

Identity, Beliefs, and Political Conflict*

Giampaolo Bonomi[†], Nicola Gennaioli[‡] and Guido Tabellini[§]

First version: October 2018 This version: May 2021

Abstract

We present a theory of identity politics that builds on two ideas. First, when policy conflict renders a certain social divide - economic or cultural - salient, a voter identifies with his economic or cultural group. Second, the voter slants his beliefs towards the stereotype of the group he identifies with. We obtain three implications. First, voters' beliefs are polarized along the distinctive features of salient groups. Second, if the salience of cultural policies increases, cultural conflict rises, redistributive conflict falls, and polarization becomes more correlated across issues. Third, economic shocks hurting conservative voters may trigger a switch to cultural identity, causing these voters to demand less redistribution. We discuss U.S. survey evidence in light of these implications.

*We are grateful to Stefanie Stancheva, four referees and Roland Benabou, Ben Enke, Giovanni Burro, Tommaso Crosta, Ed Glaeser, Elhanan Helpman, Nathan Nunn, Raffaella Piccarreta, Giacomo Ponzetto, Jesse Shapiro, Andrei Shleifer, Marco Tabellini, Pierre Yared and to participants in workshops at ASREC, Bocconi, the Bundesbank, CEPR, CIFAR, the ESM, Harvard, the IMF, Honk Kong Baptist University, National Taiwan University, National University of Singapore, CEIBS, the LSE ERINN conference, the NBER, Princeton, UPF, Università di Padova, Università Statale di Milano, the University of Namur, and to Viola Corradini, Daniele d'Arienzo, Silvia Farina, Carlo Medici, Francesca Miserocchi and Giulia Travaglini for outstanding research assistance, and Silvia Tesauero for editorial assistance. Tabellini thanks the ERC grant 741643 and Gennaioli the ERC (GA 647782) and MIUR (FARE grant) for financial support.

[†]Department of Economics, UC San Diego. E-mail: gbonomi@ucsd.edu

[‡]Department of Finance and IGIER, Università Bocconi. E-mail: nicola.gennaioli@unibocconi.it

[§]Department of Economics and IGIER, Università Bocconi; CEPR; CESifo. E-mail: guido.tabellini@unibocconi.it

1 Introduction

In the last decade, the political systems of the US and of other advanced democracies have undergone a momentous change. Economic conflict over redistribution has weakened, while conflict over cultural issues such as immigration, race, and abortion has strengthened. Consider Figure 1, constructed using data from the American National Election Surveys (ANES). We create an index of the demand for redistribution as the first principal component of two questions on public spending. We construct an index of the demand for progressive cultural policies from opinions on immigration, race relations, and abortion. To control for shifting party positions, we estimate the residuals of both variables after conditioning on the respondent's party affiliation, interacted with wave fixed effects. Results are similar in the raw data, and when considering political independents only (see the Appendix). Panel A reports the variance of these variables between 1996 and 2016. After 2008, disagreement over cultural policies has sharply increased, while disagreement over redistribution has, if anything, declined.

Figure 1 here

Panel B shows that something else has changed: preferences over redistribution have become more correlated with those on cultural policy. In 2008, a progressive cultural outlook was positively correlated with more support for redistribution, but this connection more than doubled in 2016. Wu (2020) nicely documents similar patterns using a large dataset on US college freshmen.

These trends cannot be explained by better sorting of extremists into parties (e.g. Klein 2020). Rather, they indicate population-wide changes in voter demands: cultural conflict has intensified and has become more correlated with conflict on redistribution. Another driver of increasing polarization may be growing partisan divisions and hatred (e.g. Gentzkow 2016). This mechanism likely plays a role, but it cannot easily explain why redistributive conflict – a historically partisan issue – has not itself become more intense and why these changes are also observed among political independents. Why have cultural divisions increased? Why has redistributive conflict not risen despite growing income inequality?

This paper shows that these questions can be fruitfully addressed by studying how social identities influence voters' behavior. The basic idea is that when voters abandon their class identity and re-define themselves in terms of their moral or religious values, the latter become more important to explain their beliefs in several domains. Based on social psychology (Tajfel and Turner 1979, Turner et al. 1987), we build a model of this phenomenon resting on two pillars.

First, voters can identify with their income group, upper vs. lower class, or with their

cultural group, social progressives vs. conservatives. At any point in time, voters identify with the groups that are most salient, formalized as those having the strongest policy conflict. When class conflict is stronger, voters identify with their class. When cultural conflict is stronger, they identify with their cultural group.

Second, identity causes voters to slant their beliefs towards the stereotypical views of their group. In psychology, this phenomenon is called belief polarization (Mackie 1986). It occurs because voters overweight the distinctive opinions of group members, or those of group-linked experts and media. Kahan (2015) shows that cultural identities polarize risk perceptions about environmental issues, abortion and gun control. We formalize group-stereotyped beliefs by adapting the model in Bordalo et al. (2016), and then study its implications for politics.

Our setup features two policy instruments: a distortionary income tax financing a public good and a "cultural policy" that we interpret as civil rights or control of immigration. A voter is described by two traits: expected future income and cultural progressiveness. Richer voters desire a lower tax, due to their higher tax burden. More progressive voters want a more liberal cultural policy and also a higher tax, because they like the public good more. Culture reflects deep-seated values such as moral universalism or religiosity, which shape beliefs across different domains (Haidt 2012, Enke et al. 2020). Based on his personal traits, a voter belongs to an income class, upper or lower, and to a cultural group, progressive or conservative, and he can identify with either.

We obtain three main insights. First, identity creates belief distortions that polarize ingroup-outgroup conflict along the currently salient dimension. Under class identity, lower class voters are too pessimistic about their future income, and vice-versa for upper class voters. These belief distortions boost redistributive conflict. Under cultural identity, progressive voters become even more progressive, and vice-versa for conservatives. This boosts cultural conflict. Perceived polarization is also excessive: society is divided into "us vs. them". Thus, identity can shed light on belief distortions in politics and predicts that beliefs are shaped by the changing salience of cultural, economic, or political groups (Alesina et al. 2018 a,b, Kahan 2015, Westfall et al. 2015).

Second, if the welfare relevance of cultural policy increases, due for instance to a large inflow of immigrants, identity switches from class to culture. As this happens, culture becomes a stronger driver of policy views. As a result, and in line with Figure 1, views over cultural policy become: (i) more polarized, and (ii) more correlated with views over redistribution. Critically, by obfuscating class differences, cultural identity also dampens redistributive conflict. This can explain why the latter is stable or declining after 2008 despite growing income inequality. Using ANES data, we show that – consistent with our theory – disagreement between upper and lower class voters on redistribution has sharply declined over time.

Third, economic shocks that boost conflict among cultural groups can also trigger a shift to

cultural identity. We offer two examples: skilled biased technical change and globalization. If these shocks hurt less educated and hence more conservative voters, and benefit more educated and hence more progressive voters, they make cultural cleavages more salient and can induce a switch to cultural identity. As a result, economic losers become more socially and fiscally conservative. This is consistent with the evidence in Autor et al. (2020) and Colantone and Stanig (2017), who show that losses from international trade foster support for right-wing parties.

We relate to recent work on the influence of moral universalism and fairness on policy preferences (Enke et al. 2020, Stantcheva 2020). In our paper, the importance of cultural factors is time varying. Norris and Inglehart (2019) stress the growing importance of cultural divisions. Frank (2004) vividly describes the "cultural backlash" in Kansas, a state that was Democratic in the past and then became culturally and fiscally conservative. Relative to this work, we show that the cultural divide can be amplified by specific economic shocks due to trade or technology. We also connect to recent research explaining the rise of populism as a reaction to economic distress (Guriev and Papaioannou 2020).

Murphy and Shleifer (2004), Glaeser (2005) and Wu (2020) explain changing voter beliefs based on political supply. Politicians attract voters by catering to their cultural views, and persuade them on issues they are less focused on, such as redistribution. One challenge for this approach is to explain why persuasion should work with economic losers. Our model offers a rationale: specific economic shocks have made cultural identity more salient. More broadly, cultural conflict has become more important relative to redistribution in several countries with different party systems, but subject to similar globalization and technology trends (Goodhart 2017, Evans and Mellon 2016, Guriev and Papaioannou 2020). This phenomenon points to a role of demand factors.

We also contribute to a growing literature on identity in economics. Akerlof and Kranton (2000) develop the first economic model where identity changes the payoffs of certain actions. They do not consider beliefs. The seminal paper by Shayo (2009) introduces identity in political economy. Grossman and Helpman (2020) study how views over trade policy are affected by social identities. In these papers, voters obtain utility from the welfare of their group – which affects their policy demands – and identify either with their narrow class or the broader nation. In our model identity affects beliefs, which allows us to relate to the evidence on voters' misperceptions (Flynn et al. 2017, Achen and Bartels 2016, Johnston et al. 2017, Stancheva 2020). In addition, the groups we consider – class and culture-based – are heterogeneous. Thus, identity switches do not cause social integration or disintegration, but rather a realignment of voters into different social partitions.¹

¹Bénabou and Tirole (2016) offer a different approach in which identity reflects beliefs about one-self, and beliefs adjust to improve self-image and own welfare, taking anticipatory utility into account.

A tradition in political science, started by Key (1955), studies electoral realignments (e.g., Sundquist 1983, Mayhew 2004). This work focuses on the US and seeks to explain lasting changes in party positions and in the composition of party supporters. We endogenize these realignments, abstracting from parties and their leaders.

The paper proceeds as follows. Section 2 discusses the social psychology of identity and presents our model of beliefs. Section 3 studies the causes of identity switches and how they affect beliefs and equilibrium policy. Section 4 derives testable predictions on changes in policy preferences. Section 5 illustrates the mechanism of our theory using US survey data and assesses alternative explanations for the same facts.

2 The Social Psychology of Identity and Stereotypes

The "Social Identity Perspective", the leading theory of identity and intergroup relations, combines Social Identity Theory (SIT, Tajfel and Turner, 1979), and Self Categorization Theory (SCT, Turner et al. 1987, Hogg and Abrams 2006). We discuss the psychology of identity, its connection to beliefs, and our formalization of belief distortions.

2.1 Identity Formation

According to SCT, identity is a form of self-categorization, leading people to perceive society through the lens of a group they belong to. Each person belongs to many groups: income class, religion, nation, etc., but he/she does not necessarily identify with all of them at the same time. Identity in fact depends on which social partition is salient (e.g., Hogg and Abrams 1998). For instance, at a football match one's own team is the salient group: it captures the cleavage of that moment, so it affects individual behavior more than other groups do. When participating in a union strike, though, the same person may identify with the lower class, and assimilate his behavior to it.

Which cleavage is salient depends on the so called "meta contrast ratio". A person identifies with a group of people who are: i) similar to him, but also and crucially ii) highly dissimilar from, or in conflict with, the outgroup (Oakes 1987). For instance, at a football match, the group of "team supporters" is more salient than that of "football lovers" because it reflects the cleavage of the moment. Similarly, during a strike the "lower class" is more salient than the broader group of "economic producers".

This idea naturally travels to politics. A shock, like a police officer killing a black person, renders a social partition salient, in this case that between those who believe that racial discrimination is significant and those who believe that the law has been enforced. Many people identify as members of either group. How does this affect beliefs and behavior?

2.2 Stereotyped Beliefs and Depersonalization

According to SIT and SCT, identity causes perceptions of self and others to be tainted by group features (see also Sherif 1936, Festinger 1950). Experiments on groups of objects by Tajfel and Wilkes (1963) and subsequent work in social settings (Haslam and Turner 1992) show that this occurs via two effects.

First, people stereotype groups by exaggerating differences between ingroups and outgroups. In the previous example, those who believe that the police is biased against minorities stereotype the other group as "racist", while the skeptics stereotype the others as "radicals". Society is divided into "us vs. them". Second, identity causes the individual to "depersonalize", namely to move his beliefs closer to the stereotypical group member. McGarty et al. (1992) define the stereotype as the group's most *representative* type: "...the less a person differs from ingroup members and the more he or she differs from outgroup members, the more representative he or she is of the ingroup". The stereotype is not necessarily the modal or average group trait, it must be distinctive relative to the outgroup. Thus, depersonalization creates the phenomenon of "group polarization", whereby group members hold a more extreme position on an issue than their individually expressed positions (e.g., Mackie 1986). In the example of the police officer, even moderate people slant their beliefs towards either left-wing radicalism or racism, at least to some extent, enhancing conflict.

Experiments using the "minimal group paradigm" found that even arbitrary groups affect individual behavior when they are made salient (Tajfel et al. 1971). SIT holds that, beyond affecting beliefs, identity can also yield positive self-esteem if one identifies with a high status group. Some experiments document ingroup favoritism, which can be viewed as enhancing group status. Existing political economy models of identity are centrally built on this ingredient (e.g., Shayo 2009). To highlight the new implications of our mechanism, we abstract from this effect. Status cannot be the sole driver of identity, because individuals often identify with underdog groups.²

2.3 Formalizing Distorted Beliefs

We depart from a standard political economy setup because voter's beliefs and political preferences are malleable, and change based on the salient social identity. To see how this works, consider a voter evaluating the benefit $\tilde{\psi}$ of a cultural policy such as opening to immigration or extending civil rights. The voter has a database of information in his memory, summarized by the distribution $z(\tilde{\psi}|\psi)$, where ψ is the mean of $\tilde{\psi}$. High values of $\tilde{\psi}$ capture facts or value

²There is debate on whether it is always the case that group identity leads to ingroup bias (e.g., Hinkle and Brown 1990). Ingroup bias tends to arise when discrimination is a group norm but not otherwise. A belief based mechanism like ours can arguably generate such effect.

judgments favoring a liberal policy. In the case of immigration, high $\tilde{\psi}$ may capture reports of low crime rates among immigrants or altruistic attitudes towards strangers. In the case of civil rights, high $\tilde{\psi}$ may capture statistics of minority discrimination or benefits of diversity. Low values of $\tilde{\psi}$ stand for the opposite factual and value considerations.

The average value ψ captures a voter's core belief. A voter with higher ψ is more progressive. There is a cdf $H(\psi)$ of voters, who are partitioned into two groups: the social conservatives, who have $\psi < \hat{\psi}$, and the social progressives, with $\psi > \hat{\psi}$. The threshold $\hat{\psi}$ is historically given. A voter identifies himself with one or the other group, depending on his type. Here we take group identity as given. We endogenize it in Section 3.

In standard political economy models, the voter aggregates all the information and evaluates the policy at ψ . In our model, in line with Section 2, the voter's identity cues him to retrieve facts or values that are stereotypical of his own group, downplaying the rest.

We formalize this process following Bordalo et al. (2016, BCGS henceforth): a person identified with group G overweights the trait that is relatively more frequent in this group compared to outgroup \bar{G} .³ As in McGarty et al. (1992), the stereotype is distinctive of the group, but not necessarily frequent.

Denote by $z^\theta \left(\tilde{\psi} | \psi, G \right)$ the distorted beliefs of a voter with core beliefs ψ when he identifies with group G . As we shall see later, superscript θ captures the extent of belief distortions. Following BCGS (2016), this is given by:

$$z^\theta \left(\tilde{\psi} | \psi, G \right) \propto z \left(\tilde{\psi} | \psi \right) \left[\frac{z^\theta \left(\tilde{\psi} | \bar{\psi}_G, G \right)}{z^\theta \left(\tilde{\psi} | \bar{\psi}_{\bar{G}}, \bar{G} \right)} \right]^\chi, \quad (1)$$

where $\bar{\psi}_G \equiv \mathbb{E}[\psi | G]$ is the core belief of the average member of G . The likelihood ratio in Equation (1) captures the overweighting of facts or opinions $\tilde{\psi}$ that are more frequently held by the average member of group G compared to \bar{G} . Parameter $\chi \geq 0$ captures the degree of overweighting.⁴

Relative to BCGS (2016), Equation (1) makes two innovations. First, the stereotypical opinions of social groups also affect a voter's beliefs about the world, not just his beliefs about others. Thus, group beliefs emerge as a fixed point: the beliefs about the world of average group members are distorted by the stereotypical opinions of groups; group stereotypes, in

³BCGS (2016) show that this model accounts for observed belief distortions in social domains, but also for beliefs in macro and finance (Gennaioli and Shleifer 2018). Closer to the current setting, BCGS (2019) show that gender stereotypes influence self confidence. When assessing own ability, people are too optimistic/pessimistic in domains of knowledge where their gender has a competitive advantage/disadvantage.

⁴In the BCGS model the likelihood ratio is defined using the distributions of $\tilde{\psi}$ in groups G and \bar{G} . To obtain these, one should aggregate a family of individual distributions $z^\theta \left(\tilde{\psi} | \psi, G \right)$, which is intractable. To simplify, we capture each group using the belief distribution for its average type. This assumption preserves the idea that stereotypes magnify average group differences, and yields convenient closed form solutions.

turn, reflect the distorted beliefs of group members. Second, the group G through which social reality is perceived depends on identity, which we make endogenous in the next section.

The fixed point problem that jointly determines the beliefs of average groups members and group stereotypes proves tractable if $z(\tilde{\psi}|\psi)$ is Gaussian. All proofs are in Appendix 1.

Proposition 1 *If $\chi < 1/2$ the fixed point for average group beliefs exists, it is unique and stable. Let $\theta \equiv \frac{\chi}{1-2\chi} > 0$. Identification with group G distorts the beliefs of voter ψ as follows:*

$$\psi_G^\theta = \psi + \theta (\bar{\psi}_G - \bar{\psi}_{\bar{G}}). \quad (2)$$

If $\theta = 0$, beliefs are rational and identity plays no role. If $\theta > 0$, identity distorts beliefs away from their core ψ . Identification with the conservatives, cues the voter to think about the risks of progressive policies, so he becomes even more conservative. Identification with the progressives, cues him to think about the benefits of progressive policies, so he becomes even more progressive.

Equation (2) is consistent with the evidence in Kahan (2015), showing that people with different religious or political orientations exhibit sharp factual disagreement over natural selection and climate change, even if they are knowledgeable about science. He proposes a theory of "identity protective cognition" whereby individuals appraise information in a way that buttresses beliefs associated with their ingroup.

Beliefs can also be distorted by identification with political parties. In this case, party positions cue partisan voters to retrieve party stereotypes. Our model is not inconsistent with this possibility, but here we focus on social groups. This can explain political realignments that weaken pre-existing party affiliations and that also occur for non-partisan voters.

Equation (2) also sheds light on priming effects. A long-standing tradition in studies of mass opinion holds that individual beliefs change when a political or socioeconomic group is primed (Zaller 1992, Janky 2018, Han and Wackman 2017). Equation (2) disciplines these effects: if priming works through identity, it distorts beliefs in the direction of observed disagreement between ingroups and outgroups $(\bar{\psi}_G^\theta - \bar{\psi}_{\bar{G}}^\theta)$.⁵ Priming effects are then predictably heterogeneous because they cue voters to focus on the different groups to which they belong.⁶

These priming effects are different from learning, whereby a voter deliberately combines new information with prior beliefs. As shown in the next section, if identity switches to a different social group, the voter overweights the distinctive traits of the new group he identifies with, and his beliefs change even if no information is provided.

⁵Equation (2) can in fact be rewritten as $\psi_G^\theta = \psi + \left(\frac{\theta}{1+\theta}\right) (\bar{\psi}_G^\theta - \bar{\psi}_{\bar{G}}^\theta)$.

⁶In Alesina et al. (2018a), demand for redistribution increases when subjects are provided with pessimistic information about mobility, but this effect is only present for left-wing respondents. Kuziemko et al. (2015) shows that informing poor people about their income rank in society increases their demand for inheritance taxes. These findings are consistent with Equation (2) if one considers the treatment as priming class identity.

Recent research documents pervasive voter misperceptions. Relative to right-wing respondents, left-wing respondents perceive lower social mobility (Alesina et al. 2018b), larger inequality (Gimpelson and Treisman 2018), a lower number of immigrants (Grigorieff et al. 2020), and perceive immigrants as having worse features than natives (Alesina et al. 2018a). Future work may study whether these misperceptions are determined by, and change with, prevailing social or political identities.

One key implication of Equation (2) concerns group polarization.

Corollary 1 *When all voters identify with groups G and \bar{G} , disagreement among average group members is amplified relative to disagreement in rational beliefs. Formally,*

$$\bar{\psi}_G^\theta - \bar{\psi}_{\bar{G}}^\theta = (\bar{\psi}_G - \bar{\psi}_{\bar{G}}) (1 + 2\theta) \geq \bar{\psi}_G - \bar{\psi}_{\bar{G}}.$$

Identity boosts polarization by activating stereotypes. When we endogenize identity, this result proves important to account for growing cultural conflict in Figure 1. We assume the selective recall parameter θ to be exogenous. But if political leaders or social media disseminate stereotypes, these more readily come to mind - θ increases - causing more polarization.

Our model also implies that perceived polarization is greater than actual polarization. Consider the beliefs held in society about the distorted beliefs $z^\theta \left(\tilde{\psi} | \bar{\psi}_G, G \right)$ of the average member of G , namely the second order beliefs about $\bar{\psi}_G$. Due to stereotypes, group-distinctive traits are overweighted as in (1), yielding the result below.

Corollary 2 *Denote by $\hat{\psi}_G^\theta$ and $\hat{\psi}_{\bar{G}}^\theta$ the perceived mean positions of the average group members. Group stereotypes imply:*

$$\hat{\psi}_G^\theta - \hat{\psi}_{\bar{G}}^\theta = \left(\bar{\psi}_G^\theta - \bar{\psi}_{\bar{G}}^\theta \right) \left(\frac{1 + 4\theta}{1 + 2\theta} \right) \geq \bar{\psi}_G^\theta - \bar{\psi}_{\bar{G}}^\theta.$$

When thinking about the progressives, people overweight their stereotypical members, who are very progressive. When thinking about the conservatives, their stereotypical extreme members are overweighted. Hence, perceived disagreement exceeds actual disagreement. This effect is due to stereotypical thinking, not to identity: in fact, even a non-identified voter exaggerates polarization if $\theta > 0$.

There is a large literature measuring perceived and affective polarization among US parties. Westfall et al. (2015) and BCGS (2016) show that US voters exaggerate differences in the policy views of Democrats and Republicans. These exaggerations are held by all voters, including the non politically affiliated ones, which is consistent with Corollary 2. Bordalo et al. (2020) offer intriguing new evidence: the gap between perceived and actual polarization is

especially large in those policy issues that are more salient and in which there is more actual disagreement. This is a direct implication of our model when identity is endogenized.⁷

In the current example, identity distorts beliefs and value judgments regarding the general effects of a policy, say immigration. Identity can also distort a voter's perception of his own situation, such as his income rank in society or exposure to foreign competition. When forming beliefs about these aspects, the voter overweights the facts or opinions that are more frequently held in his group compared to the outgroup, as in Equation (1). A voter identified with the lower class retrieves his group's distinctive beliefs of economic backwardness, so he becomes too pessimistic about his future income.⁸ Similarly, a voter identifying with the losers from trade retrieves examples of workers similar to him who lost their job due to foreign competition, increasing his demand for trade protection.

3 Identity and Political Conflict

We study a simple model of conflict over redistribution and cultural policy, in which voters can identify with their economic class or cultural group. We show that if cultural conflict becomes more salient, voters switch from class to cultural identity and their beliefs change accordingly. This, in turn, changes equilibrium policies.

3.1 The Model

There are two policies. The first is a proportional income tax $\tau \geq 0$ that finances a public good. It entails quadratic distortions, $-\frac{\varphi}{2}\tau^2$, $\varphi > 0$, that reduce aggregate income. The second is cultural policy q . It captures policies mainly based on values, rather than on material interests, such as abortion, civil rights, race relations and immigration.⁹ Larger q indicates a

⁷Gentzkow (2016) shows that in the US perceived polarization and distrust of political rivals has increased more than actual divergence in policy views. As we show in Proposition 4, our model can account for this effect as the result of a shift from economic to cultural identity. Ahler (2018) shows that correcting misperceptions about the out-party also reduces affective polarization. Druckman and Levendusky (2019) show that affective polarization is boosted by retrieval of extreme party stereotypes, again consistent with our model.

⁸Cruces et al. (2013) document that Buenos Aires residents misperceive their *current* income on the basis of local conditions. Voters from rich neighborhoods underestimate their income rank relative to voters from poor neighborhoods, as if each voter uses the local income distribution as a reference. Misperception of current income and social mobility may be due to many factors, including local conditions. Our point is that when class conflict becomes salient, individuals identify with their class and misperceptions become class-based. Consistent with this view, Stantcheva (2020) shows that US Republicans (Democrats) overestimate (underestimate) their income rank in society.

⁹Our conclusions also hold if we allow preferences over immigration to depend on income, provided the dependence is weaker than that on culture. Card et al. (2012) study attitudes towards immigration in the European Social Survey. They argue that the bulk of attitudes is attributable to preferences over "compositional amenities" such as cultural homogeneity, religion, language, especially across more and less educated natives. Economic factors such as the impact of immigration on wages plays a small role.

more liberal stance (e.g., more extensive civil rights or immigration).

There is a set of measure one of voters who differ in their expected relative income, ε , and their cultural traits, ψ . A voter type is summarized by vector (ψ, ε) , which is distributed in society according to the normal cdf $H(\psi, \varepsilon)$, with mean $(0, 0)$, unit standard deviations and correlation coefficient ρ . Income and social progressiveness are non-negatively correlated, $1 > \rho \geq 0$, owing for instance to education.

The policy evaluations of a voter depend on three sources of uncertainty. The first one is about the effects of q . Preferences over q follow the quadratic loss $\frac{1}{2} (q - \tilde{\psi})^2$. $\tilde{\psi}$ is the voter's preferred policy, which is Gaussian with voter-specific mean ψ . Higher ψ stands for more socially progressive culture and hence higher preferred q .

Second, voters are uncertain about their tax burden because their income $1 + \tilde{\varepsilon}$ is stochastic. $\tilde{\varepsilon}$ is Gaussian with voter-specific mean ε . Thus, a rational voter of type ε expects to earn $1 + \varepsilon$. Beliefs about $\tilde{\varepsilon}$ capture beliefs over future relative income, and higher ε implies a higher tax burden.

Third, voters are uncertain about the benefit of the public good g . Its marginal utility $\tilde{\nu}$ is also Gaussian with expected value $\nu + \beta\psi$. Since ψ has zero population mean, the average marginal value of g is $\nu > 1$. Parameter $\beta \in (0, 1)$ reflects the influence of culture on preferences for the public good. The distribution of $\tilde{\nu}$ differs from that of $\tilde{\psi}$, because these beliefs refer to different policies.

The key point here is that the voter's culture, embodied in ψ , systematically affects his preferences over q and τ . One important cultural trait is moral universalism (Haidt 2012). More universalistic people apply their value system to socially more distant individuals (Tabellini 2008, Enke et al. 2021). Thus, they are more progressive on civil rights or immigration, but they also trust the government more and they are more in favor of redistribution (Enke et al. 2020, Haidt 2012). Another relevant trait is religiosity. More religious people are more conservative on civil rights, featuring lower $\tilde{\psi}$, and they also trust the government less (Enke et al. 2021), which reduces $\tilde{\nu}$. Culture can also reflect fairness norms, which affect preferences over tax policy (Stancheva 2020) and attitudes towards minorities. These traits do not need to be independent from each other. Enke et al. (2021) show that religiosity is negatively correlated with universalism. What matters for us is that they affect voters' views over q and τ .

Since ε has zero mean in the population, aggregate income is 1. Thus, the quantity of g is equal to the tax rate τ . The expected utility of rational voter (ε, ψ) is, up to an additive constant:

$$W^{\varepsilon\psi}(\tau, q) = (1 + \varepsilon)(1 - \tau) - \frac{\varphi}{2}\tau^2 + (\nu + \beta\psi)\tau - \frac{\kappa}{2}(q - \psi)^2, \quad (3)$$

where superscript $\varepsilon\psi$ denotes the voter's type and $\kappa > 0$ captures the weight attached to

policy q . Neglecting non-negativity constraints, the rational bliss point of voter (ε, ψ) is:

$$\tau^{\varepsilon\psi} = \frac{(\nu + \beta\psi) - (1 + \varepsilon)}{\varphi}, \quad q^{\varepsilon\psi} = \psi. \quad (4)$$

More progressive voters (higher ψ) demand more redistribution and a more liberal cultural policy. Richer voters (higher ε) demand less redistribution because of their greater tax burden.

The socially optimal policy maximizes aggregate rational welfare:

$$\int W^{\varepsilon\psi}(\tau, q) dH(\varepsilon, \psi) d\varepsilon d\psi.$$

Given linear private and public consumption, and given the assumed normal cdf $H(\varepsilon, \psi)$, the socially optimal policies are:

$$\tau^0 = \frac{\nu - 1}{\varphi}, \quad q^0 = 0. \quad (5)$$

Equation (4) shows that policy disagreement combines two underlying conflicts. One is the rich vs poor conflict over taxes, captured by expected relative income ε . The second is cultural conflict, captured by ψ , which affects the evaluation of both redistributive and cultural policies. As emphasized by Sundquist (1983), major realignments in American politics have occurred when the main parties shifted their positions along the economic and cultural divides. We endogenize these realignments as the product of voters' changing identities across economic and cultural groups.

3.2 Endogenous Identity

3.2.1 Groups

Social groups are defined based on income ε and culture ψ . With respect to culture, a voter can either be socially conservative, $SC \equiv \{\psi | \psi < \widehat{\psi}\}$, or progressive, $SP \equiv \{\psi | \psi \geq \widehat{\psi}\}$, where $\widehat{\psi}$ is historically given. With respect to income, voters belong to the upper class $U \equiv \{\varepsilon | \varepsilon \geq \widehat{\varepsilon}\}$ or to the lower class $L \equiv \{\varepsilon | \varepsilon < \widehat{\varepsilon}\}$, where $\widehat{\varepsilon}$ is again historically given. A voter can identify either economically or culturally, but not both. For instance, a poor and conservative voter ($\varepsilon < \widehat{\varepsilon}$, $\psi < \widehat{\psi}$) can only identify with the lower class or the conservatives. In Appendix 3 we allow for identification with narrower groups. The analysis is more complicated, but our main results continue to hold, as we discuss below.

Group G is summarized by type $(\bar{\varepsilon}_G, \bar{\psi}_G)$, which averages the income and culture of all group members regardless of whether or not they identify with G . If ψ and ε are correlated, ingroups and outgroups differ in both income and culture. Bivariate normality of $H(\varepsilon, \psi)$

implies:

$$\bar{\varepsilon}_{SP} - \bar{\varepsilon}_{SC} = \rho (\bar{\psi}_{SP} - \bar{\psi}_{SC}), \quad (6)$$

$$\bar{\psi}_U - \bar{\psi}_L = \rho (\bar{\varepsilon}_U - \bar{\varepsilon}_L). \quad (7)$$

If $\rho > 0$, the progressives are richer than the conservatives, and the upper class is more progressive than the lower class. Of course, since $\rho < 1$, cultural groups mostly differ along culture, $\bar{\psi}_{SP} - \bar{\psi}_{SC} > \bar{\varepsilon}_{SP} - \bar{\varepsilon}_{SC}$ and economic classes along income, $\bar{\varepsilon}_U - \bar{\varepsilon}_L > \bar{\psi}_U - \bar{\psi}_L$. We make the following reasonable assumption:

$$\rho < \left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right) < \frac{1}{\rho}, \quad (A1)$$

which implies that income differences are larger between classes than between cultural groups, $\bar{\varepsilon}_U - \bar{\varepsilon}_L > \bar{\varepsilon}_{SP} - \bar{\varepsilon}_{SC}$, and viceversa for cultural differences, $\bar{\psi}_{SP} - \bar{\psi}_{SC} > \bar{\psi}_U - \bar{\psi}_L$.¹⁰

3.2.2 Metacontrast and Identity

How is a voter's identity determined? As discussed in Section 2, identity maximizes: i) similarity between oneself and the ingroups, and ii) conflict between the ingroups and the outgroups. This tradeoff shapes which social partition, either economic or cultural, is more salient. To emphasize group conflict, which is important in politics, we focus on criterion ii), subject to the constraint that the voter can only identify with his income class or cultural group, because he is naturally more similar to them than to the outgroups. As a result, all voters identify along the dimension, be it class or culture, that maximizes group conflict. Below we discuss what happens if identity reflects a full tradeoff between criteria (i) and (ii).

Conflict between groups G and \bar{G} is measured by the welfare loss that the average member of G experiences when moving from his ideal policy (τ^G, q^G) to the bliss point of the average outgroup, $(\tau^{\bar{G}}, q^{\bar{G}})$:

$$C(G, \bar{G}) = W^G(\tau^G, q^G) - W^G(\tau^{\bar{G}}, q^{\bar{G}}). \quad (8)$$

The definition uses rational bliss points, but little changes if we use stereotyped bliss points. In the Appendix we prove that conflict is measured by:

¹⁰In what follows, we discuss the implications of whether $\bar{\psi}_{SP} - \bar{\psi}_{SC} \leq \bar{\varepsilon}_U - \bar{\varepsilon}_L$. Given that $H(\psi, \varepsilon)$ is normal with mean 0 and unit variances, it can be shown that, if $\hat{\varepsilon}$ and $\hat{\psi}$ have the same sign, then

$$\bar{\psi}_{SP} - \bar{\psi}_{SC} \gtrless \bar{\varepsilon}_U - \bar{\varepsilon}_L \quad \text{as} \quad |\hat{\psi}| \gtrless |\hat{\varepsilon}|$$

See Schivardi et al. (2020) for a proof in a different context.

$$C(G, \bar{G}) = (\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})^2 + (\beta^2 + \hat{\kappa})(\bar{\psi}_G - \bar{\psi}_{\bar{G}})^2 - 2\beta(\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})(\bar{\psi}_G - \bar{\psi}_{\bar{G}}). \quad (9)$$

Term $(\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})$ captures economic conflict, term $(\bar{\psi}_G - \bar{\psi}_{\bar{G}})$ captures cultural conflict. The weight attached to the latter increases with the importance of culture for the valuation of public spending, β , and with the relative importance of cultural policy $\hat{\kappa} \equiv \kappa\varphi$. In turn, $\hat{\kappa}$ increases in the welfare weight κ of cultural policy and in tax distortions φ . If φ rises, everyone prefers lower taxes, reducing the salience of redistributive conflict.

Equation (9) is symmetric, $C(G, \bar{G}) = C(\bar{G}, G)$. Thus, if conflict between economic classes is larger than conflict between cultural groups, $C(U, L) > C(SP, SC)$, all voters identify with their economic class. Otherwise, they identify with their cultural group. Define:

$$\hat{\alpha} \equiv \frac{(1 - \beta\rho)^2 - (\beta - \rho)^2 \left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2}{\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 - \rho^2}. \quad (10)$$

We then prove:

Proposition 2 *Everyone identifies with his cultural group if $\hat{\kappa} > \hat{\alpha}$ and with his economic class if $\hat{\kappa} < \hat{\alpha}$. If $\beta \leq \rho \frac{1+\rho^2}{1+\rho^4}$ we have that: (i) $\hat{\alpha} > 0$, (ii) $\hat{\alpha}$ is strictly decreasing in $\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L}$; (iii) $\hat{\alpha}$ is strictly decreasing in ρ if $\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} > 1$.*

If $\hat{\alpha} < 0$ identity is always cultural. If $\hat{\alpha} > 0$ identity can be either economic or cultural depending on parameters. To focus on this case, which is more interesting, we impose the sufficient condition $\beta \leq \rho \frac{1+\rho^2}{1+\rho^4}$.

Identity is shaped by three forces. First, cultural identity is more likely to dominate when the importance of cultural policy relative to taxes, $\hat{\kappa}$, increases. Higher welfare weight κ of cultural policy, triggered for instance by a large inflow of immigrants or by episodes of minority discrimination, brings cultural issues top of mind relative to redistribution, promoting cultural identity. A similar effect is created by higher tax distortions φ , caused for instance by tax competition among countries, which reduce conflict over taxes.

Second, cultural identity is more likely to dominate when differences among cultural groups are large relative to those among economic groups - namely when $\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L}$ is high. By (A1), higher income inequality $\bar{\varepsilon}_U - \bar{\varepsilon}_L$ between classes increases class conflict over all policies, making class identity more likely. Stronger cultural divisions $\bar{\psi}_{SP} - \bar{\psi}_{SC}$ increase cultural conflict and reduce class cohesion, making cultural identity more likely. In an 1893 letter, Friedrich Engels argued that class struggle proved difficult in the US, because of deep ethnic cleavages within the working class. This is consistent with the idea that strong cultural divisions (large $\bar{\psi}_{SP} - \bar{\psi}_{SC}$) hinder class identity. Similarly, expansion of high education

may enhance cultural conflict between progressive elites and the rest of society (i.e. increase $\bar{\psi}_{SP} - \bar{\psi}_{SC}$), promoting cultural identity (Fukuyama 2018).

Third, the correlation ρ between income and progressiveness also plays an important role. Higher ρ promotes cultural identity if $\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} > 1$, otherwise it promotes class identity. The intuition is simple. When ρ is higher, cultural groups disagree more over redistribution and economic classes disagree more over cultural policy. Thus, higher ρ increases policy conflict among both cultural and economic groups. This effect, however, is more pronounced for the groups that are more divided. Cultural groups are more divided than economic classes if and only if $\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} > 1$, so in this case higher ρ promotes cultural identity.

Skill biased technical change is a shock that may increase ρ , because it impoverishes less educated and hence more conservative workers (Autor 2019). Appendix 2 formalizes this argument by assuming that a voter's skill endowment is positively correlated with his progressiveness ψ . We show that, if technical change determines a higher remuneration of skilled labor, income and culture become more positively correlated. This, in turn, may induce voters to switch to cultural identity, even if the shock increases income inequality.

Globalization is another possible driver of higher ρ , if losers from globalization are disproportionately conservative, which may again be due to their lower education. Appendix 2 formalizes this argument by assuming that the labor endowment of progressive voters is more tied to the export sector than that of conservative voters. As globalization determines a higher remuneration of labor in the export sector and a lower one in the import sector, the correlation between income and culture increases, which can again promote cultural identity.¹¹

We mentioned the possibility that voters may identify with narrower groups. Appendix 3 shows that the main message of Proposition 2 holds also if voters can identify both with their economic and cultural group. This means, for instance, that a voter with $\varepsilon \leq \hat{\varepsilon}$ and $\psi \leq \hat{\psi}$ can identify at the same time with the group of lower class and socially conservative people (L, SC). Two main results emerge from this analysis. First, narrow identity is often not chosen, even if available. The intuition is that it tends to reduce conflict relative to broader cultural or economic groups. The narrower group is closer to the voter, but less salient.¹²

The second result is that, even if some voters may choose narrow identity under some conditions, the key qualitative implications of Proposition 2 still hold. In particular, it remains

¹¹As shown in a previous version, trade shocks can induce a switch to cultural identity through a second mechanism. If conservatives are more exposed than progressives to import competition, while opposite income classes are equally exposed, then cultural groups (but not economic classes) also have a conflict over trade policy. Trade shocks can then intensify trade policy conflict between cultural groups, which in turn favors a switch to cultural identity.

¹²To see this, consider conservative and lower class voters in (L, SC). Relative to the lower class L , the narrower (L, SC) features lower income conflict because its outgroup contains some lower class voters, the socially progressive ones. Relative to the conservatives SC , the narrower (L, SC) features lower cultural conflict because its outgroup contains some conservative voters, the upper class ones.

true that an increase in κ promotes cultural identity. This may occur because voters switch from their class to their narrow group, rather than from class to culture. However, because the narrow group is also defined along cultural values, it remains true that culture becomes more important in shaping voters' beliefs. This preserves the gist of our results.

In our model identity switches occur suddenly and for everybody at the same time. This is an artifact of our assumptions. In a more general model where identity reflects a tradeoff between group contrast and similarity between the individual and the group average (as per criterion (i) above), the threshold for identity switches differs across individuals. As a result, the process of identity realignment triggered by growing $\hat{\kappa}$ is more gradual.¹³

3.3 Policy Preferences

How does identity affect policy preferences? As in Section 2, identity distorts beliefs through stereotypes, but beliefs now concern income prospects, $\tilde{\varepsilon}$, and cultural views, $\tilde{\psi}$. This, in turn, affects preferences over τ and q . Let $(\varepsilon_G^\theta, \psi_G^\theta)$ denote the perceived expected income and culture of voter (ε, ψ) if he identifies with group G . Let $\tau_G^{\varepsilon\psi}$ and $q_G^{\varepsilon\psi}$ denote his policy bliss points. Repeating the steps that led to Proposition 1 in Section 2, and using (4), we have:

Proposition 3 *A voter (ε, ψ) identified with group $G = SC, SP, U, L$ perceives his future expected income and his cultural trait to be:*

$$\varepsilon_G^\theta = \varepsilon + \theta (\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}}), \quad (11)$$

$$\psi_G^\theta = \psi + \theta (\bar{\psi}_G - \bar{\psi}_{\bar{G}}). \quad (12)$$

The same voter's ideal policies are given by:

$$\tau_G^{\varepsilon\psi} = \tau^{\varepsilon\psi} + \theta \frac{\beta (\bar{\psi}_G - \bar{\psi}_{\bar{G}}) - (\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})}{\varphi}, \quad (13)$$

$$q_G^{\varepsilon\psi} = q^{\varepsilon\psi} + \theta (\bar{\psi}_G - \bar{\psi}_{\bar{G}}). \quad (14)$$

The voter distorts his perceived future income and his cultural views by the belief difference between ingroups and outgroups. This, in turn, distorts his perception of the ideal

¹³When identity also depends on a voter's similarity to his groups, some voters identify based on culture, others on class. It remains true that higher salience of culture relative to income (higher $\hat{\kappa}$) favors cultural identity. But in this case, extreme voters along income and culture are more likely to change identity when a shock hits. Extreme voters feel intensely about both issues, so they readily switch to the most fitting identity. The formal analysis of that model can be found in a prior draft, available as CEPR Discussion paper 13390. We could also allow for heterogeneity in the propensity to identify with any social group (as if acquiring a social identity entailed a cost). In this case, identification with a group would be triggered only if group conflict is sufficiently large. If unidentified individuals behave rationally, the results remain very similar, as long as at least some voters are always identified in the dimension that maximizes group conflict.

policy. If the voter identifies with a group that is richer than the outgroup, his perceived tax burden is enhanced, otherwise it is dampened. If the voter identifies with a group that is more progressive than the outgroup, his perceived benefit from liberal cultural policies and universalistic redistribution is enhanced, otherwise it is dampened.

Consider a lower class and conservative voter, $\varepsilon < \widehat{\varepsilon}$, $\psi < \widehat{\psi}$. Using the above result and Equations (6) and (7), if this voter identifies with the lower class his bliss points are:

$$\tau_L^{\varepsilon\psi} = \tau^{\varepsilon\psi} + \frac{\theta(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)}{\varphi}, \quad q_L^{\varepsilon\psi} = q^{\varepsilon\psi} - \rho\theta(\bar{\varepsilon}_U - \bar{\varepsilon}_L). \quad (15)$$

Given that $\beta\rho < 1$, the voter exaggerates his demand for redistribution, $\tau_L^{\varepsilon\psi} > \tau^{\varepsilon\psi}$, especially if income inequality among classes is strong. When thinking about himself, the poverty of his social group comes to mind. Like the Marxist proletariat, he feels part of an economically oppressed class, which increases his demand for redistribution. Since $\rho > 0$, the lower class is on average conservative, so the voter also demands a stricter cultural policy $q_L^{\varepsilon\psi} < q^{\varepsilon\psi}$. Conservatism also dampens his demand for redistribution but, as we already saw, this effect is dominated because $\beta\rho < 1$.

Now suppose that cultural policy becomes salient. If the voter switches to conservative identity, his bliss points become:

$$\tau_{SC}^{\varepsilon\psi} = \tau^{\varepsilon\psi} - \frac{\theta(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC})}{\varphi}, \quad q_{SC}^{\varepsilon\psi} = q^{\varepsilon\psi} - \theta(\bar{\psi}_{SP} - \bar{\psi}_{SC}). \quad (16)$$

The voter's conservatism increases, due to (A1), so he is now even more opposed to liberal cultural policy, $q_{SC}^{\varepsilon\psi} < q_L^{\varepsilon\psi}$. He also demands less redistribution than before, $\tau_{SC}^{\varepsilon\psi} < \tau_L^{\varepsilon\psi}$.¹⁴ As the voter abandons class identity, two forces work in this direction. First, the voter is more optimistic about his future income. Second, he becomes more communitarian, which reduces his demand for universalistic redistribution.

This change in preferences may be caused by a shock that is barely related to redistribution, such as a salient conflict on abortion or a large inflow of immigrants from a different culture. Alesina et al. (2018a) find that making people think about immigrants reduces their support for redistribution, particularly if the respondent is less educated and right-wing. They argue that respondents are unwilling to redistribute towards strangers. A different interpretation is that the treatment primes cultural identity in conservative and anti-immigrant subjects. If this is true, the treatment effect should be heterogeneous: conservative subjects should reduce their demand for redistribution while progressives may even increase it, because they are willing to redistribute in favor of immigrants.

¹⁴This is obviously true if $\beta > \rho$ but it is also true if $\beta < \rho$ because, by (A1), $(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L) > (1 - \beta\rho)\rho(\bar{\psi}_{SP} - \bar{\psi}_{SC})$ which is in turn larger than $(\rho - \beta)(\bar{\psi}_{SP} - \bar{\psi}_{SC})$ due to $\rho < 1$.

3.4 Equilibrium Policy

We now study how identity affects equilibrium policy. As in standard models of probabilistic voting, two candidates commit to policy platforms ahead of the elections in order to maximize the probability of winning (cf. Persson and Tabellini 2000). We assume that all voters have the same degree of mobility across parties, so the equilibrium policy maximizes *perceived* utilitarian welfare. Let $W_d^{\varepsilon\psi}(\tau, q)$ denote the perceived expected utility of agent (ε, ψ) if he identifies based on dimension d , for $d = \tilde{\varepsilon}, \tilde{\psi}$. The equilibrium policy is defined by:

$$(\tau^*, q^*) = \arg \max_{\tau, q} \int W_d^{\varepsilon\psi}(\tau, q) dH(\psi, \varepsilon), \quad \text{for } d = \tilde{\varepsilon}, \tilde{\psi} \quad (17)$$

Suppose that everyone identifies based on class. Exploiting Proposition 3, the first order conditions of the problem imply:

$$\tau^* = \tau^\circ + \theta \frac{(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)(\pi_L - \pi_U)}{\varphi}, \quad (18)$$

$$q^* = q^\circ - \theta\rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L)(\pi_L - \pi_U), \quad (19)$$

where τ° and q° denote the socially optimal policies, and π_L, π_U denote the size of the upper and the lower class, respectively.

If $\theta = 0$, beliefs are rational and identity has no effect on equilibrium policy. When instead $\theta > 0$ identity matters. If $\hat{\varepsilon} > 0$, the lower class is the larger economic group, $\pi_L > \pi_U$. Thus, taxes are too high and, if $\rho > 0$, cultural policy is too conservative. These distortions increase with class inequality $(\bar{\varepsilon}_U - \bar{\varepsilon}_L)$. In Marxist theory, class consciousness is necessary for the poor majority to succeed. In a similar way, class identity causes lower class voters to be more radical. Opportunistic politicians then accommodate their demands because the lower class is larger. If the lower class is more conservative, cultural policy is also too restrictive, for the same reason.

Suppose now that everyone identifies based on culture. Then we have:

$$\tau^* = \tau^\circ - \theta \frac{(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC})(\pi_{SC} - \pi_{SP})}{\varphi}, \quad (20)$$

$$q^* = q^\circ - \theta(\bar{\psi}_{SP} - \bar{\psi}_{SC})(\pi_{SC} - \pi_{SP}). \quad (21)$$

Here too, if $\theta > 0$ the stereotypes of the larger group affect the equilibrium. If $\hat{\psi} > 0$, the conservative group is larger, $\pi_{SC} > \pi_{SP}$. Thus, cultural policy is too restrictive ($q^* < q^\circ$). If the average conservative demands lower spending ($\beta > \rho$) taxes are too low, otherwise they are too high.

A switch from economic to cultural identity impacts (τ, q) in a way that depends on: i)

the relative severity of economic vs. cultural conflict; ii) the relative size of different groups. If $\hat{\varepsilon} = \hat{\psi} > 0$ income inequality and cultural conflict are commensurate, $\bar{\psi}_{SP} - \bar{\psi}_{SC} = \bar{\varepsilon}_U - \bar{\varepsilon}_L$ and the dominant groups have the same size $\pi_L = \pi_{SC} > \pi_U = \pi_{SP}$. In this case, a switch from class to culture always causes cultural policy to become more conservative and taxes to fall (the latter due to the fact that $1 - \beta\rho > \beta - \rho$).

We summarize this discussion in the following:

Proposition 4 *Suppose that $\hat{\varepsilon}, \hat{\psi} > 0$.*

- i) Under class identity, $\tau^* > \tau^\circ$ and $q^* \leq q^\circ$, with strict inequality if $\rho > 0$.*
- ii) Under cultural identity, $q^* < q^\circ$, and $\tau^* \lesseqgtr \tau^\circ$ as $\beta \gtrless \rho$*
- iii) If in addition $\hat{\varepsilon} = \hat{\psi}$, a switch from class to cultural identity reduces both τ^* and q^* .*

In our model candidates are opportunistic and follow changing voters' preferences. Suppose instead that candidates are partisan and have different policy platforms, as in Alesina (1987), with the right-wing candidate being culturally and fiscally conservative, and the opposite for the left-wing candidate. Then a shift from class to culture reshuffles voters across parties. Conservative and lower class voters, that voted for the left on the basis of their class, are now attracted by the culturally conservative right-wing party. Progressive and upper class voters are instead attracted by the culturally progressive left-wing party. Piketty (2018) shows that after the 1960s similar mobility patterns have occurred in several Western democracies.

4 Empirical Predictions

In this section we illustrate the key predictions of our model. As shown in Figure 1, after 2008 disagreement over cultural policy has increased, disagreement over redistribution has remained constant or has declined, and views over cultural policy and redistribution have become more correlated. An identity switch from class to culture produces these effects. Let $\tau_d^{\varepsilon\psi}, q_d^{\varepsilon\psi}$ denote the bliss points of voter (ε, ψ) if he identifies on dimension $d = \tilde{\varepsilon}, \tilde{\psi}$.

Proposition 5 *Suppose that $\hat{\varepsilon} = \hat{\psi} > 0$. Then, an increase in κ triggering an identity switch from class to culture causes the following effects:*

- 1) the variance of ideal cultural policies increases while the variance of ideal tax rates decreases: $Var\left(q_{\tilde{\psi}}^{\varepsilon\psi}\right) > Var\left(q_{\tilde{\varepsilon}}^{\varepsilon\psi}\right)$ and $Var\left(\tau_{\tilde{\psi}}^{\varepsilon\psi}\right) < Var\left(\tau_{\tilde{\varepsilon}}^{\varepsilon\psi}\right)$.*
- 2) the correlation coefficient between $\tau_d^{\varepsilon\psi}$ and $q_d^{\varepsilon\psi}$ increases. Under cultural identity this correlation coefficient is positive, but it may be negative under class identity.*

When identity switches to culture, conflict over q intensifies because cultural beliefs polarize. Conflict over τ is subject to two effects. On the one hand, increased cultural polarization

implies more disagreement over the perceived benefits of public spending. On the other hand, dampening of class conflicts reduces disagreement over the perceived tax burden. The second effect is stronger than the first if $\hat{\varepsilon} = \hat{\psi}$.¹⁵ Thus, disagreement over τ falls.

The switch to cultural identity also causes the cultural factor driving both q and τ to become more potent. Conservative voters increase their opposition to immigration and civil rights and become less favorable to redistribution. Progressive voters move in the opposite direction. As a result, the correlation between the demand for redistribution and that for open cultural policies increases.

Proposition 5 shows that an increase in the salience κ of cultural policy can parsimoniously account for Figure 1. A version of our model with rational voters cannot do the same, because higher κ plays no role when $\theta = 0$. More generally, in the rational model no single shock can account for the facts, but a combination of shocks is needed. We return to this point in Section 5, where we discuss other possible explanations of Figure 1.

The preference changes of Proposition 5 are due to changing conflict among underlying social groups. When identity is class-based, conflict is primarily among economic groups. When it is culture-based, conflict is primarily among cultural groups. To see this in detail, let τ_d^G, q_d^G denote the average bliss points of members of group G if they identify on dimension $d = \tilde{\varepsilon}, \tilde{\psi}$.

Proposition 6 *Suppose that $\hat{\varepsilon} = \hat{\psi} > 0$. Then, an increase in κ causing identity to switch from class to culture affects the policy preferences of different social groups as follows.*

(i) *Relative to the conservatives, the progressives become even more in favor of liberal cultural policies; if $\beta > \rho$, they also become even more in favor of high taxes:*

$$q_{\tilde{\psi}}^{SP} - q_{\tilde{\psi}}^{SC} > q_{\tilde{\varepsilon}}^{SP} - q_{\tilde{\varepsilon}}^{SC} > 0 \quad (22)$$

$$\tau_{\tilde{\psi}}^{SP} - \tau_{\tilde{\psi}}^{SC} > \tau_{\tilde{\varepsilon}}^{SP} - \tau_{\tilde{\varepsilon}}^{SC} > 0 \quad \text{if } \beta > \rho. \quad (23)$$

(ii) *Relative to the upper class, the lower class becomes even more opposed to liberal cultural policies; it also becomes less favorable to high taxes:*

$$q_{\tilde{\psi}}^L - q_{\tilde{\psi}}^U < q_{\tilde{\varepsilon}}^L - q_{\tilde{\varepsilon}}^U < 0 \quad (24)$$

$$\tau_{\tilde{\psi}}^L - \tau_{\tilde{\psi}}^U < \tau_{\tilde{\varepsilon}}^L - \tau_{\tilde{\varepsilon}}^U. \quad (25)$$

As voters switch to cultural identity, their cultural traits become exaggerated. Across cultural groups, the progressives become even more progressive and vice-versa for the conservatives. As a result, their conflict over cultural policy intensifies. Conflict over τ between

¹⁵In general, as shown in the proof, the second effect dominates, provided the socially conservative groups is sufficiently large relative to the lower class, namely provided $\hat{\psi}$ is not too small relative to $\hat{\varepsilon}$.

cultural groups, $\tau_d^{SP} - \tau_d^{SC}$, also increases, provided that the influence of culture on redistributive preferences (i.e. β) is strong enough.

The opposite happens to perceived income differences between classes: lower class voters feel less poor and upper class voters less rich. Hence, their assessment of the tax burden becomes less divergent. In addition, the conservative majority of the lower class becomes more communitarian and demands less government spending, and vice-versa for the upper class. For both reasons, the lower class demands lower taxes than before, while the upper class does the opposite. Moreover, the lower class becomes more supportive of conservative cultural policies, compared to the upper class. As a result, conflict over cultural policies between opposite income groups rises.

Going back to Proposition 5, the amplification of conflict over cultural policy among economic and cultural groups explains growing disagreement over q in the population as a whole. The reduction of redistributive conflict among income classes and its increase among cultural groups explains why overall disagreement over τ falls but preferences over τ and q become more correlated over time.

5 Comparing the Theory with the Data

We now present evidence from survey data consistent with the predictions discussed above. The chain of causation in our model is:

salient issue => group identity => beliefs/policy preferences => equilibrium policy.

The effect of an identity switch on the beliefs of voters depends on their economic and cultural traits. In turn, beliefs affect voters' policy preferences. To test this mechanism, ideally one should observe voters' identities, beliefs and policy preferences. Unfortunately we do not observe identities or beliefs. As a result, we cannot rigorously test our mechanism.

Nevertheless, in this section we show that several patterns in the data are broadly consistent with our predictions, and cannot easily be accounted for by competing explanations. To begin, consider Figure 1. In the above chain of causation, the trigger of the trends is an increasing conflict among cultural groups. This could be due to greater importance of immigration and civil rights, κ , or to an economic shock hurting conservative voters (higher ρ). As identity switches from class to culture, voters' beliefs change and, by Propositions 5 and 6, we should observe the following: i) an increase in cultural conflict (measured by the population variance of q and by disagreement over q among cultural groups), ii) a decrease in redistributive conflict (measured by the population variance of τ and by disagreement over τ among classes), iii)

individual views on τ and q becoming more positively correlated in the population.

To see whether Figure 1 could be due to a switch to cultural identity, consider first whether the salience of cultural conflict has increased in the last 20 years. Using data from a repeated Pew survey between 2001 and 2018, Figure 2 plots the share of respondents that report a certain issue as one of the three most important problems facing the US.

Figure 2

The importance of race and immigration has soared from 2012 onwards, which coincides with the time in which the patterns of Figure 1 are most pronounced.¹⁶ In our model, this phenomenon corresponds to a higher welfare importance κ , and hence higher salience of cultural policy.¹⁷

A second natural step is to see whether a growing importance of cultural policies is associated with stronger cultural identity. We do not measure the strength of identification with cultural groups or with income classes. In ANES, however, respondents report "thermometer feelings" on how close they feel to certain groups. We use this information. Respondents are assigned to opposite economic classes based on a question on self-reported social class. They are assigned to opposite cultural groups based on their religiosity, a cultural trait that Enke et al. (2021) have shown to be strongly (negatively) correlated with moral universalism. Appendix 4 provides more details on the definition of social class, religiosity and all other variables used in the analysis.

As a proxy for "affective class polarization", we take the difference between the thermometer of self-classified upper and lower class respondents towards: i) unions and ii) big businesses - two symbols of class identity. As a proxy for "affective cultural polarization", we take the difference between the thermometer of religious and secular respondents towards Christian fundamentalists and Catholics. To remove the confounding effect of identification with political parties, we use the estimated residuals of these feelings, after conditioning on dummies for party affiliation interacted with wave fixed effects (patterns are stronger or similar in the unconditional data and in the sample of political independents, as shown in the Appendix).

¹⁶To highlight long term trends, we omit "the state of the economy" and "unemployment", which follow the business cycle (specially the Great Recession), and national security. Cultural issues were prominent also in the early 2000s. As emphasized by Abramowitz (2019), this is likely to reflect a continuing decline of the share of white voters over the eligible population, due to immigration from Asia and Latin America and a higher fertility of non-white. According to ANES data, (in 2016) nonwhites made up 39% of eligible voters under the age of 30, compared with only 17% of eligible voters over seventy" - Abramowitz (2019, chp 1).

¹⁷In our model there is only one cultural policy q , but the analysis is the same if there are several dimensions of cultural policy q_1, \dots, q_k whose bliss points are highly correlated because they are all driven by a voter's socially progressive or conservative stance ψ . If there is more than one cultural issue, parameter κ captures the average importance of these issues. In this case, a growing importance of several issues at the same time contributes to a strong increase in κ .

Figure 3 plots these indices over time.

Figure 3

In Panel (a), the upper class feels more distant from unions and closer to businesses compared to the lower class, but these class differences in feelings have shrunk, suggesting a reduction in affective class polarization. In Panel (b), religious respondents feel closer to Christian fundamentalists and Catholics than secular respondents, and these differences in feelings have diverged, suggesting a rise in affective cultural polarization. This is consistent with conflict cooling off among economic groups and heating up between religious vs secular groups.

The third step in the above chain of causation is the change in beliefs and policy preferences. We have already shown in Proposition 5 that our model can account for the trends of Figure 1. Here we ask whether these changes in policy preferences are tied to economic and cultural groups, as predicted by Proposition 6. A switch to cultural identity should dampen redistributive conflict among classes and enhance it among cultural groups (if $\beta > \rho$). It should also exacerbate cultural conflict among cultural groups as well as between classes. Figure 4 reports the position over redistribution and cultural policies of respondents sorted into different income classes and cultural groups. The latter are defined again based on respondents' religiosity. Here too we condition on party affiliation interacted with wave fixed effects to control for shifting party positions (the trends are stronger or similar if we do not condition on party affiliation or if we focus on political independents, see the Appendix).

Figure 4.

The trends in group conflict are in line with our predictions. In a robustness exercise, we split respondents into two equally sized groups of "conservatives" and "progressives" based on the first principal component of four questions on the importance of traditional values. The trends of Figure 4 are confirmed (see Figure A7 in the Appendix).

In sum, consistently with our model, a growing salience of cultural conflict has gone hand in hand with a convergence of feelings by social classes towards class-related groups, and with a divergence of feelings by religious/secular groups towards religious people. Moreover, and as predicted by Proposition 6, opposite economic classes now disagree less over redistribution and more over cultural policies, while opposite cultural groups disagree more in both policy domains.

Consider alternative explanations of these facts, starting from a rational version of our model. If $\theta = 0$, the variance of bliss points over q coincides with the variance of culture ψ , σ_ψ . Thus, a rational model must find an explanation for the growing disagreement over

cultural policy in Figure 1. We explain growing cultural disagreement through the growing salience of cultural policy κ , as in Figure 2. In a rational model, the salience of cultural policy has no effect on polarization, although increased salience could be a by-product of higher cultural polarization σ_ψ .

Suppose that we take for granted an increase in cultural disagreement σ_ψ , due for instance to mass education or changes in gender roles (Fukuyama 2018). Is this consistent with the remaining facts? In the rational model, a rise in σ_ψ causes higher correlation between the bliss points for q and τ , accounting for the second fact of Figure 1. The reason is that culture also shapes preferences over the public good through parameter $\beta > 0$. If progressive voters become more extreme, they want a more liberal cultural policy and more government spending, and vice-versa for the conservatives. Hence, views on these policies become more tightly connected.

This intuition, though, suggests that growing cultural disagreement σ_ψ should also cause growing disagreement over redistribution, contrary to what we observe. This is because culturally extreme individuals also hold extreme views on τ . In particular, if preferences over q and τ are positively correlated (like in the ANES data from 2008 onward) and if $\theta = 0$, then an increase in σ_ψ should cause the variance of bliss points over τ to increase. Intuitively, a positive correlation between q and τ means that a voter's culture is an important determinant of his demand for redistribution. In this case, a rise in cultural polarization should also result in higher polarization over τ .¹⁸

Thus, an exogenous increase in cultural disagreement can explain the increased coherence of views over τ and q , but not why disagreement over τ has, if anything, decreased. In our model, a switch from class to cultural identity produces these patterns via two effects. First, it polarizes cultural preferences, which is akin to increasing σ_ψ . Second, it depolarizes class conflict, which is akin to reducing the variance of expected income, σ_ε . It is this second effect that reduces the variance of tax preferences in Proposition 5. Thus, the rational model needs a combination of shocks (to σ_ψ and to σ_ε) to account for the above facts. Our model only needs growing salience of cultural conflict.¹⁹

A second set of explanations for these facts emphasizes the role of political supply. One

¹⁸Let R denote the correlation coefficient between the individual bliss points of q and τ , let σ_q and σ_τ be the standard deviation of these bliss points. Using (4), we have:

$$\begin{aligned} R &= (\beta\sigma_\psi - \rho\sigma_\varepsilon)/\varphi\sigma_\tau \\ \sigma_\tau^2 &= (\beta^2\sigma_\psi^2 + \sigma_\varepsilon^2 - 2\beta\rho\sigma_\psi\sigma_\varepsilon)/\varphi^2 \end{aligned}$$

which in turn imply: $\frac{\partial R}{\partial \sigma_q} = \beta(1 - R^2)/\varphi\sigma_\tau > 0$ and $\frac{\partial \sigma_\tau}{\partial \sigma_q} = \beta R/\varphi > 0$ if $R > 0$.

¹⁹In the rational model, conflict over τ is also affected by other parameters. An increase in tax distortions, φ , or in the correlation ρ between ε and ψ , reduces the variance of ideal tax rates. But these changes also reduce the correlation of views over taxes and cultural policy - recall that income and cultural progressiveness have opposite effects on τ .

version stresses partisan sorting. The Democratic Party has become more liberal, the Republican Party more conservative, causing extreme voters to sort into parties. Thus, partisan voters now exhibit more extreme and correlated policy views (Gentzkow 2016, Klein 2020). Another version allows for persuasion by party leaders. Parties have differentiated their platforms on immigration and civil rights, to attract voters who feel strongly about these issues. Party leaders have then persuaded voters to follow their lead on issues that voters care less about, such as redistribution (Murphy and Shleifer 2004, Glaeser 2005).

In these supply side mechanisms, persuasion should mostly affect partisan voters, but the patterns of Figure 1, Figure 3 and Figure 4 hold after controlling for party affiliation and after focusing on political independents. It is also not obvious how this mechanism can explain increasing cultural disagreement σ_ψ . One possibility, complementary to our model, is that the cultural extremism of political leaders may have increased the salience of cultural conflict, favoring a shift towards cultural identity.

Another distinctive prediction of our model is that economic shocks disproportionately hurting conservative voters should induce a switch to cultural identity. Thus, economic losers should become socially and fiscally conservative. In a rational model, instead, voters hit by economic shocks would demand more redistribution and would not change their views over cultural policy.

Autor et al. (2020) and Colantone and Stanig (2017) show that, both in the US and in Europe, losses from international trade foster support for right-wing and conservative parties. Our model can account for this effect if, as shown in Appendix 2, losers from trade are disproportionately conservative. In this case, our model predicts that losers from trade also become more socially and fiscally conservative in their policy opinions. A previous version of this paper reported evidence from CCES survey data consistent with this prediction.

6 Concluding Remarks

It is often argued that recent years have witnessed a rise in identity politics, intended as the growing importance of conflict over civil rights and minorities (Fukuyama 2018). We take a different perspective: current events underscore the role of cultural identities, but periods of class conflict have their own identity, too. As we suggest, at any given point in time voters perceive their social and political reality from the vantage point of the most salient group among the many latent ones to which they belong. We considered groups defined along economic and cultural traits, but regional, racial, or other groups are also possible. As political cleavages change, voters switch identity from their income class to their cultural, regional, or racial group. These identity switches cause a realignment of voters' beliefs across all issues in which groups systematically differ, including issues that are barely related to the identity

trigger itself.

We explored some key implications of this approach, but much more remains to be done. First and foremost, survey data, possibly combined with priming techniques, should be used to assess the link between identity and beliefs, which we did not explore here. This may also help shed light on the effects of fake or partial news, as well as on the role of digital media, such as Twitter or Facebook, that allow leaders to reach out to voters with emotional and symbolic messages that appeal to their identities.

We have focused on the consequences of a shift from class to cultural identity. But other important episodes in US political history can be interpreted as identity shifts in the opposite direction: from culture to class. One prominent example is the political realignment that took place in the late 1930s and 1940s in the US, and that pushed the Democratic Party to abandon its support for racial discrimination. As emphasized by Schickler (2016), this was the result of a grassroots movement initiated by the core groups that supported the New Deal. Industrial unions, African Americans and urban liberals in the North-East pushed for a fusion of class and race, in a joint defense of labor and civil rights. They did so out of the ideological left-wing conviction that racial division undermined class consciousness, but also because of expediency: the inflow of black immigrants from the South undermined the threat of strikes as black workers could be used as replacement in the workplace. Incorporating the blacks in the working class would remove this threat and strengthen the labor movement.²⁰ In line with this, Calderon et al. (2021) study US counties between 1940-1970 and show that where the inflow of black immigrants from the South was larger, the Democratic Party gained more votes and grassroots activism was strengthened.

We do not allow parties or politicians to shape identities. Yet political and group leaders often play an important role in this, for instance by enhancing the salience of certain groups. Marxist thinkers such as Gramsci stressed the role of the communist party and of intellectuals in fostering class awareness. Nation builders such as Bismarck used nationalism to mobilize support, and the Catholic Church promoted identity politics on the basis of religious values. Political leaders may also create new, party-based identities that supersede traditional social groups. Glaeser (2005) analyses how a leader can mobilize voters by spreading messages of hatred against a minority group. Glaeser et al. (2015) discuss how a party can energize its supporters by taking more extreme positions. We think that our demand side approach may be useful to understand these and other supply side aspects. For instance, it may be easier to identify with politicians that impersonate group stereotypes. This implies that, when polarization is strong, the most successful politicians come from the tails, not the middle.

²⁰In the words of union leader John Brophy, quoted by Schickler (2016), "Behind every lynching is the figure of the labor exploiter, the man in the corporation who would deny labor its fundamental rights". The vote of African Americans was also pivotal in several key districts in the North East, and this too induced the Democratic Party to recognize African Americans as full right members of the working class.

Successful populists will then look similar to the unskilled and unexperienced labor market outsiders that voted them in office (see in particular Dal Bo et al. 2019).

A related set of questions concerns the evolution of party systems. When does a party represent a single identity group, and when does it instead act as an ensemble of heterogeneous social identities? The US Republican party seems to represent those on the right that identify along the income dimension, and the social conservatives that identify on culture, while the Democrats stand for the opposite groups in each dimension. But this has changed at the time of major political realignments. How do party realignments interact with social identities, and how do political and social identities influence each other?

We believe that exploring these issues within the framework of identity theory opens up a new and exciting research agenda.

7 References

Abramowitz, Alan I. (2019), *"The Great Alignment - Race, Party Transformation and the Rise of Donald Trump"*, Yale University Press, New Haven

Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson, and Brendan Price (2016), "Import Competition and the Great US Employment Sag of the 2000s", *Journal of Labor Economics*, 34(S1), S141-S198.

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

Ahler, Douglas (2018), "The Group Theory of Parties: Identity Politics, Party Stereotypes , and Polarization in the 21st Century", *The Forum*, 16 (1): 3-22.

Akerlof, George , and A. Rachel E. Kranton (2000), "Economics and Identity", *The Quarterly Journal of Economics*, Volume 115, Issue 3, 715–753.

Alesina, Alberto (1987), "Macroeconomic Policy in a Two Party System as a Repeated Game", *Quarterly Journal of Economics*, 102, Issue 3, August, Pages 651–678.

Alesina, Alberto, Stefanie Stantcheva, and Armando Miano (2018a), "Immigration and Redistribution", National Bureau of Economic Research, working paper.

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso (2018b), "Intergenerational Mobility and Preferences for Redistribution", *American Economic Review* 108.2 (2018): 521-54.

Autor, David (2019), "Work of the Past, Work of the Future", *American Economic Review Papers and Proceedings*. 109: 1-32.

Autor, David, David Dorn, Gordon Hanson and Kaveh Majlesi (2020), "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure", *American Economic Review*, 110,10.

Bénabou, Roland, and Jean Tirole (2016), “Mindful Economics: The Production, Consumption, and Value of Beliefs”, *Journal of Economic Perspectives*, 30(3), 141-164.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer (2016), “Stereotypes”, *The Quarterly Journal of Economics*, 1753–1794.

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer (2019), “Beliefs about Gender”, *American Economic Review*, 109 (3): 739-73.

Bordalo, Pedro, Marco Tabellini and David Yang (2020), "Stereotypes and Politics", National Bureau of Economic Research, working paper.

Calderon, Alvaro, Vasiliki Fouka and Marco Tabellini (2021), "Racial Diversity, Electoral Preferences, and the Supply of Policy: the Great Migration and Civil Rights", Harvard University, mimeo.

Card, David, Christian Dustmann, and Preston Ian (2012), "Immigration, Wages, and Compositional Amenities", *Journal of the European Economic Association*, 10, 78-119.

Colantone, Italo and Piero Stanig (2017), "The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe", *American Journal of Political Science*, forthcoming.

Cruces, Guillermo, Ricardo Perez-Truglia, and Martin Tetaz (2013), "Biased Perceptions of Income Distribution and Preferences for Redistribution: Evidence from a Survey Experiment", *Journal of Public Economics*, 98, 100-112.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne (2019), “Economic Losers and Political Winners: Sweden’s Radical Right”, *mimeo*.

Druckman, James N., and Matthew S. Levendusky (2019), "What do we measure when we measure affective polarization?.", *Public Opinion Quarterly* 83.1: 114-122.

Enke, Ben (2020), “Moral Values and Voting”, *Journal of Political Economy*, 128, no. 10.

Enke, Ben, Ricardo Rodriguez-Padilla and Florian Zimmerman (2020a), "Moral Universalism and the Structure of Ideology", National Bureau of Economic Research, working paper.

Enke, Ben, Ricardo Rodriguez-Padilla and Florian Zimmerman (2021), "Moral Universalism: Measurement and Economic Relevance", *Management Science*, forthcoming.

Evans, Geoffrey and Jonathan Mellon (2016), "Social Class - Identity, awareness and political attitudes: why are we still working class", *British Social Attitudes* 33, NatCen Social Research 33rd Annual Report.

Festinger, Leo (1950), “Informal Social Communication”, *Psychological Review*, 57, pp. 271- 282.

Flynn, D. J., Brendan Nyhan and Jason Reifler (2017), "The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs about Politics", *Advances in Political Psychology*, 38, 1, pp. 127-150.

Frank, Thomas (2004), *What’s the Matter with Kansas? - How Conservatives Won the*

Heart of America", Picador.

Fukuyama, Francis (2018), "*Identity - The Demand for Dignity and the Politics of Resentment*", Farrar, Strauss and Giroux, New York

Gennaioli, Nicola, and Andrei Shleifer (2018), "*A Crisis of Beliefs (Investor Psychology and Financial Fragility)*", Princeton University Press.

Gentzkow, Matthew (2016), "Polarization in 2016", *Toulouse Network for Information Technology Whitepaper*, 1-23.

Gimpelson, Vladimir, and Daniel Treisman (2018), "Misperceiving inequality", *Economics & Politics*, Vol. 30 (1), 27-54.

Glaeser, Edward, (2005), "The Political Economy of Hatred", *Quarterly Journal of Economics*, Volume 120, Issue 1, February, Pages 45–86

Glaeser, Edward L., Giacomo A. M. Ponzetto, and Jesse M. Shapiro (2005), "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values," *Quarterly Journal of Economics*, 120:4, 1283–330.

Goodhart, David (2017), "*The Road to Somewhere - The New Tribes Shaping British Politics*", *Oxford University Press*.

Grigorieff, Alexis, Christopher Roth, and Diego Ubfal (2020), "Does information change attitudes towards immigrants? Representative evidence from survey experiments" *Demography* 57, 1117–1143.

Grossman, Gene M., and Elhanan Helpman (2020), "Identity Politics and Trade Policy", *The Review of Economic Studies*.

Guriev, Sergei, and Elias Papaioannou (2021), "The Political Economy of Populism", *Journal of Economic Literature*, forthcoming.

Haidt, Jonathan. (2012) *The righteous mind: Why good people are divided by politics and religion*. *Vintage*.

Han, J. and Wackman, D. B. (2017), "Partisan Self-Stereotyping: Testing the Salience Hypothesis in a Prediction of Political Polarization", *International Journal of Communication*, 11, 603–625.

Haslam, Alexander S., and John C. Turner (1992), "Context-Dependent Variation in Social Sterotyping 2: the Relationship between Frame of Reference, Self-Categorization and Accentuation", *European Journal of Social Psychology*, 22(3).

Hinkle, Steve, and Rupert Brown (1990), "Intergroup Comparisons and Social Identity: Some Links and Lacunae". In Abrams, D., and M. Hogg (Eds.), *Social Identity Theory: Constructive and Critical Advances*. New York: Springer Verlag.

Hogg, Michael A., and Dominic Abrams (1998), *Social Identification – A Social Psychology of Intergroup Relations and Group Processes*, Routledge, London.

Janky, Bela, (2018), "Salience of Political Identity and Social Attitudes", mimeo, Center

for Social Sciences, Hungarian Academy of Sciences.

Johnston, Christopher D., Howard G. Lavine, Christopher M. Federico (2017), *Open versus Closed: Personality, Identity, and the Politics of Redistribution*, Cambridge University Press

Kahan, Dan M. (2015), "Climate-Science Communication and the Measurement Problem", *Advances in Pol. Psych.*, 36, 1-43.

Key, V.O. (1955), "A theory of critical elections", *Journal of Politics* 17:3-18.

Klein, Ezra (2020), *"Why we're Polarized"*, Simon and Schuster.

Kuziemko, Ilyana, Michael Norton, Emmanuel Saez, and Stefanie Stantcheva (2015), "How Elastic are Preferences for Redistribution? Evidence from Randomized Survey Experiments", *American Economic Review*, 105, 4, 1478-1508.

Mackie, Diane M. (1986), "Social Identification Effects in Group Polarization", *Journal of Personality and Social Psychology*, 84, 932-943.

Mayhew, David (2004), *"Electoral realignments: A Critique of An American Genre"*, Yale University Press, New Heaven

McGarty, Craig, John C. Turner, Micheal A. Hogg, Barbara David, and Margaret S. Wetherell (1992), "Group Polarization as Conformity to the Prototypical Group Member", *British Journal of Social Psychology*, 31(1).

Murphy, Kevin M., and Andrei Shleifer (2004), "Persuasion in Politics", *American Economic Review Papers and Proceedings*, 94(2), 435-439.

Norris, Pippa and Ronald Inglehart (2019), *"Cultural Backlash: Trump, Brexit, and Authoritarian Populism"*, Cambridge University Press.

Oakes, Penelope J. (1987), "The Saliency of Social Categories", In Turner, J., M Hogg, P. Oakes, S. Reicher, and M Wetherell, 1987, *Rediscovering the Social Group: A Self Categorization Theory*, pp. 117-141, Oxford UK: Blackwell.

Persson, Torsten, and Guido Tabellini (2000), *"Political economics – Explaining economic policy"*, MIT Press.

Piketty, Thomas (2018), "Brahmin Left vs Merchant Right: Rising Inequality and the Changing Structure of Political Conflict. Evidence from France, Britain and the US 1948-2017", *WID.world WP 2018/7*.

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

Schickler, Eric (2016), *"Racial Realignment: The Transformation of American Liberalism, 1932-1965"*, Princeton University Press, Princeton.

Schivardi, Fabiano, Enrico Sette, and Guido, Tabellini (2020), "Identifying the Real Effects of Zombie Lending", *The Review of Corporate Finance Studies*, Volume 9, Issue 3, November, 421–429.

Sherif, Muzafer (1936), *"The Psychology of Social Norms"*, New York, Harper.

Stancheva, Stefanie (2020), "Understanding Tax Policy: How do People Reason?", National Bureau of Economic Research, working paper.

Sundquist, James L. (1983), "*Dynamics of the Party System - Alignment and Realignment of Political Parties in the United States* ", The Brookings Institution, Washington DC

Tabellini, Guido (2008), "The scope of cooperation: Values and incentives." *The Quarterly Journal of Economics* 123.3 (2008): 905-950.

Tajfel, Henri, and John Turner (1979), "An Integrative Theory of Intergroup Conflict", in W.G. Austin & S. Worcel (Eds.), *The Social Psychology of Intergroup Relations*, pp. 33-47, Monterey CA: Brooks/Cole.

Tajfel, Henri, and A.L. Wilkes (1963), "Classification and Quantitative Judgment", *British Journal of Psychology*, 54, pp. 101-114.

Tajfel, Henri, M. Billig, R. Bundy, and Claude Flament (1971), "Social Categorization and Intergroup Behavior", *European Journal of Social Psychology*, 1, 149-178.

Turner, John, Michael A. Hogg, Penelope Oakes, Steven D. Reicher, and Margareth S. Wetherell (1987), "*Rediscovering the Social Group: A Self Categorization Theory*", New York, Blackwell.

Westfall , Jacob, Leaf Van Boven, John R. Chambers, and Charles M. Judd (2015), "Perceiving Political Polarization in the United States: Party Identity Strength and Attitude Extremity Exacerbate the Perceived Partisan Divide", *Perspectives on Psychological Science*, Vol. 10(2) 145–158.

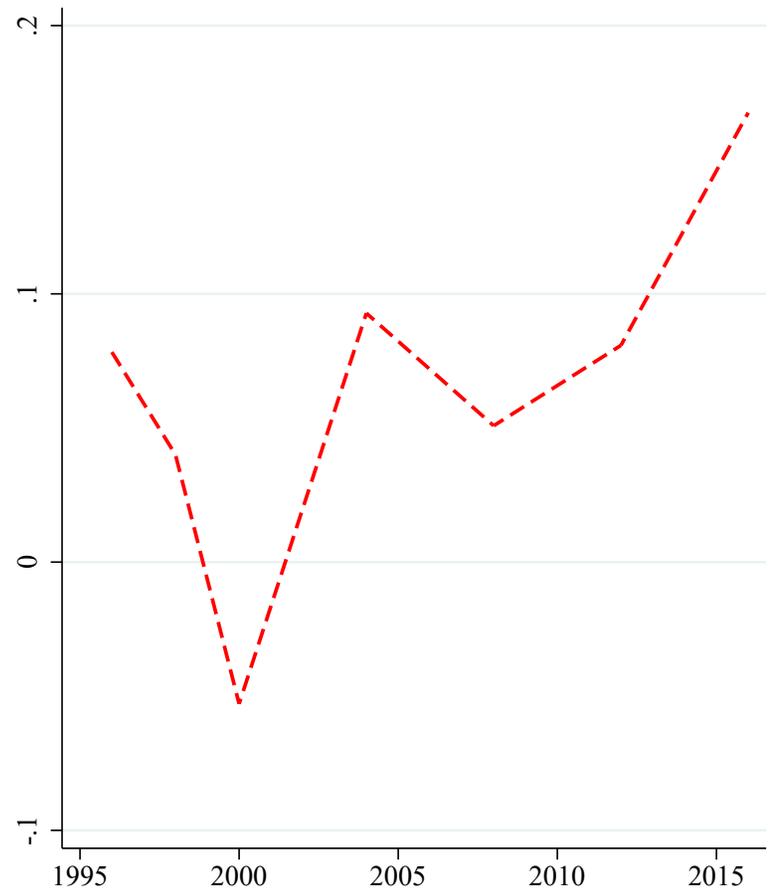
Zaller, John R. (1992), "*The nature and origins of mass opinion*". Cambridge university press.

Wu, Alex. (2020), "The Changing Nature of Political Polarization in the Unites States", Unpublished Undergraduate Thesis, *Harvard University*.

Figure 1: Population Moments



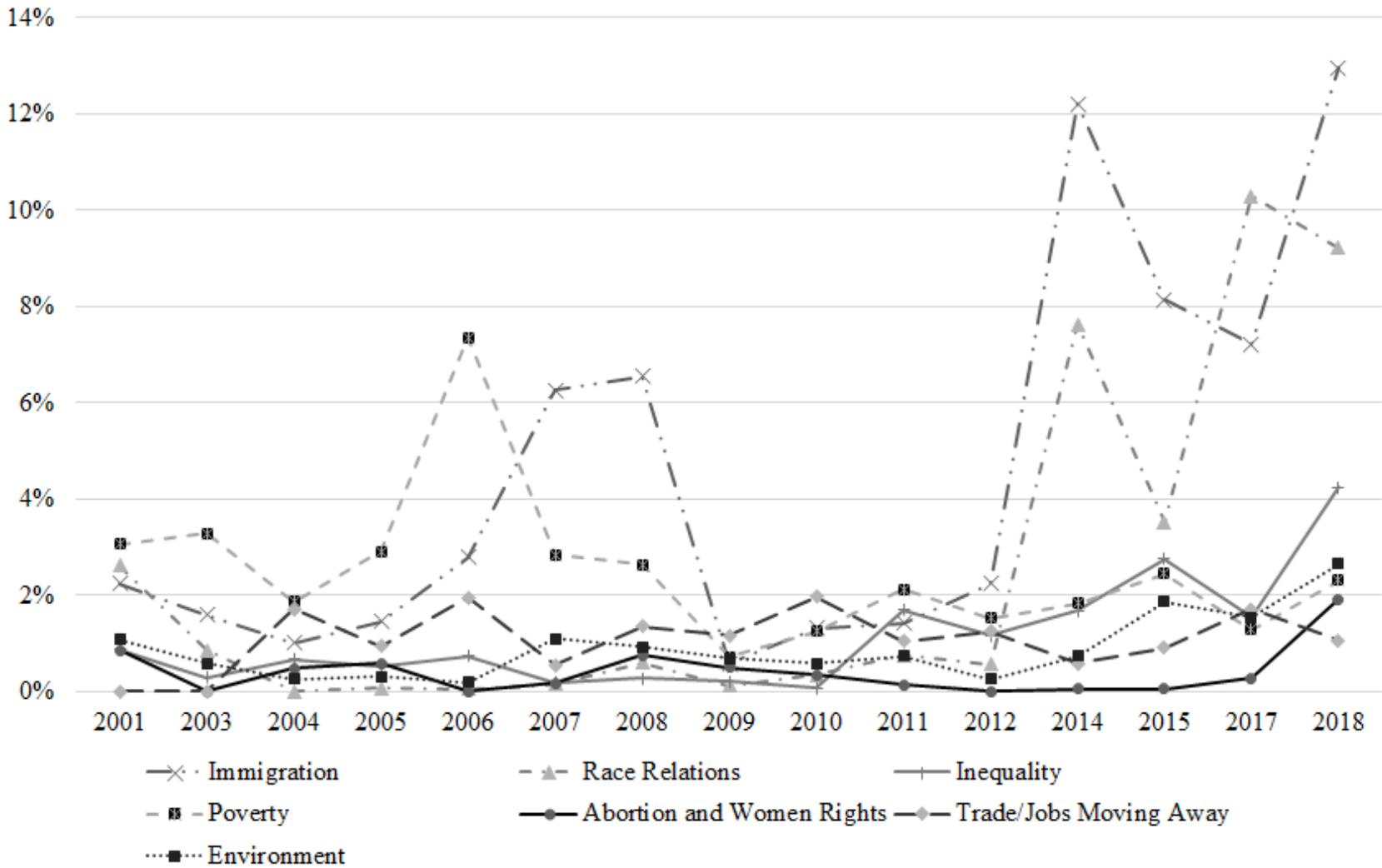
(a) Variance



(b) Correlation Coefficient

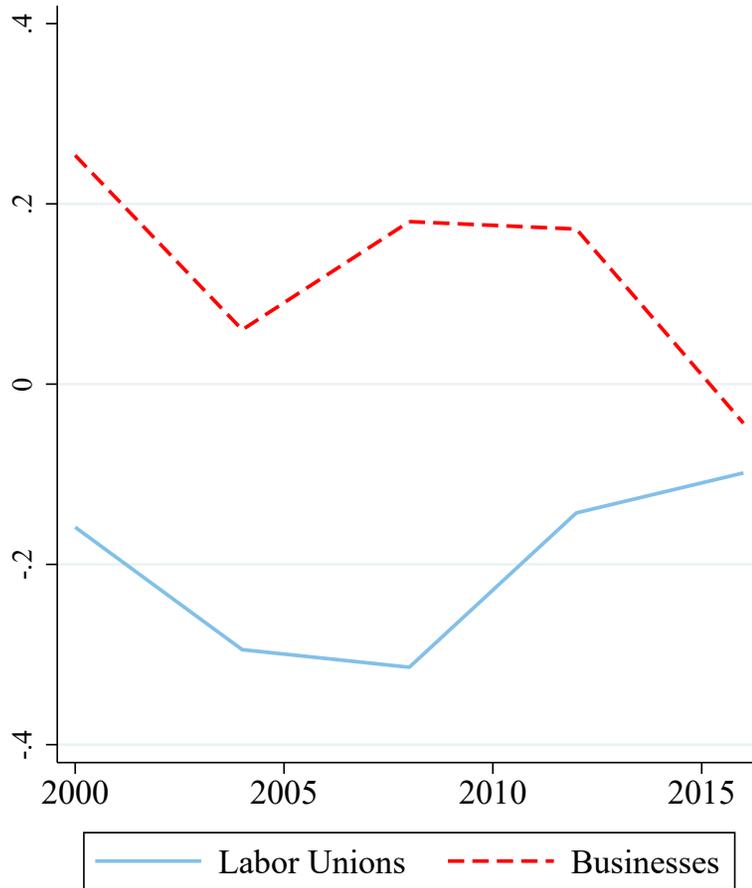
Notes: Redistribution is the first polychoric principal component of two questions on government spending and government's role in seeing to citizens' jobs and living standards; Culture is the first polychoric principal component of desired immigration levels, attitudes towards race relations and abortion policy. For these two measures, higher values correspond to more liberal attitudes. All variables are residuals after conditioning on party identity, alone and interacted with wave fixed effects. Residuals are standardized to have zero mean and unit variance across all waves. Panel (a) reports the variance of each of these two measures. Panel (b) reports the Pearson correlation coefficient for these two measures. Source: ANES.

Figure 2: Most Important Problem Facing The Country

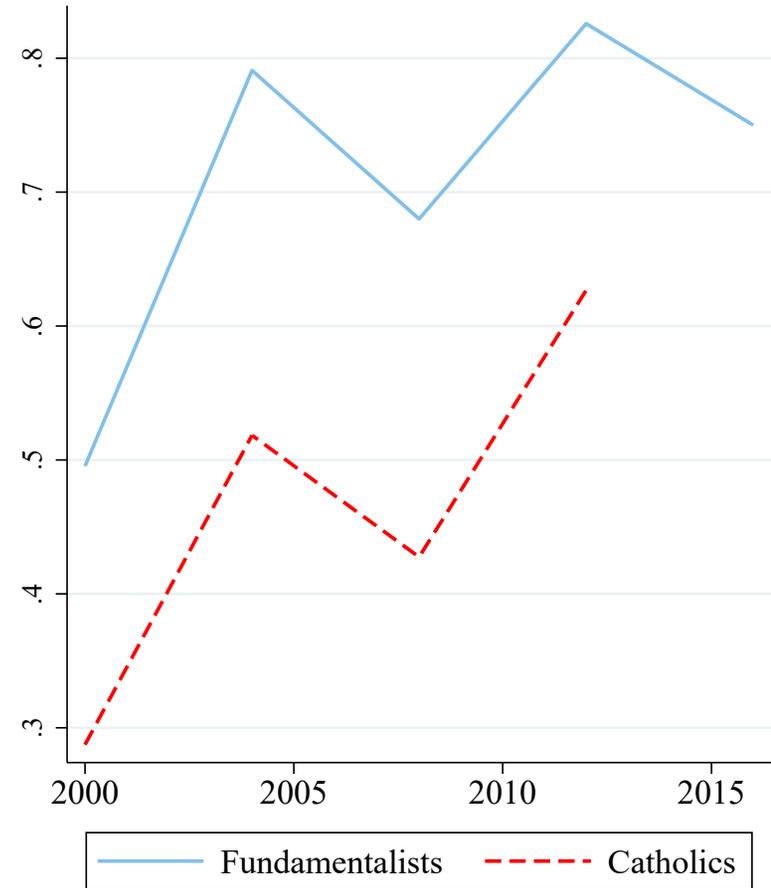


Notes: The graph shows the share of respondents mentioning the selected issues among the top three most important problems facing the US. Source: Pew Research Center public data.

Figure 3: Social Groups and Feeling Thermometer



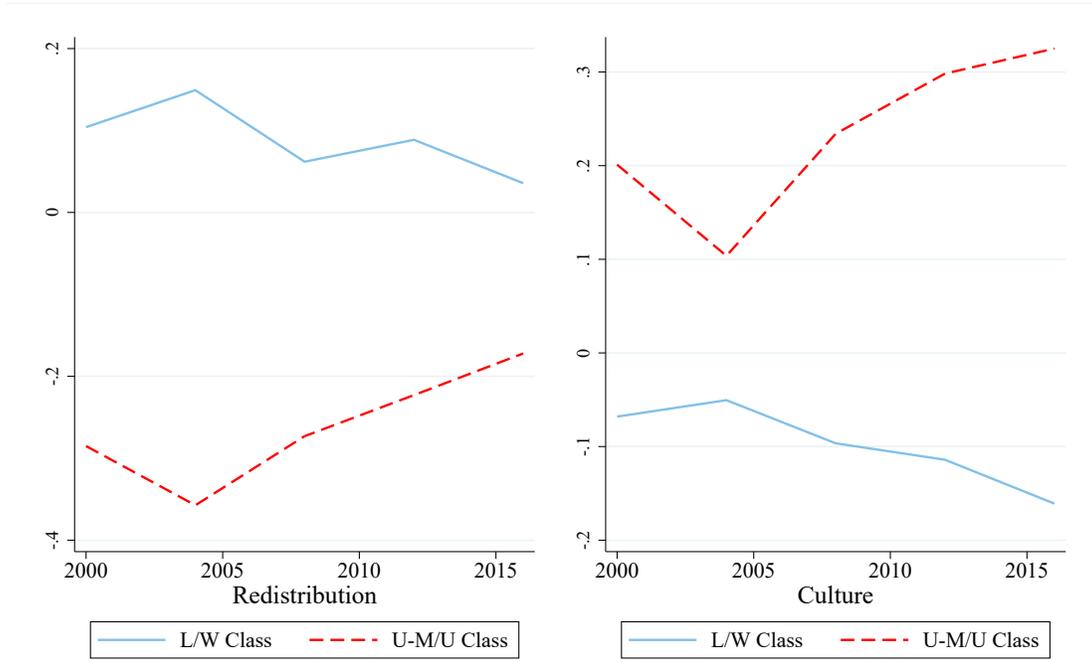
(a) Differences by Class



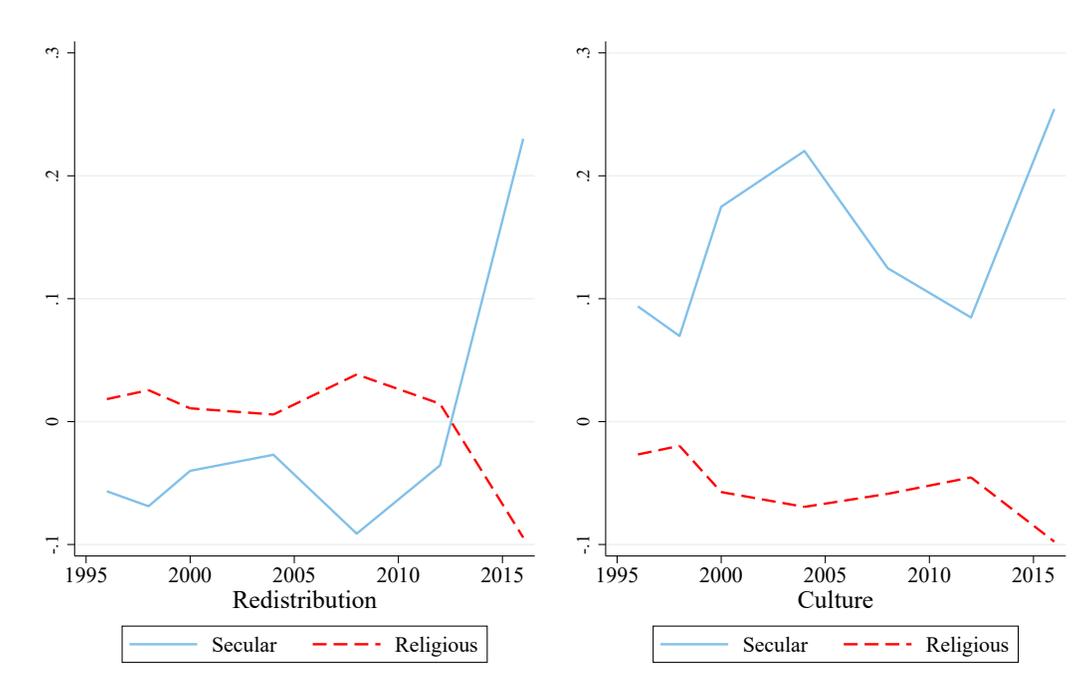
(b) Differences by Religiosity

Notes: Labor Unions, Businesses, Fundamentalists and Catholics are measures of how warm respondents' feelings towards labor unions, big businesses, Christian fundamentalists and Catholics are, respectively (higher values correspond to warmer feelings). For Labor Unions and Businesses, panel (a) plots differences between mean feelings for different social classes (upper-middle/upper class minus lower/working class). For Fundamentalists and Catholics, panel (b) plots differences between mean feelings by religiosity (religious minus secular). All thermometer variables are residuals after conditioning on party identity, alone and interacted with wave fixed effects. Residuals are standardized to have zero mean and unit variance across all waves. Source: ANES.

Figure 4: Trends in Group Conflict



(a) Class



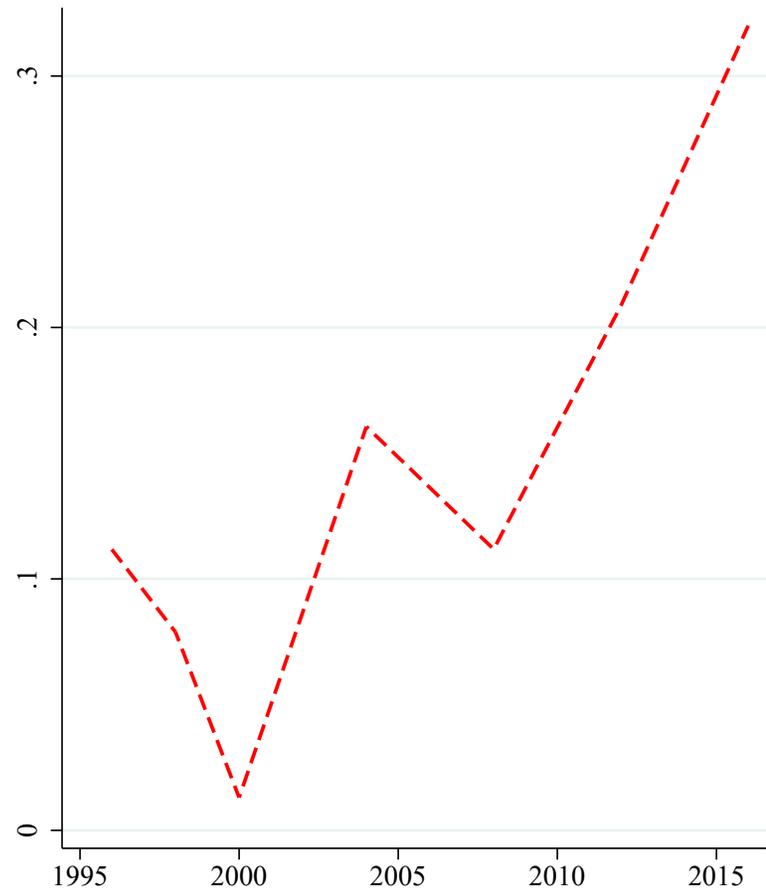
(b) Religiosity

Notes: Redistribution is the first polychoric principal component of two questions on government spending and government's role in seeing to citizens' jobs and living standards; Culture is the first polychoric principal component of desired immigration levels, attitudes towards race relations and abortion policy. For these two measures, higher values correspond to more liberal attitudes. All variables are residuals after conditioning on party identity, alone and interacted with wave fixed effects. Residuals are standardized to have zero mean and unit variance across all waves. For each variable, panel (a) reports trends in means by social class (lower class and upper middle/upper class), panel (b) reports means by religiosity. Source: ANES.

Figure A1: Population Moments (Raw Variables)



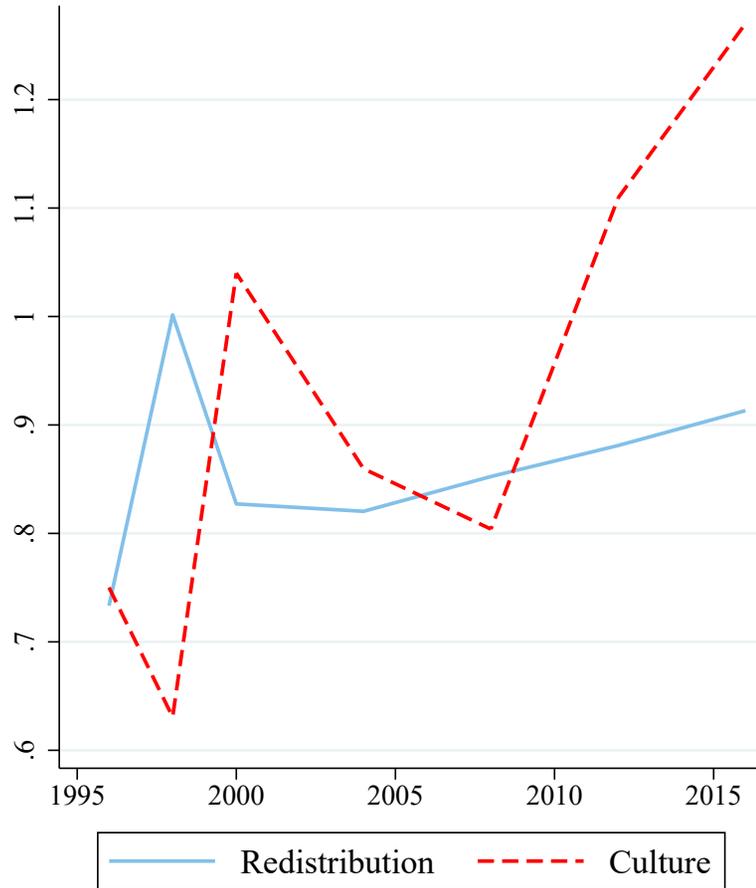
(a) Variance



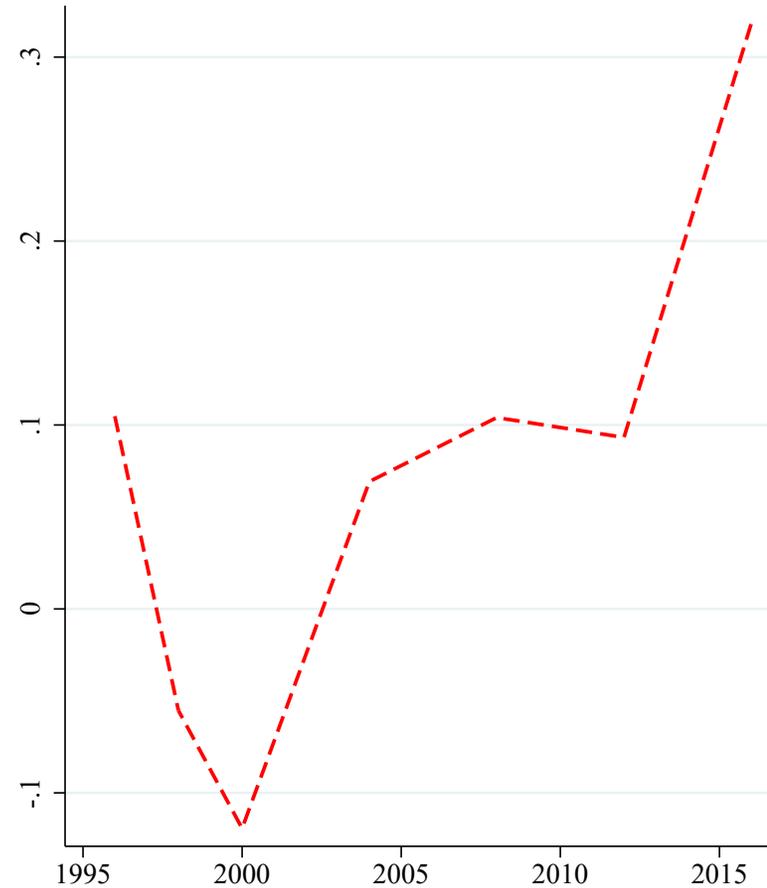
(b) Correlation Coefficient

Notes: Redistribution is the first polychoric principal component of two questions on government spending and government's role in seeing to citizens' jobs and living standards; Culture is the first polychoric principal component of desired immigration levels, attitudes towards race relations and abortion policy. For these two measures, higher values correspond to more liberal attitudes. Both measures are standardized to have zero mean and unit variance across all waves. Panel (a) reports the variance of each of these two measures. Panel (b) reports the Pearson correlation coefficient between these two measures. Source: ANES.

Figure A2: Population Moments (Independents)



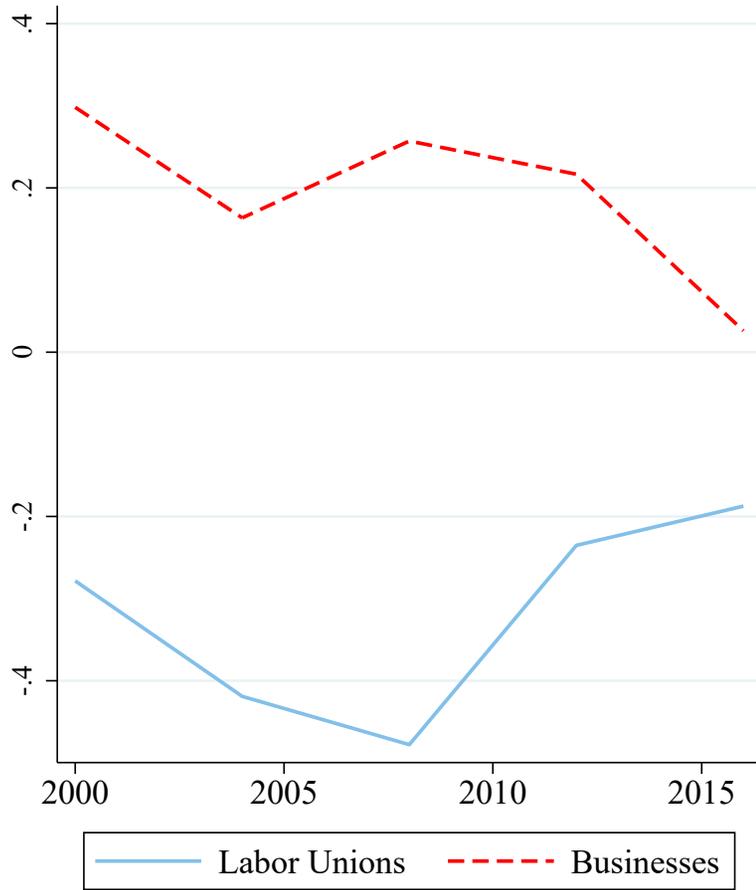
(a) Variance



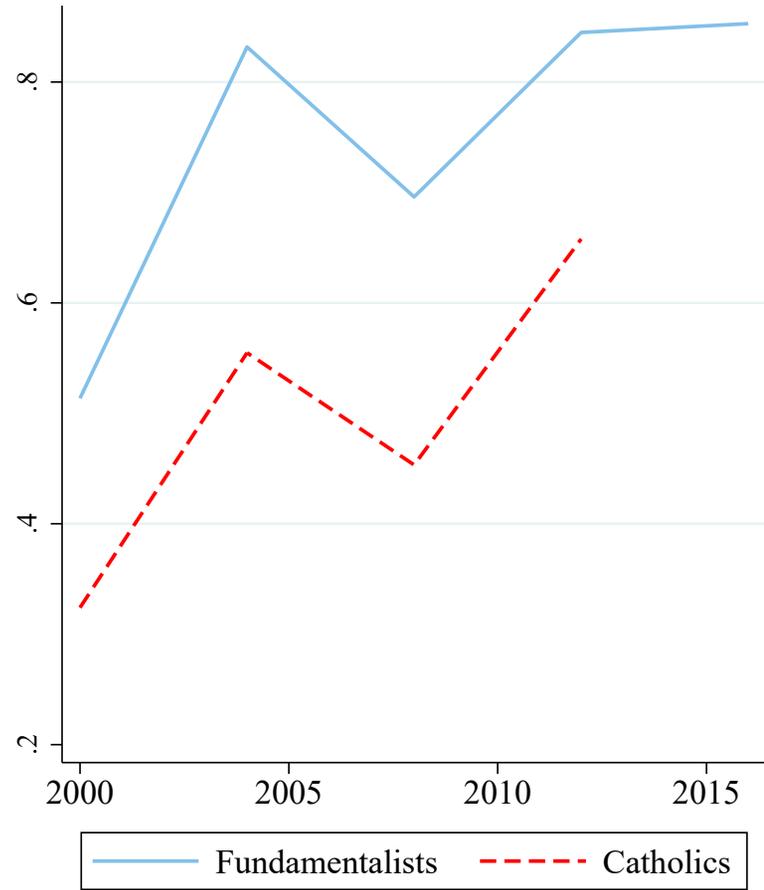
(b) Correlation Coefficient

Notes: Redistribution is the first polychoric principal component of two questions on government spending and government's role in seeing to citizens' jobs and living standards; Culture is the first polychoric principal component of desired immigration levels, attitudes towards race relations and abortion policy. For these two measures, higher values correspond to more liberal attitudes. Both measures are standardized to have zero mean and unit variance across all waves. Panel (a) reports the variance of each of these two measures. Panel (b) reports the Pearson correlation coefficient between these two measures. The sample is restricted to political independents. Source: ANES.

Figure A3: Social Groups and Feeling Thermometer (Raw Variables)



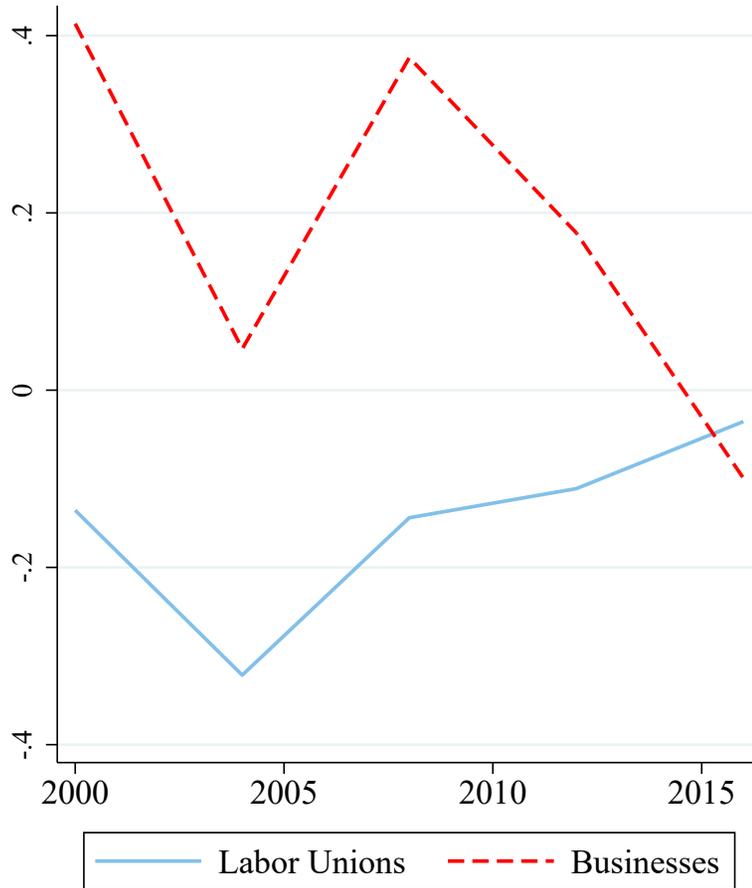
(a) Differences by Class



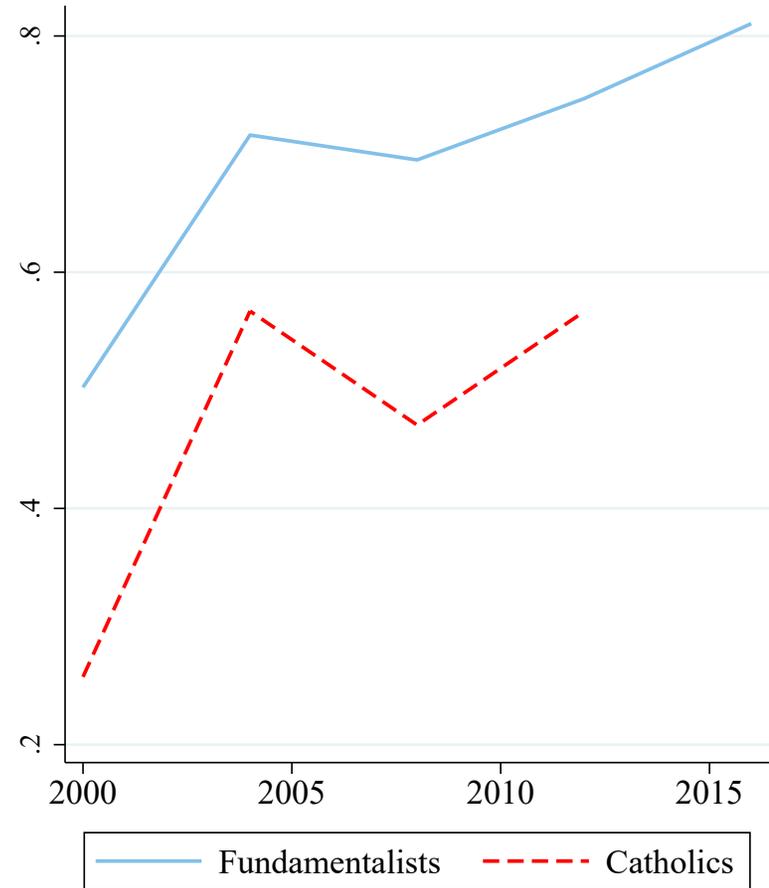
(b) Differences by Religiosity

Notes: Labor Unions, Businesses, Fundamentalists and Catholics are measures of how warm respondents' feelings towards labor unions, big businesses, Christian fundamentalists and Catholics are, respectively (higher values correspond to warmer feelings). All measures are standardized to have zero mean and unit variance across all waves. For Labor Unions and Businesses, panel (a) plots differences between mean feelings for different social classes (upper-middle/upper class minus lower/working class). For Fundamentalists and Catholics, panel (b) plots differences between mean feelings by religiosity (religious minus secular). Source: ANES.

Figure A4: Social Groups and Feeling Thermometer (Independents)



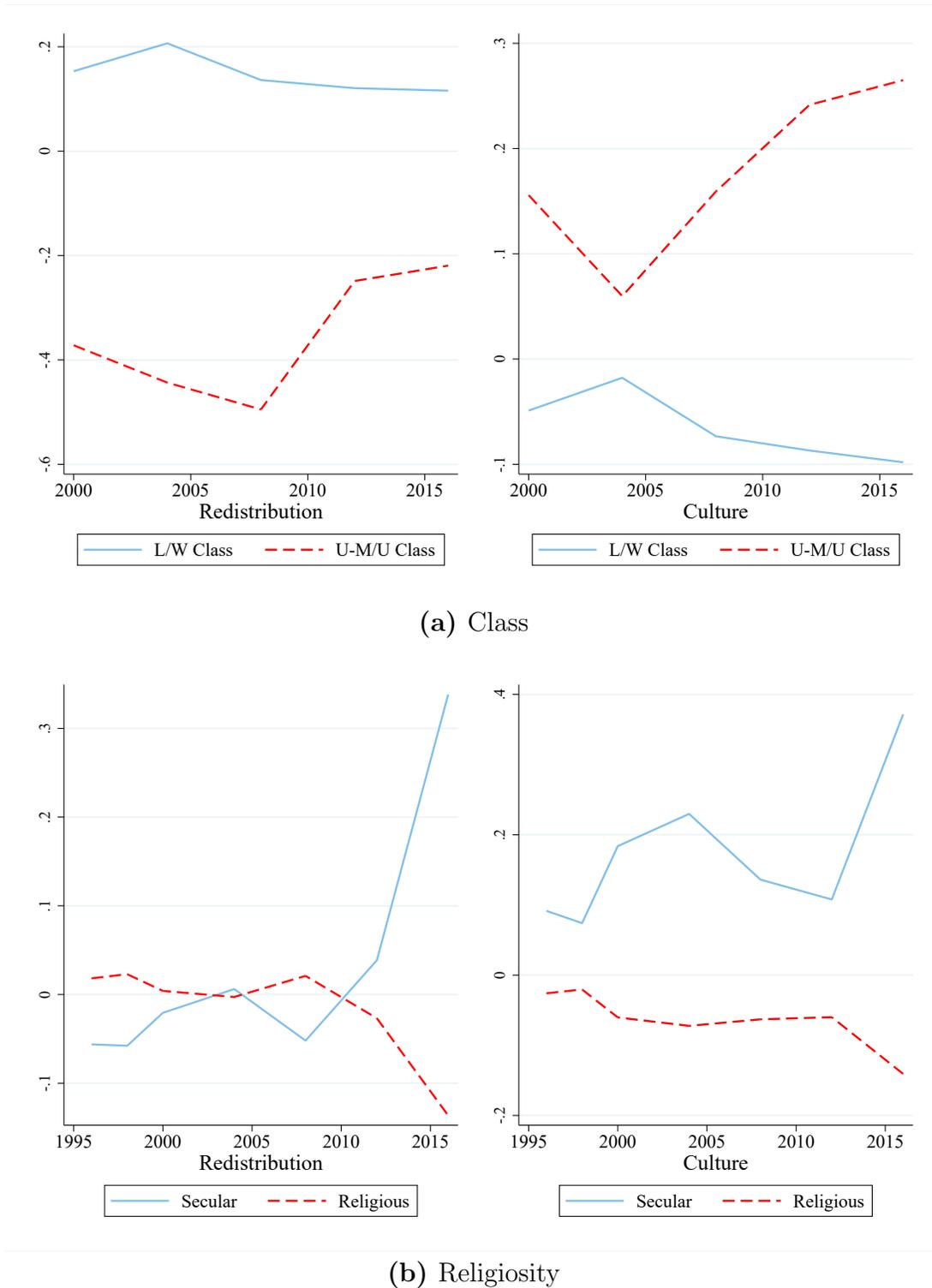
(a) Differences by Class



(b) Differences by Religiosity

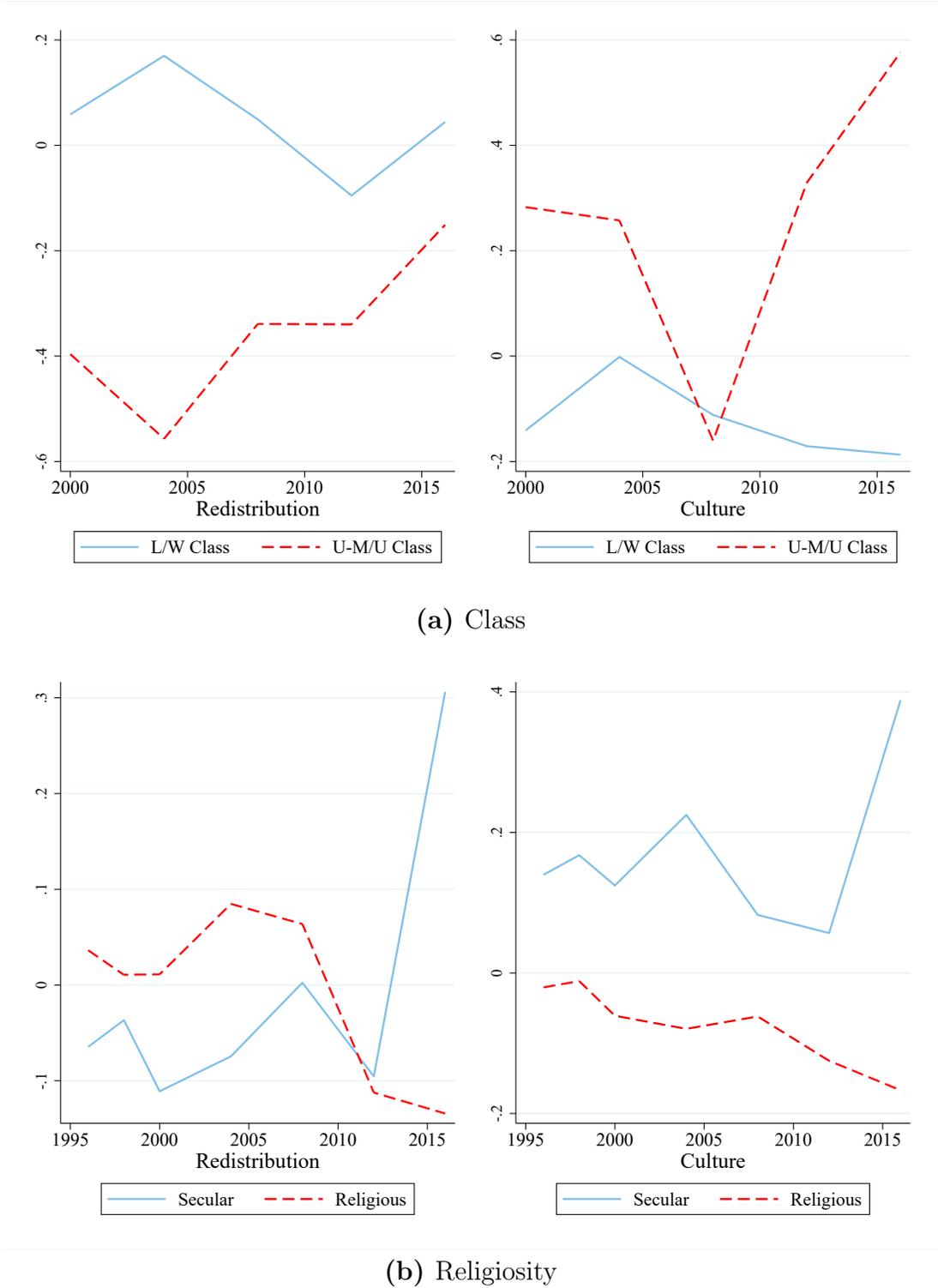
Notes: Labor Unions, Businesses, Fundamentalists and Catholics are measures of how warm respondents' feelings towards labor unions, big businesses, Christian fundamentalists and Catholics are, respectively (higher values correspond to warmer feelings). All measures are standardized to have zero mean and unit variance across all waves. For Labor Unions and Businesses, panel (a) plots differences between mean feelings for different social classes (upper-middle/upper class minus lower/working class). For Fundamentalists and Catholics, panel (b) plots differences between mean feelings by religiosity (religious minus secular). The sample is restricted to political independents. Source: ANES.

Figure A5: Trends in Group Conflict (Raw Variables)



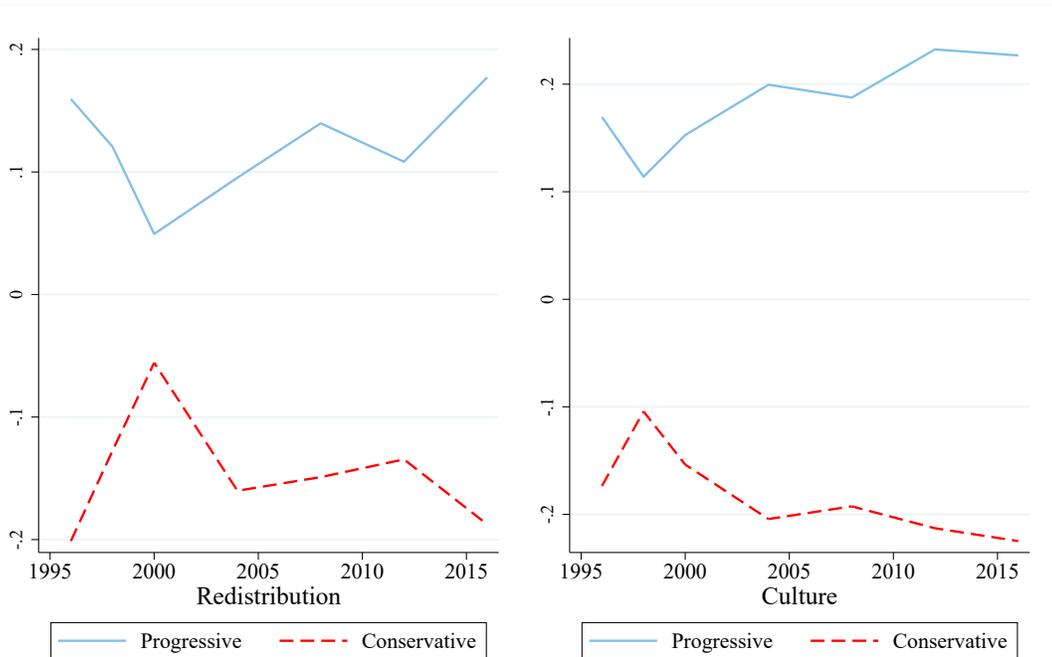
Notes: Redistribution is the first polychoric principal component of two questions on government spending and government's role in seeing to citizens' jobs and living standards; Culture is the first polychoric principal component of desired immigration levels, attitudes towards race relations and abortion policy. For these two measures, higher values correspond to more liberal attitudes. All variables are residuals after conditioning on wave fixed effects. Residuals are standardized to have zero mean and unit variance across all waves. For each variable, panel (a) reports trends in means by social class (lower class and upper middle/upper class), panel (b) reports means by religiosity. Source: ANES.

Figure A6: Trends in Group Conflict (Independents)

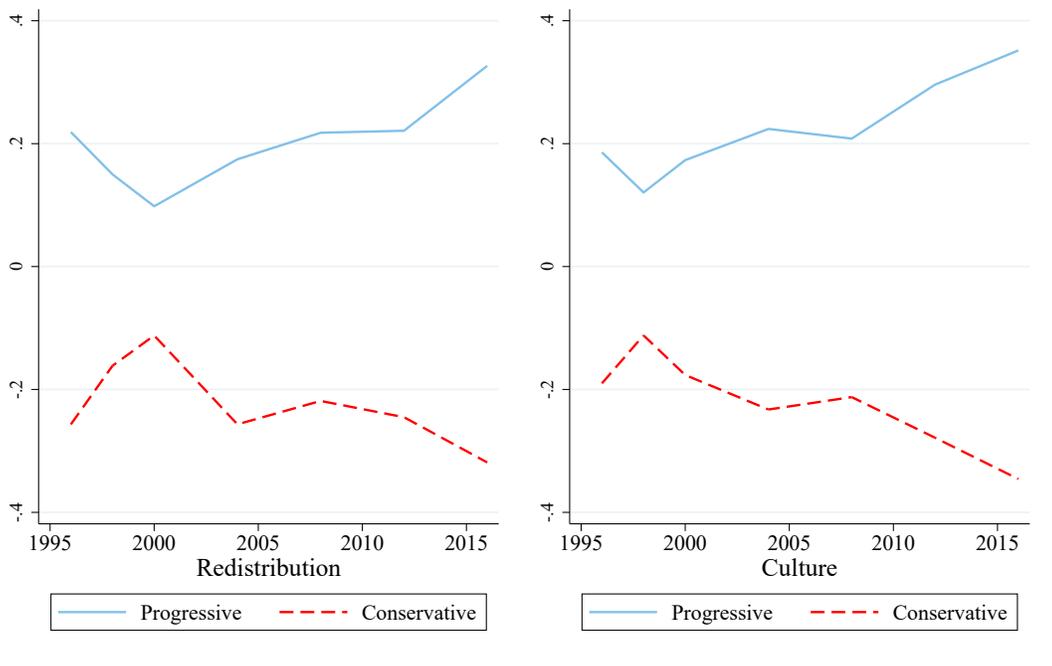


Notes: Redistribution is the first polychoric principal component of two questions on government spending and government’s role in seeing to citizens’ jobs and living standards; Culture is the first polychoric principal component of desired immigration levels, attitudes towards race relations and abortion policy. For these two measures, higher values correspond to more liberal attitudes. All variables are residuals after conditioning on wave fixed effects. Residuals are standardized to have zero mean and unit variance across all waves. For each variable, panel (a) reports trends in means by social class (lower class and upper middle/upper class), panel (b) reports means by religiosity. The sample is restricted to political independents. Source: ANES.

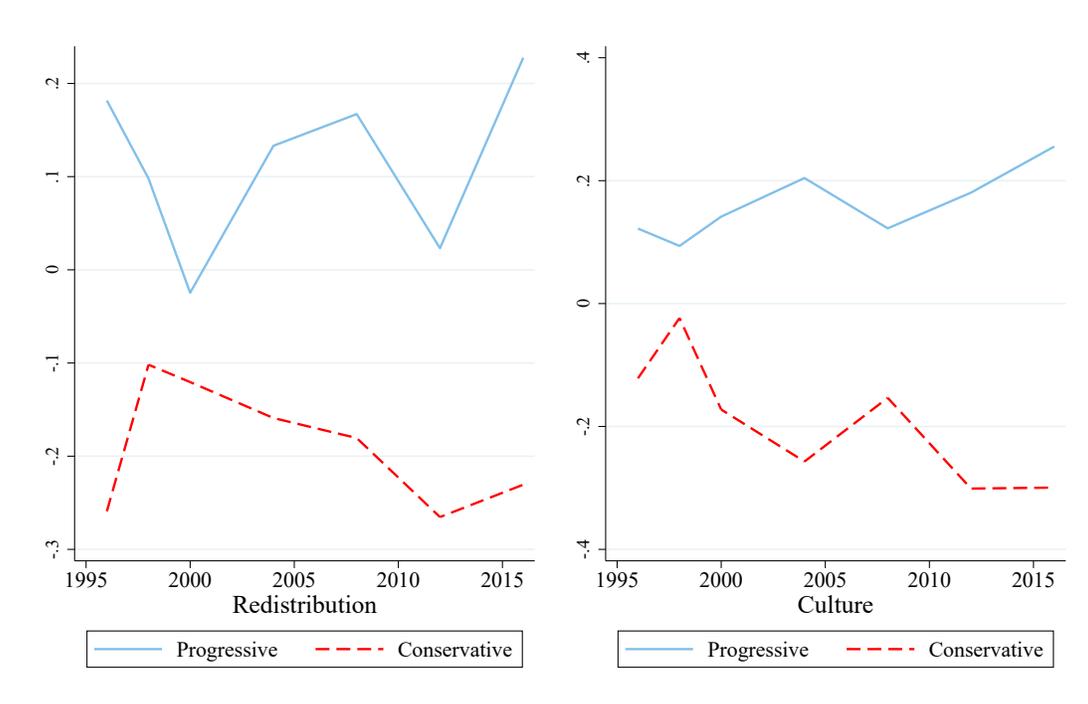
Figure A7: Trends in Group Conflict (Traditionalism)



(a) Residualized



(b) Raw Variables



(c) Independents

Notes: Redistribution is the first polychoric principal component of two questions on government spending and government's role in seeing to citizens' jobs and living standards; Culture is the first polychoric principal component of desired immigration levels, attitudes towards race relations and abortion policy. For these two measures, higher values correspond to more liberal attitudes. Traditionalism is the first polychoric principal component of questions on the value and importance of traditional values, with higher values corresponding to more traditional views. For each year, we classify as Conservative (Progressive) those scoring above (below) the median of Traditionalism in that year. In all three panels, we report trends in means for Progressive and Conservative respondents separately. In panel (a) Redistribution and Culture are residuals after conditioning on the interaction between party and wave fixed effects. In panel (b) and (c) both measures are residuals after conditioning on wave fixed effects. In panel (c) the sample is restricted to political independents. Residuals are standardized to have zero mean and unit variance across all waves. Source: ANES.

Appendix 1: Proofs

Proof of Proposition 1. Using (1), the distorted likelihood ratio between average group members is:

$$\frac{z^\theta \left(\tilde{\psi} \mid \bar{\psi}_G, G \right)}{z^\theta \left(\tilde{\psi} \mid \bar{\psi}_{\bar{G}}, \bar{G} \right)} = \frac{Z_G z \left(\tilde{\psi} \mid \bar{\psi}_G \right)}{Z_{\bar{G}} z \left(\tilde{\psi} \mid \bar{\psi}_{\bar{G}} \right)} \left[\frac{z^\theta \left(\tilde{\psi} \mid \bar{\psi}_G, G \right)}{z^\theta \left(\tilde{\psi} \mid \bar{\psi}_{\bar{G}}, \bar{G} \right)} \right]^{2\chi}, \quad (26)$$

where Z_G and $Z_{\bar{G}}$ are positive normalization constants, and where the equation defines a fixed point condition $x = f(x)$. There is only one positive fixed point, which is stable provided $f(x)$ is concave. This is ensured by $2\chi < 1$, in which case, there also exist two constants Z_G and $Z_{\bar{G}}$ such that the belief distributions $z^\theta \left(\tilde{\psi} \mid \bar{\psi}_G, G \right)$ and $z^\theta \left(\tilde{\psi} \mid \bar{\psi}_{\bar{G}}, \bar{G} \right)$ integrate to one. Then, Equation (1) becomes:

$$z^\theta \left(\tilde{\psi} \mid \psi, G \right) = Z_G \cdot z \left(\tilde{\psi} \mid \psi \right) \left[\frac{z \left(\tilde{\psi} \mid \bar{\psi}_G \right)}{z \left(\tilde{\psi} \mid \bar{\psi}_{\bar{G}} \right)} \right]^{\frac{\chi}{1-2\chi}},$$

and an equivalent expression for \bar{G} . Which is the BCGS (2016) equation with $\theta \equiv \frac{\chi}{1-2\chi}$. With Gaussian distributions this yields:

$$\psi_G^\theta \equiv \int \tilde{\psi} z^\theta \left(\tilde{\psi} \mid \psi, G \right) d\tilde{\psi} = \psi + \theta \left(\bar{\psi}_G - \bar{\psi}_{\bar{G}} \right).$$

■

Proof of Corollaries 1 and 2. The disagreement among average group types is equal to:

$$\bar{\psi}_{SP}^\theta - \bar{\psi}_{SC}^\theta = \bar{\psi}_{SP} + \theta \left(\bar{\psi}_{SP} - \bar{\psi}_{SC} \right) - \bar{\psi}_{SC} - \theta \left(\bar{\psi}_{SP} - \bar{\psi}_{SC} \right) = \left(\bar{\psi}_{SP} - \bar{\psi}_{SC} \right) \left(1 + 2\theta \right),$$

which proves Corollary 1. Denote by $\tilde{\psi}_G^\theta$ the random variable capturing distorted beliefs of the average member of G . Using Equation (1), the distorted perception of such average member is equal to:

$$z^\theta \left(\tilde{\psi}_G^\theta \right) = z^\theta \left(\tilde{\psi} \mid \bar{\psi}_G, G \right) \left[\frac{z^\theta \left(\tilde{\psi} \mid \bar{\psi}_G, G \right)}{z^\theta \left(\tilde{\psi} \mid \bar{\psi}_{\bar{G}}, \bar{G} \right)} \right]^\chi,$$

so that, given again Gaussianity, the perceived position of the average member of G is equal to:

$$\hat{\psi}_{SP}^\theta = \bar{\psi}_{SP}^\theta + \left(\frac{\theta}{1 + 2\theta} \right) \left(\bar{\psi}_{SP}^\theta - \bar{\psi}_{SC}^\theta \right),$$

where we exploited the definition $\chi \equiv \frac{\theta}{1+2\theta}$. This immediately implies that:

$$\widehat{\psi}_{SP}^{\theta} - \widehat{\psi}_{SC}^{\theta} = \left(\overline{\psi}_{SP}^{\theta} - \overline{\psi}_{SC}^{\theta} \right) \left(1 + \frac{2\theta}{1+2\theta} \right) = \left(\overline{\psi}_{SP}^{\theta} - \overline{\psi}_{SC}^{\theta} \right) \left(\frac{1+4\theta}{1+2\theta} \right).$$

■

Proof for the Conflict Function of Section 3. Recall that:

$$W^{\varepsilon\psi}(\tau, q) = (1 + \varepsilon)(1 - \tau) - \frac{\varphi}{2}\tau^2 + (\nu + \beta\psi)\tau - \frac{\kappa}{2}(q - \psi)^2,$$

which can be written as:

$$W^{\varepsilon\psi}(\tau, q) \propto A_{\varepsilon\psi} - (\tau - \tau^{\varepsilon\psi})^2 - \frac{\kappa}{\varphi}(q - \psi)^2,$$

where $\tau^{\varepsilon\psi}$ is the voter's bliss point and $A_{\varepsilon\psi}$ is a voter-dependent constant. This implies that:

$$W^G(\tau^G, q^G) - W^G(\tau^{\overline{G}}, q^{\overline{G}}) \propto (\tau^{\overline{G}} - \tau^G)^2 + \frac{\kappa}{\varphi}(\overline{\psi}_{\overline{G}} - \overline{\psi}_G)^2$$

Plugging in the expression for bliss points we get:

$$W^G(\tau^G, q^G) - W^G(\tau^{\overline{G}}, q^{\overline{G}}) \propto \frac{\beta^2}{\varphi^2}(\overline{\psi}_{\overline{G}} - \overline{\psi}_G)^2 - 2\frac{\beta}{\varphi^2}(\overline{\psi}_G - \overline{\psi}_{\overline{G}})(\overline{\varepsilon}_G - \overline{\varepsilon}_{\overline{G}}) + \frac{(\overline{\varepsilon}_G - \overline{\varepsilon}_{\overline{G}})^2}{\varphi^2} + \frac{\kappa}{\varphi}(\overline{\psi}_{\overline{G}} - \overline{\psi}_G)^2,$$

which, by collecting φ^2 in the denominator yields:

$$W^G(\tau^G, q^G) - W^G(\tau^{\overline{G}}, q^{\overline{G}}) \propto (\overline{\varepsilon}_G - \overline{\varepsilon}_{\overline{G}})^2 + (\beta^2 + \kappa\varphi)(\overline{\psi}_{\overline{G}} - \overline{\psi}_G)^2 - 2\beta(\overline{\varepsilon}_G - \overline{\varepsilon}_{\overline{G}})(\overline{\psi}_G - \overline{\psi}_{\overline{G}}),$$

which is our contrast function $C(G, \overline{G})$ in Equation (9) under the definition $\widehat{\kappa} \equiv \kappa\varphi$. ■

Proof of Proposition 2. All voters (ε, ψ) identify with their cultural group if and only if cultural contrast is larger than income contrast, $C(SP, SC) \geq C(U, L)$, which reads:

$$\begin{aligned} (\overline{\varepsilon}_{SP} - \overline{\varepsilon}_{SC})^2 + (\beta^2 + \widehat{\kappa})(\overline{\psi}_{SP} - \overline{\psi}_{SC})^2 - 2\beta(\overline{\varepsilon}_{SP} - \overline{\varepsilon}_{SC})(\overline{\psi}_{SP} - \overline{\psi}_{SC}) &\geq \\ (\overline{\varepsilon}_U - \overline{\varepsilon}_L)^2 + (\beta^2 + \widehat{\kappa})(\overline{\psi}_U - \overline{\psi}_L)^2 - 2\beta(\overline{\varepsilon}_U - \overline{\varepsilon}_L)(\overline{\psi}_U - \overline{\psi}_L). \end{aligned}$$

Exploiting correlations, the condition becomes:

$$(\overline{\psi}_{SP} - \overline{\psi}_{SC})^2 (\rho^2 + \beta^2 + \widehat{\kappa} - 2\beta\rho) \geq (\overline{\varepsilon}_U - \overline{\varepsilon}_L)^2 (1 + \beta^2\rho^2 + \widehat{\kappa}\rho^2 - 2\beta\rho),$$

which is equivalent to:

$$\widehat{\kappa} \left[\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 - \rho^2 \right] \geq (1 - \beta\rho)^2 - (\beta - \rho)^2 \left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2.$$

The left hand side is positive, for (A1) implies $\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 > \rho^2$. Thus, cultural identity prevails iff:

$$\widehat{\kappa} \geq \widehat{\alpha} \equiv \frac{(1 - \beta\rho)^2 - (\beta - \rho)^2 \left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2}{\left[\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 - \rho^2 \right]}. \quad (27)$$

If the numerator of $\widehat{\alpha}$ is negative, identity is always cultural. This occurs when:

$$\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 \geq \left(\frac{1 - \beta\rho}{\beta - \rho} \right)^2. \quad (28)$$

For $\beta < \rho$, inequality (28) is not met, for (A.1) implies $\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 < 1/\rho^2$. For $\beta > \rho$, (28) cannot also be met provided:

$$\beta \leq \beta^{**} \equiv \rho \frac{2}{1 + \rho^2}.$$

Thus, for $\beta \leq \beta^{**}$, $\widehat{\alpha} > 0$. For $\beta > \beta^{**}$, $\widehat{\alpha}$ is negative when (28) is met. We return to this issue later.

Consider now property ii). regardless of whether $\beta \leq \beta^{**}$ or $\beta > \beta^{**}$, inspection of $\widehat{\alpha}$ and of (28) immediately yield it. Consider property iii), note that after some algebra we can write:

$$\frac{\partial \widehat{\alpha}}{\partial \rho} \propto \left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^4 (\beta - \rho) - \left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 \beta (1 - \rho^2) + \rho (1 - \beta\rho).$$

By decomposing $\beta(1 - \rho^2)$ as $(\beta - \rho) + \rho(1 - \beta\rho)$, this can be factorized as:

$$\frac{\partial \widehat{\alpha}}{\partial \rho} \propto \left[\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 - 1 \right] \left[(\beta - \rho) \left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 - \rho(1 - \beta\rho) \right]. \quad (29)$$

Consider the conditions under which $\frac{\partial \widehat{\alpha}}{\partial \rho} < 0$. If $\beta < \rho$, the second term in square brackets is negative so $\frac{\partial \widehat{\alpha}}{\partial \rho} < 0$ if and only if $\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 > 1$. If $\beta > \rho$, the second term in square brackets is negative if and only if:

$$\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} \right)^2 \leq \vartheta \equiv \rho \left(\frac{1 - \beta\rho}{\beta - \rho} \right).$$

If $\beta < \beta^{**}$, $\vartheta > 1$. In this case, $\frac{\partial \hat{\alpha}}{\partial \rho} < 0$ for $\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L}\right)^2 \in (1, \vartheta]$. Note that if $\vartheta \geq 1/\rho^2$, then by assumption (A.1) $\frac{\partial \hat{\alpha}}{\partial \rho} < 0$ for all admissible $\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L}\right)^2 > 1$. The condition for $\vartheta \geq 1/\rho^2$ is:

$$\beta \leq \beta^* \equiv \rho \left(\frac{1 + \rho^2}{1 + \rho^4} \right) < \beta^{**}.$$

If $\beta > \beta^{**}$, $\vartheta < 1$. In this case, $\frac{\partial \hat{\alpha}}{\partial \rho} < 0$ for $\left(\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L}\right)^2 \in (\vartheta, 1)$.

Assuming $\beta \leq \beta^*$ ensures properties i) and iii) of the proposition. ■

Proof of Proposition 3. Take the expression for ψ_G^θ . Given a Gaussian distribution $f(\tilde{\varepsilon}|\varepsilon)$ we can define, in line with (1), the distorted beliefs over income by a voter identified with G :

$$f^\theta(\tilde{\varepsilon}|\varepsilon, G) = f(\tilde{\varepsilon}|\varepsilon) \left[\frac{f^\theta(\tilde{\varepsilon}|\bar{\varepsilon}_G, G)}{f^\theta(\tilde{\varepsilon}|\bar{\varepsilon}_G, \bar{G})} \right]^x,$$

which, repeating the logic of the fixed point, yields:

$$\varepsilon_G^\theta = \varepsilon + \theta(\bar{\varepsilon}_G - \bar{\varepsilon}_G).$$

The policy demands of a voter (ε, ψ) identified with group G are then given by:

$$\begin{aligned} \tau_G^{\varepsilon\psi} &= \frac{v + \beta\psi_G^\theta - \varepsilon_G^\theta}{\varphi}, \\ q_G^{\varepsilon\psi} &= \psi_G^\theta, \end{aligned}$$

which, by replacing the expressions for beliefs, immediately yields the proposition. ■

Proof of Proposition 4. By inspection of (18)-(21). ■

Proof of Proposition 5. Denote by $(\tau_d^{\varepsilon\psi}, q_d^{\varepsilon\psi})$ the policy demands of (ε, ψ) when he identifies along dimension $d = \tilde{\varepsilon}, \tilde{\psi}$. Neglecting constant terms, under class identity these demands are:

$$\tau_{\tilde{\varepsilon}}^{\varepsilon\psi} = \frac{\beta\psi - \varepsilon}{\varphi} - \theta(2I_U - 1) \frac{(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)}{\varphi}, \quad (30)$$

$$q_{\tilde{\varepsilon}}^{\varepsilon\psi} = \psi + \theta(2I_U - 1)\rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L), \quad (31)$$

where I_U is the indicator of U membership. Given that $var(I_U) = \pi_U(1 - \pi_U)$ and that $E(\varepsilon I_U) = \bar{\varepsilon}_U \pi_U$, and that $E(\psi I_U) = \rho \bar{\varepsilon}_U \pi_U$, and given that, due to $E(\varepsilon) = 0$, we have

$\bar{\varepsilon}_U \pi_U = (\bar{\varepsilon}_U - \bar{\varepsilon}_L) (1 - \pi_U) \pi_U$, the variance covariance matrix is equal to:

$$\begin{aligned} \text{var} \left(\tau_{\hat{\varepsilon}}^{\varepsilon\psi} \right) &= \text{var} \left(\tau^{\varepsilon\psi} \right) + 4\theta (1 + \theta) \frac{(1 - \beta\rho)^2}{\varphi^2} (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U (1 - \pi_U), \\ \text{var} \left(q_{\hat{\varepsilon}}^{\varepsilon\psi} \right) &= \text{var} \left(q^{\varepsilon\psi} \right) + 4\theta (1 + \theta) \rho^2 (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U (1 - \pi_U), \\ \text{cov} \left(\tau_{\hat{\varepsilon}}^{\varepsilon\psi}, q_{\hat{\varepsilon}}^{\varepsilon\psi} \right) &= \text{cov} \left(\tau^{\varepsilon\psi}, q^{\varepsilon\psi} \right) - 4\theta (1 + \theta) \left(\frac{1 - \beta\rho}{\varphi} \right) \rho (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U (1 - \pi_U). \end{aligned}$$

where $\pi_U = \Pr(\varepsilon \in U)$.

Neglecting constant terms, under cultural identity, the demands by (ε, ψ) are:

$$\tau_{\hat{\psi}}^{\varepsilon\psi} = \frac{\beta\psi - \varepsilon}{\varphi} + \theta (2I_{SP} - 1) \frac{(\beta - \rho) (\bar{\psi}_{SP} - \bar{\psi}_{SC})}{\varphi}, \quad (32)$$

$$q_{\hat{\psi}}^{\varepsilon\psi} = \psi + \theta (2I_{SP} - 1) (\bar{\psi}_{SP} - \bar{\psi}_{SC}), \quad (33)$$

Where I_{SP} is the indicator of social progressive membership. Given that $\text{var}(I_{SP}) = \pi_{SP} (1 - \pi_{SP})$ and that $E(\varepsilon I_{SP}) = \rho \bar{\psi}_{SP} \pi_{SP}$, and that $E(\psi I_{SP}) = \bar{\psi}_{SP} \pi_{SP}$, and given that, due to $E(\varepsilon) = 0$, we have that $\bar{\psi}_{SP} \pi_{SP} = (\bar{\psi}_{SP} - \bar{\psi}_{SC}) \pi_{SP} (1 - \pi_{SP})$, the variance covariance matrix is:

$$\begin{aligned} \text{var} \left(\tau_{\hat{\psi}}^{\varepsilon\psi} \right) &= \text{var} \left(\tau^{\varepsilon\psi} \right) + 4\theta (1 + \theta) \frac{(\beta - \rho)^2}{\varphi^2} (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP} (1 - \pi_{SP}), \\ \text{var} \left(q_{\hat{\psi}}^{\varepsilon\psi} \right) &= \text{var} \left(q^{\varepsilon\psi} \right) + 4\theta (1 + \theta) (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP} (1 - \pi_{SP}), \\ \text{cov} \left(\tau_{\hat{\psi}}^{\varepsilon\psi}, q_{\hat{\psi}}^{\varepsilon\psi} \right) &= \text{cov} \left(\tau^{\varepsilon\psi}, q^{\varepsilon\psi} \right) + 4\theta (1 + \theta) \frac{(\beta - \rho)}{\varphi} (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP} (1 - \pi_{SP}). \end{aligned}$$

Suppose that identity switched from class to culture due to an increase in κ . Then, the variance of preferred taxes decreases if and only if:

$$\begin{aligned} \text{var} \left(\tau_{\hat{\psi}}^{\varepsilon\psi} \right) &< \text{var} \left(\tau_{\hat{\varepsilon}}^{\varepsilon\psi} \right) \iff \\ (\beta - \rho)^2 (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP} (1 - \pi_{SP}) &< (1 - \beta\rho)^2 (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U (1 - \pi_U), \end{aligned}$$

which depends, among other things, on $\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L}$ and on (π_U, π_{SP}) . If $\hat{\varepsilon} = \hat{\psi}$, we have $\frac{\bar{\psi}_{SP} - \bar{\psi}_{SC}}{\bar{\varepsilon}_U - \bar{\varepsilon}_L} = 1$ and $\pi_U = \pi_{SP}$. Given that $\beta < 1$, this implies $(\beta - \rho)^2 < (1 - \beta\rho)^2$. So, disagreement over taxes falls when identity switches from class to culture.

The variance of bliss points over q increases provided:

$$\begin{aligned} \text{var} \left(q_{\hat{\psi}}^{\varepsilon\psi} \right) &> \text{var} \left(q_{\hat{\varepsilon}}^{\varepsilon\psi} \right) \iff \\ (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP} (1 - \pi_{SP}) &> \rho^2 (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U (1 - \pi_U). \end{aligned}$$

Note that by A1 $(\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 > (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \rho^2$, so this variance increases if groups are similar in size. In particular, it increases for $\hat{\varepsilon} = \hat{\psi}$. More generally, this variance increases if ρ is low enough.

Finally, consider the correlation between bliss points over τ and q . To begin, we prove that $\frac{\text{cov}(\tau_{\hat{\psi}}^{\varepsilon\psi}, q_{\hat{\psi}}^{\varepsilon\psi})}{\text{var}(q_{\hat{\psi}}^{\varepsilon\psi})} > \frac{\text{cov}(\tau_{\hat{\varepsilon}}^{\varepsilon\psi}, q_{\hat{\varepsilon}}^{\varepsilon\psi})}{\text{var}(q_{\hat{\varepsilon}}^{\varepsilon\psi})}$. This is equivalent to:

$$\frac{\text{cov}(\tau^{\varepsilon\psi}, q^{\varepsilon\psi}) + 4\theta(1+\theta) \frac{(\beta-\rho)}{\varphi} (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP})}{\text{var}(q^{\varepsilon\psi}) + 4\theta(1+\theta) (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP})} > \frac{\text{cov}(\tau^{\varepsilon\psi}, q^{\varepsilon\psi}) - 4\theta(1+\theta) \left(\frac{1-\beta\rho}{\varphi}\right) \rho (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U)}{\text{var}(q^{\varepsilon\psi}) + 4\theta(1+\theta) \rho^2 (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U)}.$$

This is equivalent to:

$$\begin{aligned} & \text{cov}(\tau^{\varepsilon\psi}, q^{\varepsilon\psi}) \rho^2 (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U) + \text{var}(q^{\varepsilon\psi}) \frac{(\beta-\rho)}{\varphi} (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}) \\ & + 4\theta(1+\theta) \frac{(\beta-\rho)}{\varphi} (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}) \rho^2 (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U) > \\ & \text{cov}(\tau^{\varepsilon\psi}, q^{\varepsilon\psi}) (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}) - \text{var}(q^{\varepsilon\psi}) \left(\frac{1-\beta\rho}{\varphi}\right) \rho (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U) \\ & - 4\theta(1+\theta) \left(\frac{1-\beta\rho}{\varphi}\right) \rho (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U) (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}). \end{aligned}$$

This can be written as:

$$\begin{aligned} & \text{var}(q^{\varepsilon\psi}) \left[\frac{(\beta-\rho)}{\varphi} (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}) + \left(\frac{1-\beta\rho}{\varphi}\right) \rho (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U) \right] \\ & + 4\theta(1+\theta) \left(\frac{1-\rho^2}{\varphi}\right) \rho (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}) (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U) > \\ & \text{cov}(\tau^{\varepsilon\psi}, q^{\varepsilon\psi}) \left[(\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}) - \rho^2 (\bar{\varepsilon}_U - \bar{\varepsilon}_L)^2 \pi_U(1-\pi_U) \right]. \end{aligned}$$

When the two groups have equal size, $\hat{\varepsilon} = \hat{\psi}$, the condition simplifies to:

$$\left[\text{var}(q^{\varepsilon\psi}) + 4\theta(1+\theta) (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}) \rho \right] \beta > \varphi \text{cov}(\tau^{\varepsilon\psi}, q^{\varepsilon\psi}),$$

which, after noticing that $\text{cov}(\tau^{\varepsilon\psi}, q^{\varepsilon\psi}) = \beta \text{var}(q^{\varepsilon\psi}) - \rho$, it is equivalent to:

$$4\theta(1+\theta) (\bar{\psi}_{SP} - \bar{\psi}_{SC})^2 \pi_{SP}(1-\pi_{SP}) \rho > -\frac{\rho}{\beta},$$

which is fulfilled. Note that when $\widehat{\varepsilon} = \widehat{\psi}$, it is also the case that the correlation between preferences over redistribution and cultural policy increases with cultural identity, because:

$$\frac{\text{cov}(\tau, q)}{\sqrt{\text{var}(q)\text{var}(\tau)}} = \frac{\text{cov}(\tau, q)}{\text{var}(q)} \frac{\sqrt{\text{var}(q)}}{\sqrt{\text{var}(\tau)}},$$

and we already established that under cultural identity $\text{var}(q)$ increases and $\text{var}(\tau)$ decreases.

■

Proof of Proposition 6. Denote by (τ_d^G, q_d^G) the distorted bliss points of the average member of G when he is identified along $d = \tilde{\varepsilon}, \tilde{\psi}$. Then, Equations (4), (13) and (14) imply:

$$\begin{aligned}\tau_d^G &= \tau^G + \theta \left(\tau^G - \tau^{\bar{G}} \right), \\ q_d^G &= q^G + \theta \left(q^G - q^{\bar{G}} \right).\end{aligned}$$

These in turn imply that policy disagreement among groups with which individuals identify is equal to:

$$\begin{aligned}\tau_d^G - \tau_d^{\bar{G}} &= \left(\tau^G - \tau^{\bar{G}} \right) (1 + 2\theta), \\ q_d^G - q_d^{\bar{G}} &= \left(q^G - q^{\bar{G}} \right) (1 + 2\theta).\end{aligned}$$

Next consider conflict between groups with which voters are not identified. Under class identity, the bliss points of the average social progressive and social conservative are equal to:

$$\begin{aligned}\tau_{\tilde{\varepsilon}}^{SP} &= \tau^{SP} - \theta \frac{(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)}{\varphi} \pi_{U|SP} + \theta \frac{(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)}{\varphi} \pi_{L|SP}, \\ q_{\tilde{\varepsilon}}^{SP} &= q^{SP} + \theta\rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L) \pi_{U|SP} - \theta\rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L) \pi_{L|SP}, \\ \tau_{\tilde{\varepsilon}}^{SC} &= \tau^{SC} - \theta \frac{(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)}{\varphi} \pi_{U|SC} + \theta \frac{(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)}{\varphi} \pi_{L|SC}, \\ q_{\tilde{\varepsilon}}^{SC} &= q^{SC} + \theta\rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L) \pi_{U|SC} - \theta\rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L) \pi_{L|SC},\end{aligned}$$

where $\pi_{X|Y}$ is the share of members of Y that belong to X . So, disagreement among cultural groups under class identity is:

$$\begin{aligned}\tau_{\tilde{\varepsilon}}^{SP} - \tau_{\tilde{\varepsilon}}^{SC} &= (\tau^{SP} - \tau^{SC}) - 2\theta \frac{(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)}{\varphi} (\pi_{U|SP} - \pi_{U|SC}) \\ q_{\tilde{\varepsilon}}^{SP} - q_{\tilde{\varepsilon}}^{SC} &= (q^{SP} - q^{SC}) + 2\theta\rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L) (\pi_{U|SP} - \pi_{U|SC}).\end{aligned}$$

Under cultural identity disagreement among cultural groups in both dimensions is $(1 + 2\theta)$ times the rational disagreement. Thus, a switch from class to culture increases their disagree-

ment over q provided:

$$\begin{aligned} q_{\hat{\psi}}^{SP} - q_{\hat{\psi}}^{SC} &> q_{\hat{\varepsilon}}^{SP} - q_{\hat{\varepsilon}}^{SC} \Leftrightarrow \\ (q^{SP} - q^{SC}) &\equiv (\bar{\psi}_{SP} - \bar{\psi}_{SC}) > \rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L) (\pi_{U|SP} - \pi_{U|SC}), \end{aligned}$$

which is true by (A.1). Consider now disagreement over τ between cultural groups:

$$\begin{aligned} \tau_{\hat{\psi}}^{SP} - \tau_{\hat{\psi}}^{SC} &\geq \tau_{\hat{\varepsilon}}^{SP} - \tau_{\hat{\varepsilon}}^{SC} \Leftrightarrow \\ (\tau^{SP} - \tau^{SC}) &\geq -\frac{(1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L)}{\varphi} (\pi_{U|SP} - \pi_{U|SC}) \Leftrightarrow \\ (\beta - \rho)(\bar{\psi}^{SP} - \bar{\psi}^{SC}) &\geq -(1 - \beta\rho)(\pi_{U|SP} - \pi_{U|SC})(\bar{\varepsilon}_U - \bar{\varepsilon}_L) \end{aligned}$$

Hence, $\tau_{\hat{\psi}}^{SP} - \tau_{\hat{\psi}}^{SC} > \tau_{\hat{\varepsilon}}^{SP} - \tau_{\hat{\varepsilon}}^{SC}$ if $\beta > \rho$, since $\pi_{U|SP} \geq \pi_{U|SC}$ given that $\rho \geq 0$.

Next, consider economic groups. Under class identity, the disagreement over tax rates and cultural policies between the average upper and lower class voters are $(1 + 2\theta)$ times their rational disagreement. Consider now disagreement under cultural identity. Bliss points are:

$$\begin{aligned} \tau_{\hat{\psi}}^U &= \tau^U + \theta \frac{(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC})}{\varphi} \pi_{SP|U} - \theta \frac{(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC})}{\varphi} \pi_{SC|U}, \\ q_{\hat{\psi}}^U &= q^U + \theta (\bar{\psi}_{SP} - \bar{\psi}_{SC}) \pi_{SP|U} - \theta (\bar{\psi}_{SP} - \bar{\psi}_{SC}) \pi_{SC|U}, \\ \tau_{\hat{\psi}}^L &= \tau^L + \theta \frac{(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC})}{\varphi} \pi_{SP|L} - \theta \frac{(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC})}{\varphi} \pi_{SC|L}, \\ q_{\hat{\psi}}^L &= q^L + \theta (\bar{\psi}_{SP} - \bar{\psi}_{SC}) \pi_{SP|L} - \theta (\bar{\psi}_{SP} - \bar{\psi}_{SC}) \pi_{SC|L}, \end{aligned}$$

where $\pi_{X|Y}$ is the share of members of Y that belong to X . So, disagreement among economic classes under cultural identity is:

$$\begin{aligned} \tau_{\hat{\psi}}^L - \tau_{\hat{\psi}}^U &= (\tau^L - \tau^U) - 2\theta \frac{(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC})}{\varphi} (\pi_{SP|U} - \pi_{SP|L}) \\ q_{\hat{\psi}}^L - q_{\hat{\psi}}^U &= (q^L - q^U) - 2\theta (\bar{\psi}_{SP} - \bar{\psi}_{SC}) (\pi_{SP|U} - \pi_{SP|L}). \end{aligned}$$

As a result:

$$\begin{aligned} q_{\hat{\psi}}^U - q_{\hat{\psi}}^L &> q_{\hat{\varepsilon}}^U - q_{\hat{\varepsilon}}^L \Leftrightarrow \\ (q^U - q^L) &> (\bar{\psi}_{SP} - \bar{\psi}_{SC}) (\pi_{SP|U} - \pi_{SP|L}) \Leftrightarrow \\ \rho(\bar{\varepsilon}_U - \bar{\varepsilon}_L) &> (\bar{\psi}_{SP} - \bar{\psi}_{SC}) (\pi_{SP|U} - \pi_{SP|L}), \end{aligned}$$

which is true if $\hat{\varepsilon} = \hat{\psi}$, since then $(\bar{\varepsilon}_U - \bar{\varepsilon}_L) = (\pi_{SP|U} - \pi_{SP|L})$ and $\rho > (\pi_{SP|U} - \pi_{SP|L})$ given

that (ε, ψ) is a standard bivariate normal (see Lemma 1 below). We also have:

$$\begin{aligned} \tau_{\hat{\psi}}^L - \tau_{\hat{\psi}}^U &< \tau_{\hat{\varepsilon}}^L - \tau_{\hat{\varepsilon}}^U \Leftrightarrow \\ -\frac{(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC})}{\varphi} (\pi_{SP|U} - \pi_{SP|L}) &< (\tau^L - \tau^U), \end{aligned}$$

which is equivalent to.

$$-(\beta - \rho)(\bar{\psi}_{SP} - \bar{\psi}_{SC}) (\pi_{SP|U} - \pi_{SP|L}) < (1 - \beta\rho)(\bar{\varepsilon}_U - \bar{\varepsilon}_L),$$

which is always true if $\hat{\varepsilon} = \hat{\psi}$ and hence $(\bar{\psi}_{SP} - \bar{\psi}_{SC}) = (\bar{\varepsilon}_U - \bar{\varepsilon}_L)$. ■

Lemma 1 *If $\hat{\varepsilon} = \hat{\psi} > 0$, then $\rho > \pi_{SP//U} - \pi_{SP//L}$*

Proof. Let $a = \hat{\varepsilon} = \hat{\psi} > 0$. Recall that $\pi_{SP//U} = \Pr(\psi > a | \varepsilon > a)$, $\pi_{SP//L} = \Pr(\psi > a | \varepsilon \leq a)$, and that (ε, ψ) are normally distributed with mean 0, variance of 1 and correlation coefficient $\rho > 0$. Define

$$g(\rho, a) := P(\psi > a | \varepsilon > a) - P(\psi > a | \varepsilon < a) = \frac{P(\psi > a, \varepsilon > a)}{1 - \Phi(a)} - \frac{P(\psi > a, \varepsilon < a)}{\Phi(a)} \quad (34)$$

Where $\Phi(a) := P(\psi < a)$. Now, we can write:

$$1 - \Phi(a) = P(\psi > a) = P(\psi > a, \varepsilon > a) + P(\psi > a, \varepsilon < a) \quad (35)$$

Then using (35) we can rewrite (34) as:

$$g(\rho, a) = \frac{P(\psi > a, \varepsilon > a)}{(1 - \Phi(a))\Phi(a)} - \frac{1 - \Phi(a)}{\Phi(a)} < \frac{P(\psi > a, \varepsilon > a)}{(1 - \Phi(a))\Phi(a)} \quad (36)$$

So, if we show that this last term is indeed smaller than $\rho \forall \rho \in (0, 1)$ then we are done. Clearly $g(0, a) = 0$ and $g(1, a) = 1$. Since at the extremes of the domain $\rho = g(\rho, a)$, we can show that (36) is smaller than ρ by showing that it is convex in ρ for $\rho \in (0, 1)$. Notice that only the numerator of (36) depends on ρ , hence we can compute:

$$\frac{\partial^2}{\partial \rho^2} \int_a^\infty \int_a^\infty \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho^2)}\right) dx dy \quad (37)$$

After computation we obtain:

$$\frac{e^{-\frac{a^2}{1+\rho}}(\rho - 1)^2(\rho + 1)(a^2(1 - \rho) + \rho + \rho^2)}{2\pi(1 - \rho^2)^{\frac{7}{2}}} > 0 \quad (38)$$

with the inequality holding since $\rho < 1$.

We show hereafter the computation to get from (37) to (38):

$$\begin{aligned}
& \frac{\partial^2}{\partial \rho^2} \int_a^\infty \int_a^\infty \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right) dx dy \\
= & \int_a^\infty \frac{\partial}{\partial \rho} \int_a^\infty \frac{\partial}{\partial \rho} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right) dx dy \\
= & \int_a^\infty \frac{\partial}{\partial \rho} \int_a^\infty \left\{ \frac{\rho \exp\left(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right)}{2\pi(1-\rho^2)^{3/2}} + \frac{\exp\left(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right) \left(\frac{xy}{1-\rho^2} - \frac{\rho(x^2+y^2-2\rho xy)}{(1-\rho^2)^2}\right)}{2\pi\sqrt{1-\rho^2}} \right\} dx dy \\
= & \int_a^\infty \frac{\partial}{\partial \rho} \int_a^\infty \exp\left(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right) \frac{-\rho^3 + xy + \rho^2 xy - \rho(x^2+y^2-1)}{2\pi(1-\rho^2)^{5/2}} dx dy \\
= & \int_a^\infty \frac{\partial}{\partial \rho} \exp\left(-\frac{a^2+y^2-2\rho ay}{2(1-\rho^2)}\right) \frac{(-1+\rho^2)(a\rho-y)}{2\pi(1-\rho^2)^{5/2}} dy \\
= & \int_a^\infty -\exp\left(-\frac{a^2+y^2-2\rho ay}{2(1-\rho^2)}\right) \\
& \frac{-a^3\rho^2 + a^2\rho(2+\rho^2)y - a(1+2\rho^2)(-1+\rho^2+y^2) + \rho y(-3+3\rho^2+y^2)}{2\pi(1-\rho^2)^{7/2}} dy \\
= & \frac{e^{-\frac{a^2}{1+\rho}}(\rho-1)^2(\rho+1)(a^2(1-\rho)+\rho+\rho^2)}{2\pi(1-\rho^2)^{7/2}}
\end{aligned}$$

■

Appendix 2: Economic Foundations for ρ

2.a Skill Biased Technical Change

We show that skilled biased technical change may increase the correlation between income and social progressiveness. Output is produced using the CES technology:

$$Y = [\gamma S^\mu + (1 - \gamma) U^\mu]^{\frac{1}{\mu}}, \mu > 0, 0 \leq \gamma \leq 1$$

where μ captures the degree of substitution between skilled labor S and unskilled labor U . When $\gamma = 1/2$ the technology is skill neutral, when $\gamma > 1/2$ it is skill biased (when $\gamma < 1/2$ it is biased in favor of unskilled labor which, as we will see, has counterfactual implications).

Each voter is described by (s, u, ψ) . s is the voter's expected endowment of skilled labor, u that of unskilled labor, which are normally distributed in the population around mean $(0, 0)$, variances $\sigma_s^2 = \sigma_u^2 = \sigma^2$, and covariance $\sigma_{su} = 0$. These assumptions on the variance covariance matrix are only made for simplicity and could be relaxed. ψ is normally distributed in the population with mean zero and $\sigma_\psi^2 = 1$. It also features covariances $\sigma_{s\psi} = \omega\sigma > 0$ and $\sigma_{u\psi} = 0$. Skilled labor is positively correlated with progressiveness, owing for instance to education. The average and hence aggregate endowment is $\bar{s} = \bar{u} = 1$.

Profit maximization at the aggregate endowment yields the skill premium:

$$\frac{\gamma}{1 - \gamma} \left(\frac{\bar{s}}{\bar{u}} \right)^{\frac{1}{\mu}} = \frac{w_s}{w_u},$$

where w_s and w_u are the wages earned by one unit of skilled and unskilled labor, respectively. Given equal aggregate endowment, this yields:

$$\frac{\gamma}{1 - \gamma} = \frac{w_s}{w_u}. \quad (39)$$

Output in the economy is equal to one, which is split between skilled and unskilled labor as follows:

$$w_s = \gamma, \quad w_u = 1 - \gamma. \quad (40)$$

The worker's labor income is equal to:

$$\gamma s + (1 - \gamma) u,$$

which is normally distributed in the population with mean 1.

The variance of workers' income in society is equal to:

$$\sigma_\varepsilon^2 = [\gamma^2 + (1 - \gamma)^2] \sigma^2.$$

The covariance between income and social progressiveness is then equal to:

$$\sigma_{\varepsilon,\psi} = \gamma\omega\sigma. \quad (41)$$

Income and social progressiveness are positively correlated. Socially progressive voters have in fact a larger skill endowment, which translates into higher income, especially if the remuneration γ of skills is higher.

As the remuneration γ of skills increases, the correlation between income and social progressiveness increases because:

$$\frac{\partial \sigma_{\varepsilon,\psi}}{\partial \gamma \sqrt{\sigma_\varepsilon^2}} = \frac{\omega\sigma^3(1-\gamma)}{\sigma_\varepsilon^3} > 0.$$

In the model of Section 3, the variance of income and of social progressiveness are fixed at one. To apply Proposition 2 to this model, then, we need to consider an increase in skill bias γ that increases the correlation ρ while holding the variance of income constant. This experiment corresponds to a marginal increase in γ starting from a skill neutral technology $\gamma = 1/2$ (and assuming that σ^2 is such that $\sigma_\varepsilon^2 = 1$). In fact, $\partial\sigma_\varepsilon^2/\partial\gamma = 2(2\gamma - 1)\sigma^2$, which is zero at $\gamma = 1/2$.

More generally, higher skill bias γ increases ρ as well as σ_ε^2 . The latter effect, if strong, may favor class identity. Thus, skilled biased technical change favors cultural identity provided the new technology exerts a strong upward effect on correlation ρ and a weaker, albeit possibly positive, effect on income inequality σ_ε^2 .

2.b Globalization

We now show that openness to trade can also increase the correlation between income and social progressiveness if voters working in import competing sectors are more conservative. A voter's utility from private consumption is:

$$c_n + U(c_x) + U(c_m),$$

where c_n is consumption of a non-tradeable good, c_x is consumption of the exported good, c_m is consumption of the imported good, where $U(c) = ac - c^2/2$, where $a > 1$. All voters have an equal endowment of one unit of the non-tradeable good. Each voter is also endowed with x units of labor in the export sector and m units of labor in the import sector. Labor endowments vary across voters.

Specifically, x and m are normally distributed around $\bar{x} = \bar{m} = 1/2$ with equal variance σ^2 and covariance $\sigma_{xm} = 0$ (these assumptions on the variance covariance matrix are also made for simplicity only). Labor is transformed into goods one to one. Labor endowments are correlated with a voter's culture ψ . In particular, export sector labor is positively correlated

with progressiveness, $cov(\psi, x) = \omega\sigma > 0$, while $cov(\psi, m) = 0$.

A voter is described by (x, m, ψ) . The government levies distortionary taxes on labor income. Nothing substantive changes if also the nontradeable good is taxed (we only need to adjust the level of the aggregate nontradeable and labor endowments). The public good is in terms of the nontradeable good.

In autarky, the prices of the tradeable goods in terms of the non-tradeable good are given by:

$$p_x = p_m = a - 1/2.$$

We assume for simplicity that the two goods have the same autarky price $p = 1$, i.e. $a = 3/2$. If the country opens up, the prices of tradeable goods change (we can think of "opening up" as the coordinated removal of non tariff barriers to imports by many countries). As in the case of skilled biased technical change, we consider price changes that leave total income constant: the price of the export goods increases by $dp_x > 0$, that of the import good decreases by the same amount, $dp_m = -dp_x$. This implies that the income of a generic voter is equal to:

$$x + m + dp_x(x - m),$$

where $dp_x = 0$ corresponds to autarky. Higher dp_x signifies more openness.

The variance of income in society is equal to:

$$\sigma_\varepsilon^2 = [(1 + dp_x)^2 + (1 - dp_x)^2] \sigma^2.$$

The covariance between income and social progressiveness is then equal to:

$$\sigma_{\varepsilon, \psi} = (1 + dp_x) \omega \sigma. \tag{42}$$

Income and social progressiveness are positively correlated, the more so the higher is openness dp_x . Socially progressive voters have in fact a larger export sector endowment, which translates into higher income, especially if the premium dp_x for the export good is higher.

As openness dp_x increases, the correlation between income and social progressiveness increases because:

$$\frac{\partial}{\partial dp_x} \frac{\sigma_{\varepsilon, \psi}}{\sqrt{\sigma_\varepsilon^2}} = \frac{2\omega\sigma^3}{\sigma_\varepsilon^3} (1 - dp_x),$$

which is positive (if $dp_x > 1$ the price of the import good would be negative).

Once again, to stay within the model of Section 3, consider an increase in dp_x that increases the correlation ρ while holding the variance of income constant. This experiment corresponds to a marginal increase in dp_x starting from autarky $dp_x = 0$ (and assuming that σ^2 is such

that $\sigma_\varepsilon^2 = 1$). In fact, $\partial\sigma_\varepsilon^2/\partial dp_x = 2dp_x\sigma^2$, which is zero at $dp_x = 0$.

Thus, openness to trade can cause a switch to cultural identity. Note that this property is not generic to all trade shock. It is only true if openness to trade hurts conservative voters while it benefits progressive voters. Furthermore, even if the trade shock increases ρ , for a large price change dp_x it will also increase σ_ε^2 . The latter effect, if strong, may favor class identity. In general, trade shocks that hurt conservative voters favor cultural identity provided they exert a strong upward effect on correlation ρ and a weaker, albeit possibly positive, effect on inequality σ_ε^2 .

Appendix 3. Multidimensional Identity

A voter is captured by (ψ, ε) as before. The voter can identify with his income group $I = U, L$, or his cultural group $C = SP, SC$, but we now also allow him to identify with his income *and* cultural group (I, C) . We call this latter case "joint identity". Each group G is summarized by its income-culture type $(\bar{\varepsilon}_G, \bar{\psi}_G)$. Under joint identity, the type of $G = (I, C)$ is $(\bar{\varepsilon}_I, \bar{\psi}_C)$, where $\bar{\varepsilon}_I$ is the average income of class I and $\bar{\psi}_C$ is the average culture of C .

Ingroup vs. Outgroup Types

In Section 3 voters are more likely to identify with $G = I, C$ the larger the income and cultural differences between it and the outgroup \bar{G} . This is also true with respect to joint identity $G = (I, C)$. Before studying identification, we characterize ingroup-outgroup differences under joint identity. As we will see, relative to the broader groups $G = I, C$, joint identity reduces ingroup-outgroup differences in income, culture, or both. This renders joint identity $G = (I, C)$ less appealing relative to identity with broader groups.

Denote by $(\bar{\varepsilon}_{(I,C)}, \bar{\psi}_{(I,C)})$ the outgroup of $G = (I, C)$. This is the average income and cultural type of the other three quadrants, formally:

$$\bar{\varepsilon}_{(I,C)} \equiv \frac{\sum_{(I',C') \neq (I,C)} \pi_{I',C'} \cdot \bar{\varepsilon}_{I'}}{\sum_{(I',C') \neq (I,C)} \pi_{I',C'}} = \frac{\pi_{I,\bar{C}} \bar{\varepsilon}_I + \pi_{\bar{I}} \bar{\varepsilon}_{\bar{I}}}{1 - \pi_{I,C}}, \quad (43)$$

$$\bar{\psi}_{(I,C)} \equiv \frac{\sum_{(I',C') \neq (I,C)} \pi_{I',C'} \cdot \bar{\psi}_{C'}}{\sum_{(I',C') \neq (I,C)} \pi_{I',C'}} = \frac{\pi_{I,C} \bar{\psi}_C + \pi_{\bar{C}} \bar{\psi}_{\bar{C}}}{1 - \pi_{I,C}}. \quad (44)$$

Outgroup income in (43) averages the income of the outgroup class $\bar{\varepsilon}_{\bar{I}}$ and the income $\bar{\varepsilon}_I$ of the ingroup class. The latter owes to the share $\pi_{I,\bar{C}}$ of voters who are culturally different but economically similar to (I, C) . Similarly, outgroup culture in (44) averages the values $\bar{\psi}_{\bar{C}}$ of the cultural outgroup with those of the cultural ingroup $\bar{\psi}_C$. The latter owes to the share $\pi_{I,C}$ of voters who are culturally similar but economically different from (I, C) .

Using Equations (43) and (44), the income and cultural differences between ingroup (I, C) and outgroup, (I, C) , are equal to:

$$\bar{\varepsilon}_{(I,C)} - \bar{\varepsilon}_{(I,C)} = \frac{\pi_{\bar{I}}}{1 - \pi_{I,C}} (\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}), \quad (45)$$

$$\bar{\psi}_{(I,C)} - \bar{\psi}_{(I,C)} = \frac{\pi_{\bar{C}}}{1 - \pi_{I,C}} (\bar{\psi}_C - \bar{\psi}_{\bar{C}}). \quad (46)$$

Income and cultural differences with outgroups load onto class inequality ($\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}$) and cultural inequality ($\bar{\psi}_C - \bar{\psi}_{\bar{C}}$), respectively. Under broader identities, $G = I, C$, these differences load on the dimension of identification (income for $G = I$ and culture for $G = C$) according to the correlation coefficient ρ . Under joint identity, conflict depends on purer income and cultural differences in society.

Crucially, income and cultural differences with the outgroup are muted. Under joint identity income differences are smaller than under class identity, $\left| \bar{\varepsilon}_{(I,C)} - \bar{\varepsilon}_{(\bar{I},\bar{C})} \right| < |\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}|$, and cultural differences are smaller than under cultural identity $\left| \bar{\psi}_{(I,C)} - \bar{\psi}_{(\bar{I},\bar{C})} \right| < |\bar{\psi}_C - \bar{\psi}_{\bar{C}}|$. This occurs because under joint identity some outgroups are similar to ingroups (captured by $\pi_{\bar{I}}/(1 - \pi_{I,C}) < 1$ and $\pi_{\bar{C}}/(1 - \pi_{I,C}) < 1$).

How does joint identity (I, C) fare in terms of ingroup-outgroup cultural differences compared to class identity $G = I$? And how does it fare in terms of income differences compared to cultural identity $G = C$? The answer depends on whether the voter belongs to a quadrant (I, C) that exhibits positive correlation between income and culture ($\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}$) ($\bar{\psi}_C - \bar{\psi}_{\bar{C}} > 0$ – i.e. when the voter is conservative-lower class (L, SC) or progressive-upper class (U, SP) – or to a quadrant that exhibits negative correlation ($\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}$) ($\bar{\psi}_C - \bar{\psi}_{\bar{C}} < 0$), i.e. when the voter is progressive-lower class (L, SP) or conservative-upper class (U, SC).

Lemma 2 *There are two cases:*

i) If $(\varepsilon - \bar{\varepsilon}_I)(\psi - \bar{\psi}_C) > 0$, the following occurs. The cultural difference between ingroup $G = (C, I)$ and its outgroup has the same sign as $(\bar{\psi}_I - \bar{\psi}_{\bar{I}})$, and it is larger in magnitude than the latter if and only if $\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| < \frac{1}{\rho} \frac{\pi_{\bar{C}}}{1 - \pi_{I,C}}$. The income difference between ingroup $G = (C, I)$ and its outgroup has the same sign as $(\bar{\varepsilon}_C - \bar{\varepsilon}_{\bar{C}})$, and it is larger in magnitude than the latter if and only if $\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| > \rho \frac{1 - \pi_{I,C}}{\pi_{\bar{I}}}$.

ii) If $(\varepsilon - \bar{\varepsilon}_I)(\psi - \bar{\psi}_C) < 0$, the following occurs. The cultural difference between ingroup $G = (C, I)$ and its outgroup has the opposite sign of $(\bar{\psi}_I - \bar{\psi}_{\bar{I}})$. The income difference between ingroup $G = (C, I)$ and its outgroup has the opposite sign as $(\bar{\varepsilon}_C - \bar{\varepsilon}_{\bar{C}})$.

Simply put: in the positive correlation quadrants, which are the most relevant ones given the positive correlation between ε and ψ , joint identity preserves the direction of ingroup vs. outgroup differences relative to broader income and cultural groups. In the negative correlation quadrants, the effect of joint identity is more drastic: it changes the direction of ingroup vs outgroup conflict along one dimension. With respect to class identity, joint identity reverses cultural conflict with the outgroup. With respect to cultural identity, it reverses income conflict with the outgroup. The benefit of joint identity is that it better captures the distinctive cultural and income traits of each quadrant. As we will see, this implies that joint identity is favored in the negative correlation quadrants (L, SP) and (U, SC).

In sum, joint identity dilutes ingroup vs. outgroup conflict along the "primary" dimension along which the broader group is defined. In addition, joint identity better captures conflict along the "secondary" dimension in the negative correlation quadrants. These aspects turn critical for characterizing the identity regime and beliefs.

Identity Choice

A voter chooses the group $G = I, C, (I, C)$ maximizing group contrast:

$$C(G, \bar{G}) \simeq (\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})^2 + (\beta^2 + \hat{\kappa})(\bar{\psi}_G - \bar{\psi}_{\bar{G}})^2 - 2\beta(\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})(\bar{\psi}_G - \bar{\psi}_{\bar{G}}), \quad (47)$$

where $(\bar{\varepsilon}_{\bar{G}}, \bar{\psi}_{\bar{G}})$ is the outgroup type. By inserting in Equation (47) the ingroup vs outgroup differences under different identities, we derive the following result.

Proposition 7 *Suppose that the correlation between income and culture is sufficiently low, $\rho \leq \rho^*$, and that $\left(\frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C}\right) < 2\frac{\pi_{SC}}{\pi_L}$. Then:*

i) All voters in the positive correlation quadrants (L, SC) and (U, SP) never choose joint identity. They identify with their class when $\hat{\kappa} < \hat{\alpha}$ and with their culture otherwise.

ii) All voters in the negative correlation quadrants (L, SP) and (U, SC) choose cultural identity when $\hat{\kappa} \geq \hat{\alpha}$, but some of them may choose class or joint identity if $\hat{\kappa} < \hat{\alpha}$, where $\hat{\alpha} \geq 0$.

When the correlation ρ between income and culture is sufficiently low and when income conflict is sufficiently low relative to cultural conflict, $\left(\frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C}\right)$ is low enough, the possibility of joint identity does not affect the main patterns of identification. First people in the positive correlation quadrants (L, SC) and (U, SP) identify with their class or their cultural group, as if joint identity was not available ($\hat{\alpha}$ is the same threshold of Section 3).²¹ Second, voters in the negative correlation quadrants (L, SP) and (U, SC) choose cultural identity provided the importance of cultural policy $\hat{\kappa}$ is high enough. Below threshold $\hat{\alpha}$, these voters may choose joint identity, or they may choose class identity for very low $\hat{\kappa}$ and joint identity for intermediate $\hat{\kappa}$. In general though, and consistent with our main model, higher $\hat{\kappa}$ reduces the prevalence of class identity and increases the prevalence of cultural identity.²²

Intuitively, voters in the negative correlation quadrants are most attracted by joint identity. A progressive lower class voter dislikes class identity due to its conservative trait, and dislikes cultural identity due to its upper class trait. He may thus identify with the narrower progressive-lower class group. The same principle holds for a conservative-upper class voter.

²¹When correlation ρ is low, voters in the positive quadrants align according to the most salient issue of the moment, maximizing contrast there. If correlation ρ becomes high but not perfect, joint identity may become fitting because the dampening factors $\frac{\pi_I}{1-\pi_{I,C}}$ and $\frac{\pi_C}{1-\pi_{I,C}}$ get closer and closer to one. However note that, as ρ gets higher and higher, all identities become similar to each other.

²²In the current model class identity may never be chosen, even if $\hat{\kappa} = 0$, because cultural disagreement is more important than income disagreement: it affects preferences over both redistributive and cultural policy.

Identity Switches, Changes in Beliefs and in Policy Demands

Consider the effect of higher salience of cultural policy $\hat{\kappa}$. Higher $\hat{\kappa}$ causes voters in the positive correlation quadrants (L, SC) and (U, SP) to switch from class to cultural identity when $\hat{\kappa}$ crosses $\hat{\alpha}$. Thus, the beliefs and preferences of these voters mimic those in our main model.

Consider voters in the negative correlation quadrants (L, SP) and (U, SC) . As $\hat{\kappa}$ becomes large, these voters switch to cultural identity, from either joint or class identity. If the switch is from class to culture, the results are the same as in Section 3. If they switch joint to cultural identity, the results are described below.

Proposition 8 *Consider voters in quadrants (L, SP) and (U, SC) . Under joint identity, their beliefs about income and culture are equal to:*

$$\begin{aligned}\varepsilon_{L,SP}^\theta &= \varepsilon + \theta \left(\frac{\pi_U}{1 - \pi_{L,SP}} \right) (\bar{\varepsilon}_L - \bar{\varepsilon}_U) & \psi_{L,SP}^\theta &= \psi + \theta \left(\frac{\pi_{SC}}{1 - \pi_{L,SP}} \right) (\bar{\psi}_{SP} - \bar{\psi}_{SC}), \\ \varepsilon_{U,SC}^\theta &= \varepsilon + \theta \left(\frac{\pi_L}{1 - \pi_{U,SC}} \right) (\bar{\varepsilon}_U - \bar{\varepsilon}_L) & \psi_{U,SC}^\theta &= \psi + \theta \left(\frac{\pi_{SP}}{1 - \pi_{U,SC}} \right) (\bar{\psi}_{SC} - \bar{\psi}_{SP}).\end{aligned}$$

If $\hat{\kappa}$ increases enough that the identity of these voters switches to culture, their beliefs change to:

$$\begin{aligned}\varepsilon_{L,SP}^\theta &\rightarrow \varepsilon_{SP}^\theta = \varepsilon + \theta \rho (\bar{\psi}_{SP} - \bar{\psi}_{SC}) & \psi_{L,SP}^\theta &\rightarrow \psi_{SP}^\theta = \psi + \theta (\bar{\psi}_{SP} - \bar{\psi}_{SC}), \\ \varepsilon_{U,SC}^\theta &\rightarrow \varepsilon_{SC}^\theta = \varepsilon + \theta \rho (\bar{\psi}_{SC} - \bar{\psi}_{SP}) & \psi_{U,SC}^\theta &\rightarrow \psi_{SC}^\theta = \psi + \theta (\bar{\psi}_{SC} - \bar{\psi}_{SP}).\end{aligned}$$

Provided ρ is low enough, income conflict among these voters dampens, formally $\varepsilon_{SC}^\theta < \varepsilon_{U,SC}^\theta$ and $\varepsilon_{SP}^\theta > \varepsilon_{L,SP}^\theta$, while cultural conflict accentuates, formally $\psi_{SC}^\theta < \psi_{U,SC}^\theta$ and $\psi_{SP}^\theta > \psi_{L,SP}^\theta$.

A switch from joint to cultural identity triggers the same change in beliefs as a switch from class to cultural identity in our main model. Provided ρ is low enough, as the progressive-lower class voters abandon their narrow group and become culturally identified, their income extremism weakens and their cultural progressiveness magnifies. Likewise, as the conservative-upper class voters leave their narrow group and identify with their culture, they become less polarized on income and more polarized on their conservatism. In terms of policy preferences, polarization in the demand for redistribution may go up or down, while polarization over cultural policy goes up. Views on cultural policy and redistribution become more correlated, because culture is a more important determinant of both views.

We cannot fully characterize the identity regime of voters in the negative correlation quadrants, so we cannot rule out the possibility that, as the salience of cultural policy $\hat{\kappa}$ increases

from low to intermediate, their identity switches from class to joint. Even in this case, however, there is a monotonic shift towards lower income conflict and stronger cultural conflict.

To see this, consider the (L, SP) voters first. As their identity switches from class to joint, their beliefs about income switch from the lower class prospect ε_L^θ to the less pessimistic joint prospect $\varepsilon_{L,SP}^\theta$. Along cultural values, the same voters switch from the mildly socially conservative stance of the lower class ψ_L^θ (which is especially mild if ρ is small), to the strongly socially progressive stance $\psi_{L,SP}^\theta$. Thus, their cultural preferences become more extreme and more correlated with income.

Consider next the voters in (U, SC) . As they switch from income to joint identity, they abandon the upper class income prospect ε_U^θ and embrace the less optimistic prospect of their narrower group $\varepsilon_{U,SC}^\theta$. They also switch from the mild progressiveness of the upper class ψ_U^θ (which is especially mild if ρ is small), to a strongly conservative stance $\psi_{L,SC}^\theta$. Also these voters, then, become economically less extreme, culturally more extreme, and more correlated in their economic and cultural beliefs.²³

The bottom line is simple: both types of voters behave similarly as if they were switching from economic to cultural identity: their income polarization drops, their cultural extremism increases. Joint identity can thus be viewed as an intermediate stage between income and cultural identity. Voters choose it when neither of the two dimensions is salient enough, but it preserves a monotonic progression from income to cultural conflict as the salience of cultural policy $\hat{\kappa}$ increases.

Proof of Proposition 7. The contrast under income, cultural, and joint identification is respectively equal to: ■

$$C(I, \bar{I}) = (\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}})^2 [(1 - \beta\rho)^2 + \hat{\kappa}\rho^2], \quad (48)$$

$$C(C, \bar{C}) = (\bar{\psi}_C - \bar{\psi}_{\bar{C}})^2 [(\beta - \rho)^2 + \hat{\kappa}], \quad (49)$$

$$C\left[(I, C), (\bar{I}, \bar{C})\right] = \quad (50)$$

$$= (\bar{\psi}_C - \bar{\psi}_{\bar{C}})^2 \left(\frac{\pi_{\bar{C}}}{1 - \pi_{I,C}} \right)^2 \left[\left(\beta - (2I_{>0} - 1) \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \frac{\pi_{\bar{I}}}{\pi_{\bar{C}}} \right)^2 + \hat{\kappa} \right] \quad (51)$$

where $I_{>0}$ is a dummy variable equal to 1 if the voter belongs to a group (I, C) featuring positive correlation between attributes, $(\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}})(\bar{\psi}_C - \bar{\psi}_{\bar{C}}) > 0$, and equal to zero in the negative correlation case.

By using (48) and (50), one finds that the voter at (I, C) prefers class identity to joint

²³It is immediate to see that this argument is also valid for voters located in the positive correlation quadrants when ρ is high enough that even these voters may choose joint identity when $\hat{\kappa}$ is intermediate. Overall, then, even if joint identity is allowed for, an increases in $\hat{\kappa}$ exerts very similar effects to those in our main model.

identity iff:

$$\hat{\kappa} \leq \alpha_*(I, C) \equiv \frac{\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 (1 - \beta\rho)^2 - \left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 \left[\beta - (2I_{>0} - 1) \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right| \frac{\pi_I}{\pi_C} \right]^2}{\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 \rho^2}, \quad (52)$$

where we assume that $\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 \rho^2 > 0$ for all (I, C) . This latter condition is satisfied provided ρ is low enough. Indeed, $\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 \rho^2$ is equal to zero when $\rho = 0$ and increases in ρ until $\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 > 0$. The term $\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2$ starts positive, it then increases in ρ for $I_{>0} = 1$ and decreases in ρ for $I_{>0} = 0$. At $\rho = 1$ we have $\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 < 1$. Thus, for $\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right| \geq 1$ (which we have assumed in Propositions 4, 5 and 6), there is threshold $\rho_1 > 0$ such that, for $\rho \leq \rho_1$, we have $\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 \rho^2 > 0$.

By using (49) and (50), one finds that the voter at (I, C) prefers cultural to joint identity iff:

$$\hat{\kappa} \geq \alpha^*(I, C) \equiv \frac{\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 \left[\beta - (2I_{>0} - 1) \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right| \frac{\pi_I}{\pi_C} \right]^2 - (\beta - \rho)^2}{1 - \left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2}. \quad (53)$$

On the basis of our prior analysis, and in particular owing to Proposition 2, it is evident that for $\hat{\kappa} > \max[\hat{\alpha}, \alpha^*(I, C)]$ all voters belonging to (I, C) identify culturally. By the same token, for $\hat{\kappa} \leq \min[\hat{\alpha}, \alpha_*(I, C)]$, all voters belonging to (I, C) identify with their class. In these cases, voters in (I, C) behave identically to the voters in our baseline model.

Consider now the two separate leading cases. First, the case in which (I, C) features positive correlation, $(\bar{\varepsilon}_I - \bar{\varepsilon}_I)(\bar{\psi}_C - \bar{\psi}_C) > 0$, namely $I_{>0} = 1$. In this case, one can rule out joint identity by imposing:

$$\frac{\alpha^*(I, C) < \alpha_*(I, C)}{\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 \left[\beta - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right| \frac{\pi_I}{\pi_C} \right]^2 - (\beta - \rho)^2} < \frac{\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 (1 - \beta\rho)^2 - \left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 \left[\beta - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right| \frac{\pi_I}{\pi_C} \right]^2}{\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 \rho^2},$$

which is equivalent to:

$$\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 \left[\beta - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right| \frac{\pi_I}{\pi_C} \right]^2 \left[1 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 \rho^2 \right] < (1 - \beta\rho)^2 \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 \left[1 - \left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 \right] + (\beta - \rho)^2 \left[\left(\frac{\pi_C}{1 - \pi_{I,C}} \right)^2 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_I}{\bar{\psi}_C - \bar{\psi}_C} \right|^2 \rho^2 \right].$$

A sufficient condition for the above inequality is that:

$$\left(\frac{\pi_{\bar{C}}}{1 - \pi_{I,C}}\right)^2 \left[\beta - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \frac{\pi_{\bar{I}}}{\pi_{\bar{C}}} \right]^2 \left[1 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right|^2 \rho^2 \right] < (1 - \beta\rho)^2 \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right|^2 \left[1 - \left(\frac{\pi_{\bar{C}}}{1 - \pi_{I,C}}\right)^2 \right],$$

which can in turn be written as:

$$\left(\frac{\beta - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \frac{\pi_{\bar{I}}}{\pi_{\bar{C}}}}{1 - \beta\rho} \right)^2 \left(\frac{\pi_{\bar{C}}}{1 - \pi_{I,C}} \right)^2 \left[1 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right|^2 \rho^2 \right] < \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right|^2 \left[1 - \left(\frac{\pi_{\bar{C}}}{1 - \pi_{I,C}}\right)^2 \right].$$

The first term in round brackets on the left hand side increases in β provided $\rho \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \frac{\pi_{\bar{I}}}{\pi_{\bar{C}}} \leq 1$. We assume for now that this is the case. This implies that if the inequality holds at $\beta = 1$, it also holds at any $\beta < 1$. As a result, given $\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \geq 1$, a sufficient condition is in turn:

$$\left(\frac{\pi_{\bar{C}}}{1 - \pi_{I,C}} \right)^2 \left[1 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \frac{\pi_{\bar{I}}}{\pi_{\bar{C}}} \right]^2 < (1 - \rho)^2,$$

which leads to the even more stringent condition:

$$\left(1 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \frac{\pi_{\bar{I}}}{\pi_{\bar{C}}} \right)^2 < (1 - \rho)^2,$$

which is equivalent to:

$$\rho < \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \left(\frac{\pi_{\bar{I}}}{\pi_{\bar{C}}} \right) < 2 - \rho.$$

Under the maintained assumption $\pi_L \geq \pi_{SC}$ and hence $\left(\frac{\bar{\varepsilon}_U - \bar{\varepsilon}_L}{\bar{\psi}_{SP} - \bar{\psi}_{SC}} \right) \geq 1$, considering quadrants (U, SP) and (L, SC) this condition becomes:

$$\rho < \rho_2 \equiv \min \left[\left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \left(\frac{\pi_U}{\pi_{SP}} \right), 2 - \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \left(\frac{\pi_L}{\pi_{SC}} \right) \right],$$

which can be fulfilled provided the necessary condition $2 > \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \left(\frac{\pi_L}{\pi_{SC}} \right)$ is satisfied, which we assume to be the case. We also have the previous condition $1 \leq \rho \left| \frac{\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}}}{\bar{\psi}_C - \bar{\psi}_{\bar{C}}} \right| \left(\frac{\pi_L}{\pi_{SC}} \right)$ ensuring sufficiency for evaluating the inequality at $\beta = 1$. This condition boils down to $\rho \leq \rho_3$, where ρ_3 is a suitable threshold. This implies that, provided $\rho < \rho^* \equiv \min(\rho_1, \rho_2, \rho_3)$, voters in the quadrants (U, SP) and (L, SC) never choose joint identity and hence behave exactly as in the main model.

Consider now the case in which (I, C) features negative correlation, $(\bar{\varepsilon}_I - \bar{\varepsilon}_{\bar{I}})(\bar{\psi}_C - \bar{\psi}_{\bar{C}}) < 0$, namely $I_{>0} = 0$. By inspecting Equation (50) and by comparing it with (48) and (49) one can see that it is not possible to rule out joint identity. In particular, even if $\rho < \rho^*$, it is possible that for sufficiently low levels of $\hat{\kappa}$ that joint identity prevails over class identity for voters in this quadrant. It is in particular not possible to find conditions that ensure that $\alpha^*(I, C) < \alpha_*(I, C)$ is positive for all $\rho < \rho^*$. Thus, there exist parameter constellations in which, provided $\hat{\kappa}$ is low enough, joint identity prevails. Of course, as $\hat{\kappa}$ becomes large enough that $\hat{\kappa} > \hat{\alpha} \equiv \max[\alpha^*(U, SC), \alpha^*(L, SP)]$, cultural identity prevails in all these quadrants.

Appendix 4. Empirical Analysis

Trends in Political Conflict: PEW and ANES

The data used to create Figure 2 are publicly available on the *Pew Research Center* website. Specifically, we use data from the following surveys: *June 2018 Political Survey*, *January 2017 Political Survey*, *December 2015 Political Survey*, *December 2014 Political Survey*, *March 2012 Political Survey*, *December 2011 Political Survey*, *February 2010 Political Survey*, *February 2009 Political and Economic Survey*, *January 2008 Political Survey*, *September 2007 Political Survey*, *January 2006 News Interest Index*, *January 2005 News Interest Index*, *July 2004 Foreign Policy and Party Images*, *April 2003 Iraq Poll*, *February 2001 News Interest Index*. All such surveys are conducted on nationally representative samples of US adults aged 18 or more, with size ranging from 1303 individuals in 2010 to 2009 individuals in 2004. Survey weights are used to enhance representativeness at national level.

For the analysis of the most important problem we rely on the following question: "What do you think is the most important problem facing the country today? [Record up to three responses, in order of mention]." The question is open-ended, but in the public release of the datasets answers have been classified in roughly 55 macro categories, with only minor changes in classification over time. We further aggregate the categories "Abortion" and "Rights of Women Under Attack/Rolling Back" in the macro category "Abortion and Women Rights". To create the trends, we consider for each of the selected issues the share of respondents including such issue among their first three mentions.

All other figures use data from the *American National Election Studies* (ANES), and in particular of the surveys carried out in 1996, 1998, 2000, 2004, 2008, 2012 and 2016. We use the version of the variables available in the Cumulative Dataset of December 2018, and complement such information with data from the yearly releases when required. Following standard practice of dynamic analyses on ANES, we restrict the analysis to the Face-to-Face sample. Results are robust if we add the WEB sample, available for years 2012 and 2016. Given that the target population of the analysis consists of all adult individuals living in the US, in computing aggregates and running regressions we use individual survey weights, which ensure that the sample is representative of the US adult population, at national level. Individual survey weights are rescaled so that each wave/year has a cumulative weight of one. Yearly sample sizes range from roughly 1200 individuals in 2004 to about 2300 individuals in 2008. Below we describe the questions and variables used in the analysis.

To measure policy opinions we rely on the following questions:

Redistr. Spending "Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel

that it is important for the government to provide many more services even if it means an increase in spending. Where would you place yourself on this scale?" Answers are given on a seven-point scale, and recoded so that the variable is increasing in respondents' desired size of government.

Redistr. Assist "Some people feel that the government in Washington should see to it that every person has a job and a good standard of living. Others think the government should just let each person get ahead on his/their own. Where would you place yourself on this scale?" Answers are given on a seven-point scale, and recoded so that the variable is increasing in respondents' desired government assistance.

Immigration "Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be [1. increased a lot; 2. increased a little; 3 left the same as it is now; 4 decreased a little; 5. decreased a lot]?" Answers are in a scale from 1 to 5, following the order in which they appear in the question. We reverse the scale so that higher values correspond to more liberal views.

Race Relations Index constructed from the following two questions (Group Thermometer): "Still using the thermometer, how would you rate the following group: Blacks." "Still using the thermometer, how would you rate the following group: Whites." Answers are collected on a 0-100 scale, and answers higher than 97 are coded as 97. 0 represents the "coldest" (most averse) feelings, while 100 is "warmest" feelings. Our index of race relations is simply the difference between the rating given to black people and the one given to white people.

Abortion "There has been some discussion about abortion during recent years. Which one of the opinions on this page best agrees with your view? You can just tell me the number of the opinion you choose. [1. By law, abortion should never be permitted; 2. The law should permit abortion only in case of rape, incest, or when the woman's life is in danger; 3. The law should permit abortion for reasons other than rape, incest, or danger to the woman's life, but only after the need for the abortion has been clearly established; 4. By law, a woman should always be able to obtain an abortion as a matter of personal choice]."

Using the questions described above, we construct the following indexes, based on the polychoric correlation matrix computed pooling the seven waves together.

Redistribution First polychoric principal component extracted from "Redistr. Spending" and "Redistr. Assist". It correlates positively with both measures.

Culture First polychoric principal component extracted from "Immigration", "Race Relations" and "Abortion". It correlates positively with all three measures.

Each of the two principal components described above are then regressed on individual party affiliation and wave fixed effects using the following regression, estimated with OLS,

$$y_i = \alpha + \alpha_t + \sum_{g \in \{D,R\}} \gamma_{gt} party_g + \epsilon_i$$

where y_i is the attitude/preference of individual i , α_t are wave fixed effects, $party_g$ are group dummy variables for people identifying as Democrats or Republicans (the omitted category being political independents); the coefficients of these group dummies, γ_{gt} , are wave-specific. We use standardized residuals from these regressions as our final measures of individual opinions/policy preferences in all the analyses carried out on ANES data, except for figures A1, A2, A5, A6 and panels (b) and (c) of Figure A7, in the Appendix. In Figure A1, we use the variables "Redistribution" and "Culture" without residualizing; in Figure A2, the sample is restricted to people who identify as Independents, and therefore we do not residualize. In Figure A5 and in panel (b) of Figure A7 we do not condition on party affiliation, but only on wave fixed effects. The latter are used to highlight the trends in contrast between different groups in each year. Similarly, in Figure A6 and in panel (c) of Figure A7, when focusing on political independents, we use residuals after conditioning on wave fixed effects only. Prior to plotting the trends of figures 1, 4, A1, A2, A5, A6 and panels (b) and (c) of Figure A7, the residuals of "Redistribution" and "Culture" are standardized to have mean 0 and unit variance on the pooled sample (1996-2016).

Note that in Figure A3 and Figure A4, residualization on wave fixed effects is redundant, since we look at yearly differences in average thermometers between social groups. See the definition of "Affective Cultural Polarization" and "Affective Class Polarization" (below) for a description of the construction of figures 3, A3 and A4.

The other variables used in the analysis are the following:

Social Class "There's been some talk these days about different social classes. Most people say they belong either to the middle class or the working class. Do you ever think of yourself as belonging in one of these classes?" Depending on the answer, the following follow-up questions are asked: (i) "Well, if you had to make a choice, would you call yourself middle class or working class?"; (ii) "Well, if you had to make a choice, would you call yourself middle class or working class?"; (iii) "Would you say that you are about average middle/working class or that you are in the upper part of the middle/ working class?" We aggregate answers "Lower Class (Volunteered)", "Average Working", "Working" and "Upper Working" in the macro category "Lower/Working Class"; answers "Lower Middle" and "Average Middle" in the macro category "Middle Class"; and "Upper Middle" and "Upper (Volunteered)" in "Upper Middle/Upper Class". The question is not asked in 1996 and 1998.

Religiosity "Do you consider Religion to be an important part of your life?" Answers are binary, and we code "Yes" as 1, "No" as 0. We classify individuals as "Religious" if "Religiosity" is equal to 1 and as "Secular" if "Religiosity" is equal to 0.

Party "Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?". A small share of respondents answers "None/No preference". we code Independents and these respondents in the same category.

Affective Class Polarization To construct the indexes of class polarization used in Figure 3, we rely on the following two questions: "Still using the thermometer, how would you rate the following group: Big Businesses."; "Still using the thermometer, how would you rate the following group: Labor Unions." For both variables, answers are collected on a 0-100 scale, and answers higher than 97 are coded by the ANES staff as 97. 0 represents the "coldest" (most averse) feelings, while 100 is "warmest" feelings. Using the same specification presented above, we regress the two measures on the interaction between party affiliation and wave dummy variables, and compute the corresponding residuals. Residuals are standardized to have mean 0 and unit variance on the pooled sample. For each of the two residualized measures and each year, we construct the indexes of "Affective Class Polarization" by taking the difference in average feelings between "Lower/Working Class" and "Upper Middle/Upper Class". In Figure A3, we replicate the analysis without residualizing feelings. In Figure A4, the sample is restricted to people who identify as Independents, and therefore we do not condition on political affiliation. In both cases, the thermometer variables are still standardized to have 0 mean and unit variance on the pooled sample.

Affective Cultural Polarization To construct the indexes of cultural polarization used in Figure 3, we rely on the following two questions: "Still using the thermometer, how would you rate the following group: Christian Fundamentalist."; "Still using the thermometer, how would you rate the following group: Catholics." The question is not asked in 1996, 1998 and 2016. For both variables, answers are collected on a 0-100 scale, and answers higher than 97 are coded by the ANES staff as 97. 0 represents the "coldest" (most averse) feelings, while 100 is "warmest" feelings. Using the same specification presented above, we regress the two measures on the interaction between party affiliation and wave dummy variables, and compute the corresponding residuals. Residuals are then standardized to have mean 0 and unit variance on the pooled sample. For each of the two residualized measures and each year, we construct the indexes of "Affective Cultural Polarization" by taking the difference in average feelings between respondents classified as "Religious" and "Secular". In Figure A3, we replicate the analysis without residualizing feelings. In Figure A4, the sample is restricted to people who identify as Independents, and therefore we do not condition on political affiliation. In both cases, the thermometer variables are still standardized to have 0 mean and unit variance on the pooled sample.

Traditionalism First polychoric principal component of four questions asking if respondents agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat or disagree strongly with each the following statements: (i) "The newer lifestyles are contributing

to the breakdown of our society"; (ii) "The world is always changing and we should adjust our view of moral behavior to those changes"; (iii) "This country would have many fewer problems if there were more emphasis on traditional family ties"; (iv) "We should be more tolerant of people who choose to live according to their own moral standards, even if they are very different from our own". Answers are given on a 5-point scale, ranging from "Agree strongly" to "Disagree strongly". The first principal component correlates positively with (ii) and (iv) and negatively with (i) and (iii).

Conservative/Progressive In each year, respondents are classified as "Conservative" if they score (strictly) above the median of the distribution of Traditionalism in that year. Respondents are classified as "Progressive" if they score (weakly) below the median of Traditionalism in that year.