : region fixed effects
- ε_{ic} : error term

6 Results

When we compare trust levels in political institutions between people who have experienced discrimination or harassment (D/H) due to their political opinions and those who have not, we observe that the former group typically achieves lower average scores, as shown in Figure 1. This figure reveals a strong negative correlation between experiencing D/H events and current levels of institutional trust. However, this relationship cannot be claimed to be causal. As explained previously, people who initially have lower trust in institutions might display higher levels of political participation or activism and, therefore, be more prone to experiencing discrimination.

We attempt to isolate a causal effect by matching people who have experienced a D/H event with people who have similar demographic characteristics and political traits but have not experienced such events. This approach allows us to approximate the counterfactual outcomes that we cannot directly observe in our data. To evaluate how this matching affects our analysis, we begin by assessing the initial balance in characteristics between people who have experienced D/H events and those who have not.

Figure 1: Trust in Political Institutions by Political D/H Experience



Note: The figure shows the distribution of Trust in Political Institutions Index scores for individuals who experienced political discrimination or harassment (D/H) versus those who did not. Index scores are derived from Logistic PCA of trust responses across nine institutional categories: local authorities, national authorities, police, prosecutors, public defense attorneys, judges, magistrates, political parties, and Parliament members. Red areas show 95% confidence intervals of median scores. Scores were re-scaled to fit a scale between 1-10.
Source: Eurovoices General Population Poll 2024

6.1 Assessing balance between groups

Appendix B shows balance statistics for the whole sample prior to matching. As we can observe, several covariates display standardized mean differences above the thresholds suggested by the literature—0.1 or, more conservatively, 0.05. Of particular note are the high values for average civic participation, interest in politics, and political alignment with the incumbent party, suggesting that the political profiles of people who have suffered D/H due to their political opinions and those who have not are systematically different. Similarly, age and ethnic composition differ significantly between both groups on average, indicating that the demographic composition of the groups is also systematically different. These insights are validated by the variance ratios and empirical cumulative distribution functions (eCDFs).¹²

The imbalances found between both groups suggest that estimates calculated using the full study sample would carry significant bias due to systematic differences in demographic and political characteristics. Therefore, a matching procedure to preprocess the data is a valid approach to reduce these biases.

After performing the matching procedure, we reassess the balance of covariates between the treatment and control groups. The results are presented in Table 3. As we can see, the matching procedure has significantly improved the balance of covariates between both groups. The standardized mean differences for all covariates are now below 0.05, indicating that the treatment and control groups are now comparable in terms of their demographic and political characteristics. This is further supported by the variance ratios, which are now close to 1 for all covariates, and the empirical cumulative distribution function (eCDF) statistics, which are also close to zero for all covariates. Appendix C provides a love plot that visually summarizes the balance of covariates before and after matching.

As a result of the matching procedure, our sample size was reduced from 44,235 individuals to 14,004 individuals. The final matched sample consists of 7,002 individuals who experienced D/H events and 7,002 individuals who did not. This reduction in sample size is a common outcome of matching procedures, as some individuals may not have suitable matches in the control group or may fall outside the common support region. This is the study sample we use to estimate the causal effect of political discrimination and harassment on trust in political institutions.

¹²The Std. Mean Diff. is the difference in group means standardized to a single scale for all covariates. The literature suggests values below 0.1—or more conservatively, below 0.05—as indicators of good balance (Ali et al., 2011; Stuart et al., 2013). The Var. Ratio compares the variance of a covariate between the treatment and control groups. Ratios close to 1 indicate good balance (Austin, 2009). Empirical Cumulative Distribution Function Statistics (eCDF) assess balance beyond the mean by comparing the full distribution of each covariate across groups (McCaffrey et al., 2004). Values close to zero indicate good balance.

Table 3: Post-Matching Covariate Balance Assessment

	Mean D/H Exp.	Mean No D/H Exp.	Std. Mean Diff.	Var. Ratio	eCDF Mean	eCDF Max
Distance	0.270	0.265	0.031	1.122	0.001	0.027
Female	0.468	0.461	0.014	-	0.007	0.007
Age	46.268	46.255	0.001	1.000	0.004	0.013
Rural	0.244	0.243	0.002	-	0.001	0.001
Constrained Fin. Situation	0.281	0.269	0.027	-	0.012	0.012
Higher Education	0.352	0.358	-0.013	-	0.006	0.006
Employed	0.579	0.584	-0.010	-	0.005	0.005
Married	0.570	0.580	-0.019	-	0.010	0.010
Foreigner	0.025	0.025	0.002	-	0.000	0.000
Ethnic Minority	0.513	0.512	0.003	-	0.001	0.001
Interest in Politics	0.663	0.658	0.011	-	0.005	0.005
Political Ideology	5.504	5.485	0.008	1.153	0.017	0.033
Political Al. with Incumbent	0.117	0.121	-0.012	-	0.004	0.004
Civic Participation Score	-2.499	-2.570	0.015	1.015	0.010	0.024
<i>Sample sizes:</i>	<i>Control</i>	<i>Treated</i>				
All	37,233	7,002				
Matched	7,002	7,002				
Unmatched	30,184	0				
Discarded	47	0				

Note: The Std. Mean Diff. is the difference in group means standardized to a single scale for all covariates. The literature suggests values below 0.1—or more conservatively, below 0.05—as indicators of good balance (Ali et al., 2011; Stuart et al., 2013). The Var. Ratio compares the variance of a covariate between the treatment and control groups. Ratios close to 1 indicate good balance (Austin, 2009). Empirical Cumulative Distribution Function Statistics (eCDF) assess balance beyond the mean by comparing the full distribution of each covariate across groups (McCaffrey et al., 2004). Values close to zero indicate good balance.

Source: Eurovoices General Population Poll 2024

6.2 The effect of political discrimination and harassment on trust in political institutions

To estimate the effect of political discrimination and harassment on trust in political institutions, we estimate four different specifications of Equation 1. The first specification includes only the treatment variable (experiencing a D/H event) and region fixed effects. The second specification adds demographic covariates, while the third specification adds both demographic and political covariates. The fourth specification replicates the third one, but it fits the model using the full unmatched sample. The results of these estimations are presented in Table 4.

Table 4: Effect of Political D/H Experiences on Institutional Trust

	(I)	(II)	(III)	(IV)
D/H Experience	-0.194***	-0.191***	-0.189***	-0.159***
Num.Obs.	14 004	14 004	14 004	44 235
Adj. R.sq.	0.074	0.090	0.110	0.101
Adj. Within R.sq.	0.022	0.039	0.060	0.046
Region FE	X	X	X	X
Dem. Cov.		X	X	X
Pol. Cov.			X	X
Sample	Matched	Matched	Matched	Full

Note: Table shows the results of regressing having experienced political discrimination or harassment (D/H) on the level of trust in political institutions, measured as an index score. Model (I) includes only the treatment variable (D/H event) and region fixed effects. Model (II) adds demographic covariates, while model (III) and (IV) adds both demographic and political covariates. Model (IV) uses the full unmatched sample to fit the regression. *, **, and *** represent statistical significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. Standard Errors are clustered at the regional level and paired-units weights were used for regressions using the matched sample. Estimates shown in the table have been transformed using an exponential transformation to interpret them as a percentage change.

Source: Eurovoices General Population Poll 2024

As we can observe, the coefficient for the D/H event variable is negative and statistically significant in all specifications. This indicates that experiencing political discrimination or harassment has a negative effect on trust in political institutions. Furthermore, the magnitude of this effect is consistent across specifications, ranging from -0.159 to -0.194, depending on the model. This means that individuals who have experienced political discrimination or harassment tend to have between 16% and 19% lower levels of trust in political institutions compared to those who have not experienced such events. This effect is large in the context of public opinion data and suggests a meaningful deterioration in institutional trust linked to experiences of political unfairness or targeting.

Interestingly, the magnitude of the coefficient is slightly smaller when using the full unmatched sample, which suggests that the bias introduced by using the full sample tends to underestimate the absolute effect of political discrimination and harassment on trust in political institutions.

Using a matched sample has allowed us to improve the precision of our estimates. However, achieving a balance in covariates using a matching procedure suggests that the treatment and control groups are comparable in ways that are more aligned with the assumptions necessary for causal inference (Branson, 2021). However, balance represents a necessary but not sufficient condition for causality, as unobserved confounders may still bias our estimates. Therefore, we perform a sensitivity analysis to assess the robustness of our results to potential violations of the unconfoundedness assumption.

6.3 Sensitivity analysis

We assess the robustness of our results to violations of the unconfoundedness assumption using the sensitivity analysis framework of Cinelli and Hazlett (2020). This method estimates the minimum strength of unobserved confounding needed to eliminate our treatment effect and evaluates how results vary under different confounding scenarios. Table 5 presents these sensitivity analysis results.

According to our results, the robustness value ($RV_{q=1}$) is 13.8%, which means that an unobserved confounder—orthogonal to the current set of covariates—would need to explain at least 13.8% of the variance in experiencing political discrimination or harassment (D/H) to nullify our estimated treatment effect.

Similarly, the robustness value for testing the null hypothesis that the treatment effect is zero ($RV_{q=1, \alpha=0.05}$) is 12.4%. This means that an unobserved confounder—orthogonal to the current set of covariates—would need to explain at least 12.4% of the variance in experiencing D/H events to make our treatment effect statistically insignificant at the 5% level.

Finally, the partial R^2 of the treatment ($R^2_{Y \sim D | \mathbf{X}, D}$) is 3.5%, which indicates that an unobserved confounder would need to explain at least 3.5% of the variance in trust in political institutions to nullify our estimated treatment effect.

These levels are relatively high. However, without accurate domain knowledge, it is very difficult to assess how strong the unobserved factors not captured in our data might be. To provide context, we benchmark these results against political alignment with the incumbent party—a covariate strongly related to institutional trust. Political alignment serves as a reasonable upper bound since few omitted variables would likely have stronger associations with institutional trust and D/H experiences than alignment with the governing party.

Table 5 shows the benchmarking results for political alignment with the incum-

bent party (*incpp*). According to the analysis, political alignment explains only 1.8% of the variance in trust in political institutions, which is significantly lower than the estimated partial R^2 and both RV values. Therefore, we can conclude that our results are robust to the possibility of unobserved confounding.

Table 5: Sensitivity Analysis: Main Results

Treatment:	Est.	S.E.	t-value	$R^2_{Y \sim D \mathbf{X}}$	$RV_{q=1}$	$RV_{q=1, \alpha=0.05}$
<i>D/H Experience</i>	-0.209	0.012	-17.523	2.2%	13.8%	12.4%
df = 13,880		<i>Bound (2x incpp):</i> $R^2_{Y \sim Z \mathbf{X}, D} = 1.8\%$, $R^2_{D \sim Z \mathbf{X}} = 0\%$				

Source: Eurovoices General Population Poll 2024

Our analysis further shows that results remain valid even with an omitted confounder 15 times stronger than political alignment (see Appendix D). Moreover, the treatment effect magnitude remains stable even under this extreme confounding scenario. This reinforces our conclusion that the results are robust to unobserved confounding, at least within the bounds established by our political alignment benchmark.

6.4 Polarization as a transmission channel

As mentioned earlier, we hypothesize that political polarization serves as a transmission channel through which political discrimination and harassment affect trust in political institutions. To test this hypothesis, we conduct a mediation analysis using the framework developed by Imai et al. (2011) and Imai and Yamamoto (2013). This analysis aims to determine how much of the total effect of D/H experiences on institutional trust operates through political polarization. The framework is based in the estimation of the system presented by equations 2 and 3 along with a bootstrapping procedure to test the mediation channel.

Table 6 presents three key estimates: the total effect, the average causal mediation effect (ACME), and the average direct effect (ADE). As we can observe, experiencing political discrimination or harassment leads to a large and highly significant drop in institutional trust: the Total Effect of -0.863 (on a 0–10 scale) tells us that, on average, those who report discrimination score 0.863 points lower in trust than those who do not. This effect is equivalent to the 18.9% decrease presented in Table 4, column (III).

To better understand the mechanisms behind this decline in trust, the analysis separates the effect into two components: the part that operates indirectly through increased political dissonance (the Average Causal Mediation Effect, or ACME), and the part that operates directly, independently of dissonance (the Average Direct Effect, or ADE).

The estimated ACME is -0.0063 , meaning that only about 0.006 points of the

decline in trust can be attributed to a pathway where discrimination increases political dissonance, which in turn lowers trust. By contrast, the ADE is -0.857 , nearly identical to the total effect, which shows that the vast majority of the impact of discrimination on trust occurs directly, without being mediated by political dissonance.

In other words, less than 1% of the total effect of political discrimination on institutional trust can be explained by changes in political dissonance. Although this indirect effect is statistically significant, it is substantively negligible. This implies that political dissonance does act as a mechanism, but it plays only a very minor role. Most of the loss in trust stems from other pathways—perhaps emotional responses like anger or fear, cognitive assessments of systemic injustice, or broader alienation from the political system that are not captured by the dissonance variable.

Table 6: Mediation Analysis Results

	Estimate	95% CI Lower	95% CI Upper	p-value
ACME	-0.006	-0.013	-0.001	0.02*
ADE	-0.857	-0.945	-0.770	0.00***
Total Effect	-0.863	-0.951	-0.776	0.00***
Prop. Mediated	0.007	0.001	0.015	0.02*

Note: Table shows the results of the mediation analysis. The ACME (Average Causal Mediation Effect) measures the indirect effect of having experienced D/H events on trust in political institutions through our measure of affective polarization (*mediator*). The ADE (Average Direct Effect) measures the direct effect on the outcome. The Total Effect is the sum of the ACME and ADE. The Prop. Mediated is the proportion of the total effect that is mediated through the mediator. *, **, and *** represent statistical significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. Estimates shown in the table tested using a bootstrapping approach with 10,000 replications.

Source: Eurovoices General Population Poll 2024



6.5 Heterogenous effects by political alignment

Given that our outcome variable is individual trust in political institutions, we hypothesize that there might be heterogeneous effects since political alignment with the incumbent party is a strong predictor of institutional trust. To test this hypothesis, we estimate the effect of political discrimination and harassment separately for two groups based on their political alignment: those aligned with the incumbent party and those who are not.

As Table 7 shows, experiencing a D/H event reduces trust in political institutions regardless of political alignment. However, the magnitude of this effect is significantly larger for individuals who are not aligned with the incumbent party. According to our estimates, experiencing a D/H event is associated with a 20% decrease in institutional trust for non-aligned individuals, compared to

a 14% decrease for those aligned with the incumbent party. The differences in magnitudes is statistically different at the 95% confidence level.

Table 7: Effect of Political D/H Experiences on Institutional Trust by Political Alignment

Pol. Alignment	Estimate	$Pr(> z)$	95% CI Lower	95% CI Lower	Visualization
Non-Incumbent	-0.195***	0.000	-0.215	-0.175	
Incumbent	-0.138***	0.000	-0.189	-0.084	

Note: Table shows the estimated coefficient of having experienced political discrimination or harrasment (D/H) on the level of trust in political institutions for two different groups depending on their political alignment: non-incumbent and incumbent. The estimation used the matched sample of 14,004 respondents, region fixed effects, demographic and political covariates, and clustered standard errors at the regional level. *, **, and *** represent statistical significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. Estimates shown in the table have been transformed using an exponential transformation to interpret them as a percentage change.

Source: Eurovoices General Population Poll 2024

We wanted to further explore the relationship between political alignment and the effect of political discrimination and harassment on trust in political institutions. Given the results presented in the previous section, we consider the possibility that the heterogenous effects found in politically aligned individuals might be affected by their level of political dissonance. To answer this question, we further increased the level of disaggregation by estimating the effect of political discrimination and harassment on trust in political institutions for four different groups depending on their political alignment and their level of political dissonance.¹³

As Table 8 shows, for individuals with a low level of political dissonance, experiencing a D/H event does not has any effect on their level of trust in political institutions if they are politically aligned with the incumbent party, while their counterparts that are not politically aligned experience a 18% decrease in trust in political institutions.

On the other hand, for individuals with a high level of political dissonance, the effect remains high between 18% and 21% for both politically aligned and non-aligned individuals. Interestingly, the effect is not statistically different between the two groups.

7 Discussion and Final Remarks

¹³The level of political dissonance was categorized in either high or low depending if the individual's measure was above or below the median of the full study sample, respectively.

Table 8: Effect of Political D/H Experiences on Institutional Trust by Level of Political Dissonance and Political Alignment

Pol. Diss.	Pol. Alignment	Estimate	$Pr(> z)$	95% CI Lower	95% CI Upper	Visualization
Low	Non-Incumbent	-0.180***	0.000	-0.209	-0.149	
	Incumbent	-0.085	0.000	-0.168	0.006	
High	Non-Incumbent	-0.211***	0.000	-0.238	-0.183	
	Incumbent	-0.179***	0.000	-0.241	-0.111	

Note: Table shows the estimated coefficient of having experienced political discrimination or harrasment (D/H) on the level of trust in political institutions for four different groups depending on their political alignment and their level of political dissonance. The estimation used the matched sample of 14,004 respondents, region fixed effects, demographic and political covariates, and clustered standard errors at the regional level. *, **, and *** represent statistical significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. Estimates shown in the table have been transformed using an exponential transformation to interpret them as a percentage change.

Source: Eurovoices General Population Poll 2024

A Survey methodology and sample characteristics by country

Country	Scope	Regions	Method	Sample Size	Data Collection
Austria	NUTS-1	3	Online	2,310	02/27-04/09/24
Belgium	NUTS-1	3	Online	1,521	03/01-04/12/24
Bulgaria	NUTS-1	2	Face to Face	1,500	01/19-02/24/24
Croatia	NUTS-2	4	Face to Face	2,000	02/15-03/28/24
Cyprus	NUTS-0	1	Face to Face	1,010	11/10/23-03/20/24
Czechia	NUTS-2	8	Online	2,024	02/26-04/10/24
Denmark	NUTS-2	5	Online	2,522	03/01-04/23/24
Estonia	NUTS-0	1	Online	1,044	02/26-04/11/24
Finland	NUTS-2	4	Online	2,218	02/26-04/14/24
France	NUTS-1	13	Online	6,195	02/26-04/16/24
Germany	NUTS-1	16	Online	8,055	03/01-04/24/24
Greece	NUTS-1	4	Face to Face	2,009	01/08-04/04/24
Hungary	NUTS-1		Online	1,750	02/07-02/29/24
Ireland	NUTS-2	3	Online	1,288	02/01-04/30/24
Italy	NUTS-1	5	Online	3,844	03/01-04/01/24
Latvia	NUTS-2	1	Face to Face	1,003	03/01-03/31/24
Lithuania	NUTS-1	2	Face to Face	1,500	03/01-04/21/24
Luxembourg	NUTS-0	1	Online	605	02/02-02/23/24
Malta	NUTS-0	1	Face to Face	500	01/02-03/13/24
Netherlands	NUTS-1	4	Online	3,050	03/01-04/17/24
Poland	NUTS-1	7	Face to Face	5,252	03/01-05/10/24
Portugal	NUTS-1	3	Face to Face	1,562	01/23-04/02/24
Romania	NUTS-2	4	Face to Face	3,018	03/09-05/12/24
Slovakia	NUTS-2	4	Online	2,200	02/07-02/29/24
Slovenia	NUTS-2	2	Online	1,019	02/26-04/10/24
Spain	NUTS-1	7	Online	3,536	02/26-04/17/24
Sweden	NUTS-1	3	Online	1,554	02/26-04/11/24

Source: Eurovoices General Population Poll 2024

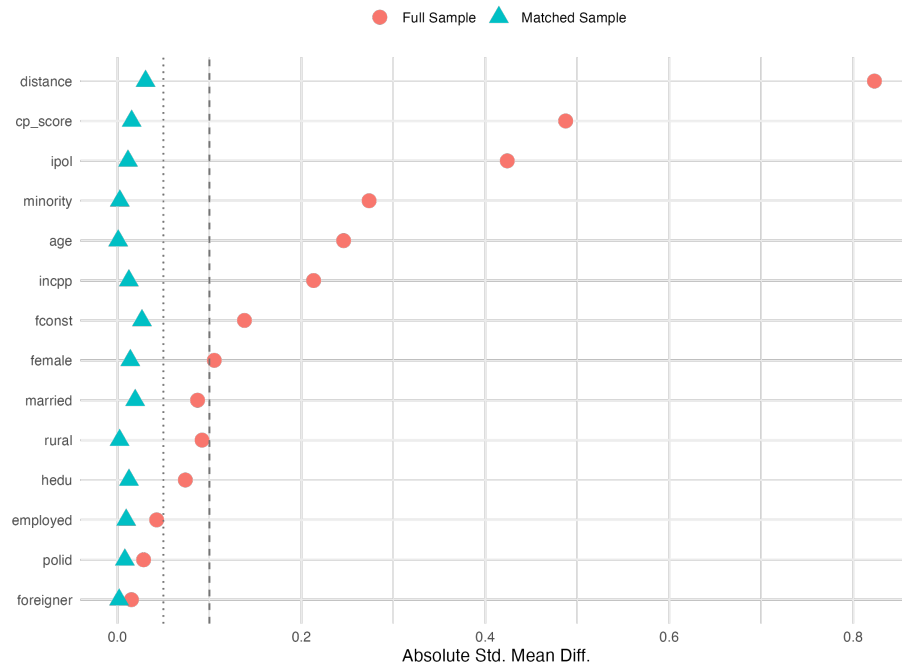
B Pre-Matching Covariate Balance Assessment

	Mean D/H Exp.	Mean No D/H Exp.	Std. Mean Diff.	Var. Ratio	eCDF Mean	eCDF Max
Distance	0.270	0.137	0.823	1.929	0.264	0.396
Female	0.468	0.521	-0.105	-	0.053	0.053
Age	46.268	50.309	-0.246	0.979	0.051	0.102
Rural	0.244	0.283	-0.092	-	0.039	0.039
Constrained Fin. Situation	0.281	0.219	0.138	-	0.062	0.06
Higher Education	0.352	0.317	0.074	-	0.035	0.035
Employed	0.579	0.558	0.043	-	0.021	0.021
Married	0.570	0.614	-0.087	-	0.043	0.043
Foreigner	0.025	0.022	0.015	-	0.002	0.002
Ethnic Minority	0.513	0.376	0.274	-	0.137	0.137
Interest in Politics	0.663	0.463	0.424	-	0.200	0.200
Political Ideology	5.504	5.435	0.028	1.206	0.021	0.041
Political Al. with Incumbent	0.117	0.186	-0.213	-	0.069	0.069
Civic Participation Score	-2.499	-4.734	0.487	1.721	0.145	0.259

Note: The Std. Mean Diff. is the difference in group means standardized to a single scale for all covariates. The literature suggests values below 0.1—or more conservatively, below 0.05—as indicators of good balance (Ali et al., 2011; Stuart et al., 2013). The Var. Ratio compares the variance of a covariate between the treatment and control groups. Ratios close to 1 indicate good balance (Austin, 2009). Empirical Cumulative Distribution Function Statistics (eCDF) assess balance beyond the mean by comparing the full distribution of each covariate across groups (McCaffrey et al., 2004). Values close to zero indicate good balance.

Source: Eurovoices General Population Poll 2024

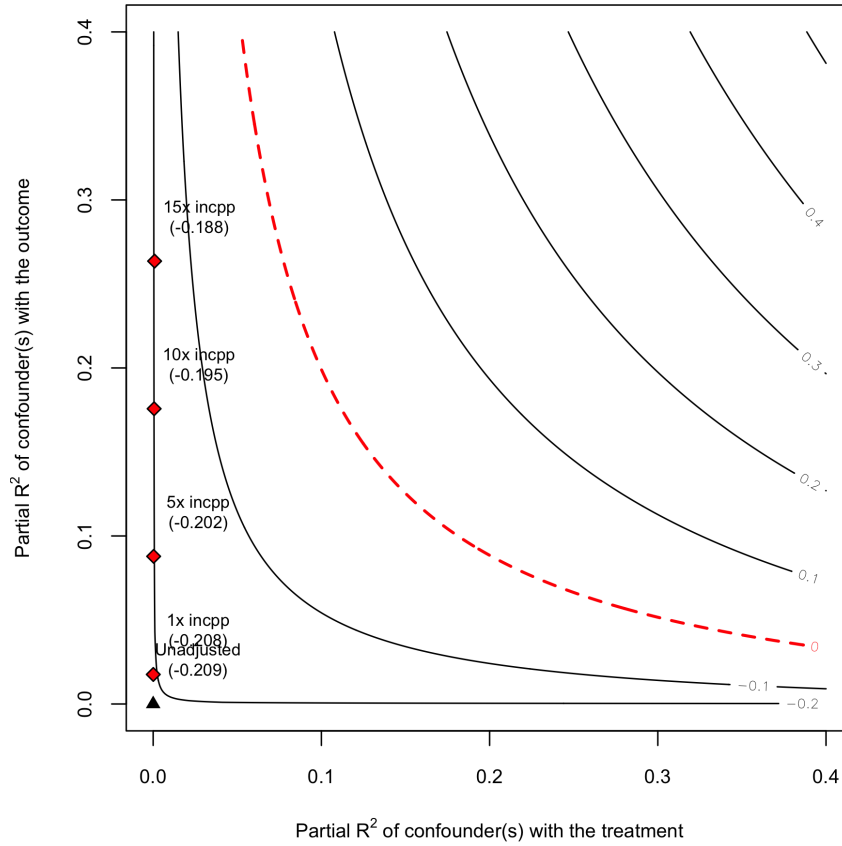
C Matching Covariate Balance Assessment



Note: The figure shows the Absolute Std. Mean Diff. for each covariate before (full sample) and after matching. The dashed and dotted lines represent the thresholds of 0.1 and 0.05, respectively, for good balance. The figure indicates that the matching procedure has significantly improved the balance of covariates between the treatment and control groups.

Source: Eurovoices General Population Poll 2024

D Sensitivity contour plots of point estimates



Note: The figure shows hypothetical sensitivity contours for the point estimates of the treatment effect of political discrimination and harassment on trust in political institutions. The contours represent the hypothetical residual share of variation of the treatment that an unobserved confounding explains (x-axis) and the the hypothetical partial R^2 of an unobserved confounding with the outcome (y-axis). The red line represents the level at which the unobserved confounding would nullify our results. The magnitude of the coefficient of interest (ATT) is shown within parenthesis.

Source: Eurovoices General Population Poll 2024