

# DISCUSSION PAPER SERIES

DP13390  
(v. 3)

## **Identity, Beliefs, and Political Conflict**

Nicola Gennaioli and Guido Tabellini

**PUBLIC ECONOMICS**



# Identity, Beliefs, and Political Conflict

*Nicola Gennaioli and Guido Tabellini*

Discussion Paper DP13390  
First Published 13 December 2018  
This Revision 10 July 2019

Centre for Economic Policy Research  
33 Great Sutton Street, London EC1V 0DX, UK  
Tel: +44 (0)20 7183 8801  
[www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Public Economics

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Nicola Gennaioli and Guido Tabellini

# Identity, Beliefs, and Political Conflict

## Abstract

We present a theory of identity politics that builds on two ideas. First, voters identify with the social group whose interests are closest to theirs and that features the strongest policy conflict with outgroups. Second, identification causes voters to slant their beliefs of self and others toward group stereotypes. The theory yields two main implications: i) voters' beliefs are polarized along the distinctive features of salient groups; ii) economic shocks that render new groups salient bring about large and non standard changes in beliefs and policies across many issues. In particular, exposure to globalization or cultural changes may induce voters to switch identities, dampening their demand for redistribution and exacerbating conflicts in other social dimensions. We show that survey evidence is broadly consistent with these implications.

JEL Classification: N/A

Keywords: N/A

Nicola Gennaioli - [nicola.gennaioli@unibocconi.it](mailto:nicola.gennaioli@unibocconi.it)  
*Department of Finance and IGIER, Università Bocconi and CEPR*

Guido Tabellini - [guido.tabellini@unibocconi.it](mailto:guido.tabellini@unibocconi.it)  
*Department of Economics and IGIER, Università Bocconi; CESifo; CIFAR and CEPR*

# Identity, Beliefs, and Political Conflict<sup>\*</sup>

Nicola Gennaioli<sup>†</sup> and Guido Tabellini<sup>‡</sup>

First version: October 2018 This version: July 2019

## Abstract

We present a theory of identity politics that builds on two ideas. First, voters identify with the social group whose interests are closest to theirs and that features the strongest policy conflict with outgroups. Second, identification causes voters to slant their beliefs of self and others toward group stereotypes. The theory yields two main implications: i) voters' beliefs are polarized along the distinctive features of salient groups; ii) economic shocks that render new groups salient bring about large and non standard changes in beliefs and policies across many issues. In particular, exposure to globalization or cultural changes may induce voters to switch identities, dampening their demand for redistribution and exacerbating conflicts in other social dimensions. We show that survey evidence is broadly consistent with these implications.

---

<sup>\*</sup>We are grateful to Giampaolo Bonomi, Ben Enke, Ed Glaeser, Pierre Jared, Elhanan Helpman, Nathan Nunn, Raffella Piccarreta, Giacomo Ponzetto, Jesse Shapiro, Andrei Shleifer, Marco Tabellini, Pierre Yared and to participants in workshops at Bocconi, the Bundesbank, CIFAR, Harvard, the LSE ERINN conference, the NBER, Princeton, UPF, Università Statale di Milano, the University of Namur, and to Giampaolo Bonomi, Viola Corradini, Daniele d'Arienzo, Carlo Medici, Francesca Misericocchi and Giulia Travaglini for outstanding research assistance. Tabellini thanks the ERC grant 741643 and Gennaioli the ERC (GA 647782) and MIUR (FARE grant) for financial support.

<sup>†</sup>Department of Finance and IGIER, Università Bocconi.  
E-mail: nicola.gennaioli@unibocconi.it

<sup>‡</sup>Department of Economics and IGIER, Università Bocconi; CEPR; CESifo; and CIFAR.  
E-mail: guido.tabellini@unibocconi.it

# 1 Introduction

During the last decades the political systems of advanced democracies have witnessed momentous change. The classical left-right divide on redistribution has weakened, and new dimensions of political conflict such as immigration and civil rights have emerged. This is puzzling because during this period income inequality has increased, so one would have expected the demand for redistribution to soar. By contrast, many poor and less educated voters have gradually moved to the right, while richer and more educated voters have moved to the left (Piketty 2017). Even more puzzling, this phenomenon is correlated with the very exposure to adverse economic shocks. In the US, Autor et al. (2017) show that areas exposed to import competition have become more polarized and conservative. In Europe, Colantone and Stanig (2017, 2018 a,b) and Anelli et al. (2018) show that areas hurt by free trade or skilled biased technical change have moved toward nationalistic and socially conservative positions.

One striking illustration of this change in political demands is offered by Figure 1 below, which is constructed using data from ANES. It plots the views on government spending and immigration of white U.S. respondents who self define as belonging to the working class and to the middle/upper class.

Figure 1

Working class people favor higher government spending and oppose immigration, relative to members of the upper middle/upper class. However, during the last 15 years the disagreement over the size of government has almost disappeared, while that over immigration has widened. Similar patterns emerge in UK data (Evans and Mellon 2016). Why this drop in the demand for redistribution, and this increase of polarization over new non-economic dimensions despite growing income gaps?

This paper argues that a promising way to understand these phenomena is to consider voters' identity and its effect on beliefs. Inspired by the social psychology of identity (Tajfel and Turner 1979, Turner et al. 1987), we offer a theory in which a voter's beliefs are tainted by the stereotype of the group he identifies with. In turn, identification is driven by the "meta contrast" ratio: the voter identifies with the group of people similar to him, and that stand in starkest contrast with the outgroup. When redistributive conflict is strong, less educated and poorer voters identify with low income people. As a result, they form stereotypes of excessively low social mobility, and demand large redistribution. The reverse occurs for more educated and richer voters. If instead conflict over culture or globalization becomes strong, less educated and poorer voters identify with social conservatives or losers from trade. Due to this identity switch, their views on these issues become more radical, while views on rich-poor redistribution more moderate. Critically, economic shocks hitting socially conservative or

trade averse strata enhance similarity within these groups. As a result, such shocks may favor identification away from income class, reducing the demand for redistribution. More generally, by inducing specific belief distortions, endogenous social identity alters and propagates the effects of economic and social shocks.

Besides offering an explanation for puzzling shifts in policy demands, our theory helps shed light on growing evidence of systematic and persistent distortions in voters' beliefs (Flynn et al. 2017, Achen and Bartels 2016, Johnston et al. 2017). In particular, there is evidence that partisan identity or political ideology is systematically correlated with distortions in beliefs over aggregate social mobility (Alesina et al. 2018a), immigration (Alesina et al. 2018b), inequality (Gimpelson and Treisman 2018) and global warming (Kahan 2014). Consistently with our predictions, political beliefs become more polarized when group identities are primed (e.g., Janky 2018, Han and Wackman 2017). Moreover, and also in line with our theory, perceived polarization between opposite political groups exceeds actual polarization in policy views (Westfall et al 2015, Bordalo et al. 2016), particularly if political conflict becomes more salient (Bordalo et al. 2019b).

We develop our analysis in three steps. In Section 3 we study a simple model of class conflict to illustrate how identity influences beliefs through group stereotypes. Stereotypes are modelled as in Bordalo et al. (2016), namely they exaggerate distinctive group features. There are two predefined economic groups: Upper and Lower class, consisting of people having different income prospects. Everyone is exogenously identified with either of them. A voter identified with the Lower class slants his beliefs toward the stereotype of his group, a poor individual. Thus, he becomes too pessimistic about his social mobility, enhancing his demand for redistribution. The reverse happens for identification with the Upper class. This way, identity enhances actual and perceived polarization in the demand for redistribution. In a standard probabilistic voting model, these belief distortions influence the political platforms offered by opportunistic candidates. Thus, identity creates a multiplier whereby class-conflict leads to class-distorted beliefs that in turn breed stronger conflict.

Section 4 endogeneizes voters' identities, and shows that now certain non-economic shocks can reduce the demand for redistribution. We add to the previous setting cultural heterogeneity among voters, and a civil rights/immigration policy over which voters disagree. Voters can now identify with their income class (Upper vs Lower) or their cultural group (Progressive vs Conservative). Due to the meta contrast ratio, non economic shocks affect identity and hence beliefs about the two policies. Suppose that - due to an increase in education - socially progressive voters become more extreme; or that - due to an inflow of immigrants - the welfare relevance of immigration policy increases. These shocks increase conflict between opposite cultural groups, and also disagreement within economic classes. Due to both effects, the cohesion of the lower economic class drops, causing socially conservative lower class voters

to switch to cultural identity. As a result of this identity switch, these voters become more conservative and abandon class-based stereotypes on social mobility. Thus, a non-economic shock increases cultural disagreement and reduces the demand for redistribution.

Section 5 studies how even certain adverse economic shocks can drive identity away from class, reducing the demand for redistribution. This occurs when the shock disproportionately hits a certain group, such as socially conservative voters. For instance, skilled biased technical change hits voters who are on average more conservative because they are less educated or live in declining regions. Key of this shock is to render cultural groups more economically homogeneous, increasing their cohesion. Under some conditions, then, identities switch to culture so that low income people demand less redistribution, despite increasing inequality. A similar effect arises when we add trade policy to our model. A common nationalistic identity among losers from trade weakens class cohesion, giving rise to demand for protectionism and lower support for income redistribution. We also find that if, consistent with the data, losers from trade tend to be more socially conservative, exposure to globalization exerts correlated effects, consistent with extant evidence: it reduces demand for redistribution while it increases demand for protectionism, control of immigration, and constraints on civil rights.

In section 6 we take some of these predictions to the data. First, we explore time patterns in US survey data. The perceived importance of immigration and race has risen in recent years. Consistently with our model, this was accompanied by a strong increase of polarization over immigration and some dampening of polarization on government spending. We also find that people exposed to imports from China became more willing to accept cuts in public spending, more averse to immigrants, and consider abortion issues more important, consistently with our model. Finally, using a longitudinal survey on France we show that between 2013 and 2017 political conflict changed from left vs right to nationalist vs cosmopolitan. This is reflected in how people voted in Presidential elections and in how their attitudes changed, with a dampening of redistributive conflict and a clustering of conflict over nationalism, immigration and civil rights, as predicted by the theory.

A long tradition in the social sciences argues that group identity and group-tainted beliefs play a key role in political change. According to Marx and Engels, individual workers should identify with the proletariat, viewing themselves as part of a historical class struggle, rather than as carriers of specific cultural or regional traits. For nationalists like Mazzini or Herder individuals should instead downplay economic or religious differences and view themselves as part of the broader national community. Lipset and Rokkan (1960) describe the evolution of Western party systems as reflecting shifting identities across salient groups such as income classes, religious versus secular groups, center versus periphery, etc.

Akerlof and Kranton (2000) develop the first economic model where identification with a group changes the payoffs of certain actions. They do not consider beliefs. A few recent papers

introduce identity into political economy models (e.g., Shayo 2009, Helpman and Grossman 2018). In these papers, voters obtain positive utility from the status of their group, causing them to internalize the welfare of the latter. In our model identification is not related to status, so a voter is not penalized from identifying with the lower class, and affects policy demands through beliefs. We view these papers as highlighting complementary mechanisms.<sup>1</sup> One advantage of our approach is that it sheds light on belief distortions.

A different approach has been pursued by Benabou and Tirole (2011), (2016). They view identity as reflecting beliefs about one-self, and belief formation as motivated also by the desire to improve self-image and individual welfare, taking anticipatory utility into account. This idea has produced valuable insights on many phenomena, including cross country differences in redistribution (Benabou and Tirole 2006). Relative to these papers, our approach can help explain puzzling facts such as group stereotypes, group-stereotyped beliefs about self, and the notion that the beliefs of certain voters sometimes seem to run counter their individual interests.

A strand of work studies changing patterns of polarization in survey data (see Layman et al. 2006, Gentzkow 2016) and the growing importance of cultural conflict in the US. Several papers emphasize the role of parties and leaders (Glaeser et al. 2005, Glaeser and Ward 2006). Supply side factors are certainly important, but they alone cannot explain changes in voter beliefs, such as in Figure 1, particularly if they are common to a large number of countries. Of course, beliefs may change also in response to political messages. Future work should merge psychologically founded mechanisms for belief change like the one we outline here, with persuasion by politicians. Murphy and Shleifer (2004) argue that political entrepreneurs can mobilize existing social networks such as trade unions, and change their beliefs in unrelated domains. Our view of identity is broader than active membership. It is centered on perceived similarities and entails stereotyping of others. It helps explain large changes in political demands that may even dismantle existing parties or organizations.

Finally, our work is related to a rapidly growing literature that seeks to explain the rise of nationalism and populism as a reaction to economic distress and economic insecurity (Guiso et al 2017, Gidron and Hall 2017, Dal Bo et al. 2018).

---

<sup>1</sup>Fryer and Jackson (2008) study the implications of categorization for discrimination and identities. In line with the Social Identity Approach we instead emphasize stereotypes (Bordalo et al. 2016) and focus on different applications. Diermeyer and Li (2018) study a model of electoral competition with retrospective voting, where voters pay more attention to the policies of the candidate they identify with. This can induce polarization even if candidates are entirely opportunistic. Some papers also provide evidence supporting the implications of social identities, for private consumption (Atkin et al. 2019) or for political behavior and cultural traits (Grosfeld et al. 2013).



## 2 The Social Psychology of Identity and Beliefs

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 1998). The general idea is that when a social cleavage is salient, identification with one of the relevant groups anchors perceptions of others, of self, and behavior. Experiments using the "minimal group paradigm" found that even arbitrary groups influence behavior in this way (Tajfel, Billig, et al. 1971). We now discuss the building blocks of this mechanism.

### 2.1 Identification

According to SCT, identification with a group is a form of self-categorization in society. An individual belongs to many groups: income class, religious group, nation, etc, but he does not identify with all of them at the same time. As shown by several experiments (e.g., Hogg et al. 1998), identification and behavior depends on which groups are salient. For instance, at a soccer match one's own team is the salient group of identification, as it affects behavior more than other groups.

According to SCT, the salience of a group depends on its accessibility and fit (Oakes, 1987). At a soccer match, the group of "team supporters" and that of "soccer lovers" are accessible, in the sense that they easily come to mind, the group of "churchgoers" much less so. In this context, though, the most fitting group is "team supporters", for it reflects the current social cleavage. The importance of cleavages for salience gives rise to the so called "meta contrast ratio": an individual tends to identify with a group of people who are: i) similar to him, but also ii) highly dissimilar from/in conflict with, the outgroup.<sup>2</sup>

As we will see, this idea intuitively translates to politics. During times of redistributive conflict, one's own income class offers a fitting identity for political action. During times of regional or cultural clash, class identity subdues and is replaced by regional/cultural ones.

### 2.2 Stereotyped Beliefs and Depersonalization

How do groups and identity influence perceptions and behavior? According to SIT and SCT, identification causes individuals to perceive others and oneself in terms of their group (see also Sherif 1936, Festinger 1950). Early experiments on abstract groups by Tajfel and Wilkes

---

<sup>2</sup>The importance of cross-group differences in attenuating within group differences, reminiscent of the early Tajfel and Wilkes (1963) experiments, plays a key role in theories of categorization, starting from Rosch (1978). See also Tversky (1977) for a close notion of contextual similarity.

This approach remedies some difficulties of SIT, in which identity is a way to achieve self esteem. This notion clashes with the fact that individuals often identify with low status groups. According to the meta contrast ratio, the mere presence of a cleavage prompts identification, not (only) group status.

(1963), and subsequent work in social settings (Haslam and Turner 1992) highlight two effects. First, individuals stereotype groups by exaggerating differences with the outgroup while dampening within group differences. Society is divided in "us versus them". Second, when an individual identifies with a group he "depersonalizes", slanting his beliefs toward the perceived group view, namely the stereotype.

McGarty et al. (1992) defines such stereotype as follows: "...the less a person differs from ingroup members and the more he or she differs from out group members, the more representative he or she is of the in-group." Thus, the stereotype is not necessarily the most frequent trait in the group; it must also be infrequent in, or contrasting with, the out-groups.

We formalize stereotypes as in BCGS (2016), who build upon Kahneman and Tversky's representativeness heuristic. Suppose that two groups  $G$  and  $\bar{G}$  partition society along a certain trait, for instance education, with  $G$  denoting the group of highly educated people, and  $\bar{G}$  its complement.  $X$  is a random variable, e.g. individual income, distributed in the population with density  $f(X)$ . BCGS (2016) define a certain value  $X$  of the random variable to be more stereotypical of group  $G$  if it scores higher in the likelihood ratio:

$$\frac{f(X|G)}{f(X|\bar{G})}, \quad (1)$$

where  $f(X|G)$  is the actual distribution of  $X$  in group  $G$ . The stereotypical values of  $X$  for  $G$  are not only likely in this group. They are also unlikely in the outgroup  $\bar{G}$ . This captures precisely the above definition, yielding one key implication: when groups differ in the tails, the stereotype is more extreme than the average group position in the direction of true group differences. This is the so called "kernel of truth" property of stereotypes (BCGS 2016).<sup>3</sup>

BCGS (2016) assume that the stereotyped distribution of  $X$  in group  $G$  takes the form:

$$f^\theta(X|G) = f(X|G) \left[ \frac{f(X|G)}{f(X|\bar{G})} \right]^\theta Z, \quad (2)$$

where  $Z \geq 0$  is a constant ensuring that  $f^\theta(X|G)$  adds up to one and  $\theta > 0$  captures the strength of stereotypes. When  $\theta$  increases, stereotypical traits are overweighed more.

We extend this idea by two ways. First, we endogenize  $G$  on the basis of group similarity. Second, we model depersonalization. To do so, in our setting  $X$  is a feature of an individual, such as his future income, or his opinion about immigrants. The individual is genuinely uncertain, and has to form beliefs about  $X$ . As we discuss in Section 3, he does so by overweighing in his own uncertainty the values of  $X$  that are stereotypical of the group  $G$  with which he identifies (in a way isomorphic to (2)).

---

<sup>3</sup>BCGS (2016) offer empirical support for this notion, which also accounts for laboratory puzzles in probability judgments and for belief distortions in macro and finance (see Gennaioli and Shleifer 2018).

This approach implies that when groups differ in the tails, perceptions of groups and perceptions of self become more extreme, in the direction of true differences between groups. Consistent with this idea, BCGS (2018) show that the logic of Equation (2) captures beliefs about others and self in the context of gender groups. In their experiment, subjects assess uncertain ability in different domains of knowledge. They find that beliefs about women are too pessimistic, and women are too little self-confident, in domains where women do worse than men as a group, e.g. mathematics and sports. By contrast, beliefs about women are too optimistic, and women are too self-confident, in domains where they do better as a group such as emotion recognition, or Disney movies. Perceptions of groups and of self are stereotyped across many domains in line with true gender differences, as per the kernel of truth. Thus, beliefs are more extreme than reality, especially if gender is made more salient.<sup>4</sup>

We now apply these ideas to politics. When thinking about difficult questions such as "What is the optimal tax rate?", "Should same-sex adoptions be allowed?", or "Is Globalization good?", individuals do not just look at their material interests and information. They also think about the social group they feel similar to, so that their beliefs also incorporate group stereotypes. When the salience of groups changes, the voter's identification can also change, influencing his beliefs across several domains.

### 3 Identity and Economic Conflict

We introduce identity in the simplest political economy model of income taxation and public good provision. Economic conflict and class identity have been central in western political systems during the 20th century. The leading and most extreme example is communist class struggle. But even in the US, Campbell et al. (1966) show that during the New Deal many working class voters turned Democratic because they identified with this party as it represented "the common man". Here we take identification with one's economic class as given and study its effects on beliefs and policy. In Section 4 we endogenize identity and explain the shifting salience of economic conflict.

There is a measure one of individuals, each of which earns stochastic income  $1 + \tilde{\varepsilon}$ , where random variable  $\tilde{\varepsilon}$  has an individual-specific mean  $\varepsilon$  capturing the voter's rationally expected

---

<sup>4</sup>This mechanism accounts for the well documented phenomenon of "group polarization": the tendency for group members to embrace a more extreme position on a particular issue than their individually expressed positions (e.g., Mackie 1986). According to SCT, this phenomenon arises because identification leads individuals to sharpen differences in opinion away from outgroups. Our model captures this idea.

SIT holds that depersonalization can also yield positive self-esteem from the status of one's own group. In line with this notion, some experiments document ingroup favoritism, but the generality of these findings has been challenged (e.g., Hinkle and Brown 1990). It is now believed that ingroup bias arises only when discrimination is a group norm. We later show that our belief-based mechanism can generate a taste for ingroup success, but in a way consistent with the importance of group norms and beliefs.

income. Type  $\varepsilon$  is distributed in society according to cdf  $H(\varepsilon)$ , and it has zero mean, so aggregate income is 1. Beliefs about  $\varepsilon$  can be viewed as beliefs about own social mobility.

The individual derives utility from private consumption and a public good. The latter is financed by a proportional income tax  $\tau \geq 0$ , which entails quadratic distortions  $-\frac{\varphi}{2}\tau^2$ ,  $\varphi > 0$ , that fall on the government budget constraint. The utilities of private and public consumption are linear, with marginal utilities equal to 1 and  $\nu > 1$  respectively. Under rational beliefs, the expected utility of a generic income type  $\varepsilon$  is equal to:

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

where the last term is the utility from the public good. The rational bliss point of type  $\varepsilon$  is

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

The marginal utility  $\nu$  of the public good is large enough that even individuals with high  $\varepsilon$  are at an interior optimum, so that the optimal tax strictly decreases in expected income.

Given linear private consumption, the socially optimal tax rate  $\tau^0$  maximizes the welfare of the average individual with  $\varepsilon = 0$ . Thus, by (4),

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

### 3.1 Group Identity and Beliefs

Society is partitioned in two exogenous groups that differ in expected income: the lower class  $L \equiv \{\varepsilon | \varepsilon < \hat{\varepsilon}\}$ , and the upper class  $U \equiv \{\varepsilon | \varepsilon \geq \hat{\varepsilon}\}$ , where  $\hat{\varepsilon}$  is an historically given class boundary. For simplicity, we assume that all voters identify with their economic class. Average income differences across classes may be due to differences in the ownership of a productive factor (physical or human capital), to social connections, etc.

Identification causes depersonalization: perceptions of self align with the distinctive feature of one's own class. Marxists call this phenomenon class consciousness: the aware proletarian views his condition not as a product of his specific traits, but as reflecting a permanent class struggle. To formalize this notion, we apply equation (2) to individual conditions. Denote by  $f(\tilde{\varepsilon}|\varepsilon)$  the true income density of a voter with mean income  $\varepsilon$ . Then, identification with group  $G$  causes this voter to overweight the income realizations stereotypical of this group, in the sense of the likelihood ratio in (1). Formally, the beliefs of a voter  $\varepsilon$  identified with group

$G$  are modelled as the distorted distribution:

$$f^\theta(\tilde{\varepsilon}|\varepsilon, G) = f(\tilde{\varepsilon}|\varepsilon) \left[ \frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})} \right]^\theta Z. \quad (6)$$

where  $\theta, Z > 0$  and the group conditioned distribution  $f(\tilde{\varepsilon}|G) \propto \int_{\varepsilon \in G} f(\tilde{\varepsilon}|\varepsilon) dH(\varepsilon)$  is aggregated from all members of each class. Here we set  $\theta$  to be the same for all individuals, but this assumption can easily be relaxed and yield additional implications.

Stereotyped beliefs exaggerate conflict. Let  $\varepsilon_G^\theta = \int \tilde{\varepsilon} f^\theta(\tilde{\varepsilon}|\varepsilon, G) d\tilde{\varepsilon}$  denote the future mean income perceived by a type  $\varepsilon$  identified with group  $G$ . We prove the following result.

**Proposition 1** *Suppose that the pdf of income  $f(\tilde{\varepsilon}|\varepsilon)$  exhibits the Monotone Likelihood Ratio Property (MLRP). Then*

$$\varepsilon_L^\theta < \varepsilon < \varepsilon_U^\theta \quad (7)$$

*and belief distortions for an atomistic voter with generic mean income  $\varepsilon^*$  and density  $f(\tilde{\varepsilon}|\varepsilon^*)$  increase if: i)  $\theta$  increases; ii) the income distribution becomes more polarized in society, in the sense that the income pdf of any other voter  $\varepsilon \in G$  is replaced by  $f'(\tilde{\varepsilon}|\varepsilon) \propto k_G(\tilde{\varepsilon}) f(\tilde{\varepsilon}|\varepsilon)$ , where  $k_U(\tilde{\varepsilon})$  is increasing and  $k_L(\tilde{\varepsilon})$  is decreasing in  $\tilde{\varepsilon}$ .*

Holding fixed true future mean income  $\varepsilon$ , lower class identity causes people to be too pessimistic about their income prospects, upper class identity causes them to be too optimistic. As the two classes are made more unequal (see property ii)), in the sense that high income becomes more likely in  $U$  and less likely in  $L$ , these distortions amplify for a fixed type  $\varepsilon^*$ .

Belief distortions are due to the kernel of truth property of stereotypes, as highlighted by the comparative statics on  $\theta$ . Under MLRP, the stereotypical condition for the lower class is the risk of becoming poor, that for the upper class is the prospect of becoming rich. As I identify with a class characterized by lower upward mobility, I underestimate my own upward mobility prospects. As in Marxist analysis, the self conscious proletarian views himself as the member of an oppressed class, so he discounts the possibility of individually improving his own condition. The reverse occurs for upper class people.

Alesina et al. (2018a) show that in several western countries leftwing voters are less optimistic than rightwing voters about social mobility. This evidence concerns aggregate mobility, whereas in our model voters have distorted beliefs over their individual future conditions. Nevertheless, these two notions of social mobility are probably correlated, as individuals are likely to extrapolate their perceived subjective condition to aggregate variables.<sup>5</sup>

---

<sup>5</sup>Cruces et al. (2009) 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 neighborhood, as if each voter uses the local income distribution as a reference. Misperception of current

The properties of Proposition 1 are sufficient for most of our results, and in fact the proof of Proposition 4 shows that they extend to a multi-dimensional setting. In what follows, however, we work with an approximation that yields convenient closed form expressions, namely we replace the stereotyped distribution in (6) with :

$$f^\theta(\tilde{\varepsilon}|\varepsilon, G) = f(\tilde{\varepsilon}|\varepsilon) \left[ \frac{f(\tilde{\varepsilon}|\bar{\varepsilon}_G)}{f(\tilde{\varepsilon}|\bar{\varepsilon}_G)} \right]^\theta Z, \quad (8)$$

where  $\bar{\varepsilon}_G \equiv \mathbb{E}[\varepsilon|G]$  denotes the average income type in group  $G$ . Thus, we replace the group-conditioned distribution  $f(\tilde{\varepsilon}|G)$  on the RHS of (6) with the distribution for the average group type,  $f(\tilde{\varepsilon}|\bar{\varepsilon}_G)$ . This is a minor departure: it neither affects average income in  $G$  nor the MLRP condition that characterizes stereotypes. Its upside is that it buys tractability.

**Corollary 1** *Suppose that  $f(\tilde{\varepsilon}|\varepsilon)$  is normal with variance  $\sigma^2$ . Then, Equation (8) implies that a type  $\varepsilon$  identified with group  $G = L, U$  believes his future expected income to be:*

$$\varepsilon_G^\theta = \varepsilon + \theta(\bar{\varepsilon}_G - \bar{\varepsilon}_G). \quad (9)$$

As in Proposition 1, belief distortions are more severe when groups are more unequal, as reflected in a larger distance  $|\bar{\varepsilon}_U - \bar{\varepsilon}_L|$ , and when  $\theta$  is higher. To obtain this characterization we need that all distributions in Equation (8) belong to the same normal class. In the remainder of the paper, we model beliefs using Equation (9).<sup>6</sup>

By distorting beliefs, identity distorts policy evaluations. The expected utility of  $\varepsilon$  is:

$$W^{\varepsilon\theta}(\tau|G) = (1 + \varepsilon_G^\theta)(1 - \tau) + \nu(\tau - \frac{\varphi}{2}\tau^2), \quad (10)$$

which is also distorted by stereotypes. SIT, and its formalizations in economics (e.g., Shayo 2009), holds that identification causes individuals to favour ingroup members relative to the outgroup. Our belief-based mechanism creates similar effects, because  $\varepsilon_G^\theta$  depends on income polarization across classes. Indeed, Equation (10) can be written as follows:

$$W^{\varepsilon\theta}(\tau|G) = W^\varepsilon(\tau) + \theta \left[ W^G(\tau) - W^{\bar{G}}(\tau) \right], \quad (11)$$

where  $W^G(\tau)$  refers to the rational welfare of an individual with expected income equal to the group average,  $\bar{\varepsilon}_G$ . An individual may support a policy benefitting ingroups at the expenses of outgroups, increasing the gap  $W^G(\tau) - W^{\bar{G}}(\tau)$ , even if it reduces own welfare  $W^\varepsilon(\tau)$ . The

---

income and social mobility may be due to many factors, including local conditions. Our point is that when class conflict becomes salient (or social classes are primed), individuals identify with their class and misperceptions become class-based.

<sup>6</sup>We can avoid the nuisance of negative income levels by assuming that income is lognormally distributed.

intuition is that such policies focus the individual on states that are *distinctively* good for his group, also in the sense that the competing group is worse off.<sup>7</sup>

### 3.2 Actual and Perceived Polarization

We now illustrate some implications of our model for policy preferences. Let  $\tau^{\varepsilon\theta}$  denote the distorted bliss point of type  $\varepsilon$ , obtained by replacing mean income  $\varepsilon$  with  $\varepsilon_G^\theta$  in (4). Then, the difference in the distorted bliss points of an average rich and an average poor voter is:

$$\tau^{L\theta} - \tau^{U\theta} = (\tau^L - \tau^U)(1 + 2\theta) \quad (12)$$

where  $\tau^G$  is the rational bliss point of the average member of group  $G$ , and  $\tau^{G\theta}$  is the distorted bliss point of the same average group member. Stereotypes boost polarization. The lower class exaggerates the risk of poverty and demands too much redistribution, the upper class is too optimistic and demands too little redistribution. If exposure to digital media such as Facebook or Instagram strengthens stereotypical thinking  $\theta$ , then it also leads to more polarization. Although rational bliss points are not directly observable, Equation (12) has empirical content: it implies that polarization over redistribution should increase when social classes (or conflicting parties) are primed.<sup>8</sup>

Our theory also implies that voters should hold an exaggerate perception of partisan polarization. This is consistent with Westfall et al. (2015) and Bordalo et al. (2016), who show that US voters perceive an exaggerated difference in the policy views of Democrats and Republicans. To see how identity produces this effect, consider two groups, *Left* and *Right*, where *Left* contains a higher share of lower class voters than *Right*. These groups can capture parties with opposite ideological positions, or they could be defined in other ways. Denote by  $\tau^{Left,\theta}$  and  $\tau^{Right,\theta}$  the actual average positions (distorted by stereotypes) of the *Left* and

---

<sup>7</sup>This property holds in general. Suppose that  $f(u|q)$  is the distribution of utility at policy  $q$  for the individual,  $f(u|q, G)$  is the distribution of ingroup  $G$  utility and  $f(u|q, \overline{G})$  is the distribution of outgroup  $\overline{G}$  utility. The perceived expected utility under  $q$  according to the formula in footnote 5 is equal to:

$$\mathbb{E}^\theta(u|q, G) = \mathbb{E}(u|q) + Z \cdot cov\left(u, \left[\frac{f(u|q, G)}{f(u|q, \overline{G})}\right]^\theta\right),$$

where  $\mathbb{E}(u|q)$  is the rational expectation, and  $Z$  a normalizing constant. Once again, policies that increase the probability for the outgroups to achieve high utility states cause the ingroup to focus on low utility outcomes. These policies are thus disliked even if they increase true utility  $\mathbb{E}(u|q)$ .

<sup>8</sup>Alesina et al. (2018a) show demand for redistribution increases when subjects are provided pessimistic information about mobility, but this effect is only present for left wing respondents. Similarly, Kuziemko et al. (2015) show that informing low income individuals about their income rank in society increases their demand for inheritance taxes. The authors view their findings as reflecting information provision, but our analysis suggests that they may also be due to the fact that information primes class identity. We leave it to future work to assess this hypothesis also through direct measurement of identity.

*Right* groups. Using (2) to assess the distribution of beliefs in  $G = \text{Left}, \text{Right}$  we find:

**Corollary 2** *Denote by  $\hat{\tau}^{\text{Left},\theta}$  and  $\hat{\tau}^{\text{Right},\theta}$  the perceived average positions of the Left and the Right. When  $\theta > 0$ , perceived polarization is excessive,  $\hat{\tau}^{\text{Left},\theta} > \tau^{\text{Left},\theta} > \tau^{\text{Right},\theta} > \hat{\tau}^{\text{Right},\theta}$ .*

When thinking about the *Left*, beliefs overweigh the overly represented lower class voters. When thinking about the *Right*, beliefs overweigh the overly represented upper class voters. Thus, the left is stereotyped as being too redistributionist, the right as being too laissez faire. The "us versus them" logic of group conflict enhances perceived partisan polarization. Bordalo et al. (2019) provide support for this prediction. They show that the gap between perceived and actual polarization between Democrats and Republicans is larger on policy issues where party stereotypes are further apart in the precise sense of the likelihood ratio in (1).

Gentzkow (2017) shows that in the US perceived polarization and distrust of political rivals (so called affective polarization) has increased more than actual divergence in policy views, consistent with the above result. A contributing factor may be a rise in  $\theta$ , fueled by stronger stereotyping in traditional or social media. Ahler (2018) shows that correcting misperceptions about the out-party also reduces affective polarization.

### 3.3 Political Equilibrium

Consider the implications of identity for policy. We assume, as in standard models of probabilistic voting, that two candidates commit to policy platforms ahead of the elections in order to maximize the probability of winning (cf. Persson and Tabellini 2000). To isolate the role of beliefs, we assume that all voters have the same degree of mobility across parties. In this case, the equilibrium policy maximizes perceived utilitarian welfare:

$$\tau^* = \arg \max_{\tau} \int W^{\varepsilon\theta}(\tau|G) dH(\varepsilon).$$

By replacing (10) in the above objective, we obtain that the equilibrium policy satisfies:

$$\tau^* = \tau^0 - \theta \frac{(\bar{\varepsilon}_L - \bar{\varepsilon}_U)(\pi_L - \pi_U)}{\nu\varphi} \quad (13)$$

where  $\tau^0$  is the socially optimally policy defined in (5) and  $\pi_G = \Pr(\varepsilon \in G)$  is the population share of group  $G = L, U$ . Since  $(\bar{\varepsilon}_L - \bar{\varepsilon}_U) < 0$ , we obtain the following immediate result:

**Proposition 2** *If the Lower (Upper) class are the majority, then the equilibrium tax is larger (smaller) than in the utilitarian benchmark. Else, the equilibrium tax is socially optimal.*

Due to distorted beliefs, the lower class demands too much taxation and the upper class too little taxation. These effects tend to offset each other, because political competition leads



parties to converge to the average voter. On net, however, policy is distorted in favor of the more numerous group. Here we realistically assume that the lower class voters outnumber the upper class ones (i.e., the threshold  $\hat{\epsilon}$  is large enough). The direction of the policy distortion, though, is not key. The important point is that beliefs and preferences are shaped by identity, which alters the effect of economic change, as we discuss below.<sup>9</sup>

### 3.4 Discussion

In our model, identity amplifies conflict according to the multiplier:

$$\text{conflict} \Rightarrow \text{identity} \Rightarrow \text{beliefs} \Rightarrow \text{conflict}.$$

Conflict among heterogeneous social groups polarizes beliefs. Each group may have a majority of moderates but their perceived, stereotypical, types are more extreme. As identified voters move toward these perceived group positions, they become more extreme, too. By changing beliefs, conflict breeds more conflict, which amplifies economic shocks. Indeed, in Equation (13) stereotypes  $\theta$  amplify the effect of inequality ( $\bar{\epsilon}_U - \bar{\epsilon}_L$ ). Furthermore, as we will see in Sections 4 and 5, economic shocks can cause prevalent identities to drift away from class, altering the demand for redistribution in unconventional ways. In this way, identity can trigger large changes in beliefs and political outcomes.

In our simple model, the demand for redistribution is distorted by a clash between lower and upper class voters. To improve realism, we could allow for the middle class to act as a third group of identification, consistent with survey data showing that many people self-classify as middle class. We conjecture that the logic of stereotypes would cause voters in this group to shrink their perceived incomes toward the middle, and in particular toward an income level that is slightly lower than their true class average. This would create a large and cohesive group supporting redistribution, consistent with the intuition of Director’s law. We expect the broad qualitative properties of our model to be robust to this modification.<sup>10</sup>

---

<sup>9</sup>A large literature (eg. Alesina and Glaeser 2005) studies why in certain countries demand for redistribution is puzzlingly low. Here we do not analyze this phenomenon, but identity may play a role. As we discuss in the next section, if the poor tend to be an ethnically distinct minority, middle-low income individuals may be reluctant to identify with the lower class. If they identify instead with the upper-middle class, they will have an exaggerated perception of upward income mobility. These voters are thus naturally hostile to taxation, the more so the larger is inequality ( $\bar{\epsilon}_U - \bar{\epsilon}_L$ ). In other words, our model directs attention to how groups of identification are historically defined.

<sup>10</sup>Think of each voter in one of the three classes (lower, middle, upper) as slanting his income prospect toward the stereotype of his group, defined relative to the income distribution in the rest of society. The lower and upper class still feature stereotypically low and high incomes, respectively. When thinking about the middle class, then, two things must be noted. First, the average middle class income is lower than average income in society (due to right skewness of the income distribution). Second, income distribution within the middle class has thinner tails than in the rest of society. Thus, middle class voters likely exhibit a moderate downward distortion of their income but also converge to a tighter distribution.

Another simplification here is that political parties do not have partisan or ideological positions, and voters identify with social groups. Historically, social groups have been important, but identification with parties clearly also plays a role. This issue can be studied by introducing partisan politicians in the model. In such a model, divergent party platforms may separate society into two conflicting groups (e.g. *Left* and *Right*, as above), and endogenous identification with these groups may reinforce perceived and actual conflict, in turn fostering divergence. This mechanism offers an alternative perspective on extremist political platforms relative to Glaeser et al. (2005). We leave this issue to future work.

Finally, we have modelled beliefs over subjective conditions. Exactly the same formalism, and similar results, obtain when we apply Equation (6) to uncertainty about aggregate variables. In such a case, however, group based stereotypes are best modelled as reflecting ideological differences between groups, rather than differences in the individual economic prospects of group members. The distinction between distorted beliefs over subjective vs aggregate variables is thus related to the distinction between identification with social vs political groups (i.e. to social vs political identities).

## 4 Economic vs. Cultural Identity

How do individuals form their social identity? This question arises whenever political conflict also concerns non-economic dimensions, such as civil rights, religion, etc, so that class identity may be displaced by more salient cleavages. We now show that in this case non-economic changes, such as immigration or the diffusion of education, can reduce class cohesion and induce voters to identify with their cultural group. By altering voters' stereotypes, this new identity enhances cultural disagreement while reducing disagreement over redistribution.

To capture this mechanism, we introduce another policy instrument,  $q$ , which we interpret as a civil rights or an immigration policy, with higher  $q$  denoting a more progressive policy. We say that conflict over  $q$  is "cultural" to distinguish it from "economic" conflict over taxes  $\tau$ . Fitting real world policies in this dichotomy is a matter of degree. Taxes may be used to finance cultural programs, and culturally charged policies like immigration have economic effects. We define cultural policies as those having strong effects on the values and norms of society, but more limited effects on income distribution.<sup>11</sup>

Individual preferences over  $q$  are represented by the quadratic loss  $\frac{\kappa}{2} (q - \tilde{\psi})^2$ , where  $\kappa > 0$  captures the weight of policy  $q$  relative to taxation  $\tau$ . Individuals are uncertain about the consequences of cultural policy, which we capture by assuming that  $\tilde{\psi}$  is distributed according

---

<sup>11</sup>Card, Dustmann and Preston (2012) study attitudes toward immigration in the European Social Survey. 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.

to pdf  $z(\tilde{\psi}|\psi)$ , where  $\psi$  denotes the individual-specific mean. A voter with higher  $\psi$  is more progressive, and the average of  $\psi$  in the population is zero. Differences in  $\psi$  could be due to factual disagreement (over the effect of same-sex adoption, the number of immigrants, etc.), or to different value judgments. Beliefs about  $\tilde{\psi}$  can be stereotyped. To use the convenient closed form of equation (9) we also assume  $z(\tilde{\psi}|\psi)$  to be normal.

Now a voter type is summarized by his expected social progressiveness-income pair  $(\psi, \varepsilon)$ , distributed in society according to the joint cdf  $H(\psi, \varepsilon)$ . Here we assume that  $\psi$  and  $\varepsilon$  are uncorrelated in the population, but this assumption is relaxed in Section 5.

The rational expected utility of type  $(\psi, \varepsilon)$  from the policy vector  $(\tau, q)$  is equal to:<sup>12</sup>

$$W^{\varepsilon\psi}(\tau, q) = (1 + \varepsilon)(1 - \tau) - \nu\left(\tau - \frac{\varphi}{2}\tau^2\right) - \frac{\kappa}{2}(q - \psi)^2, \quad (14)$$

The bliss point for taxes only depends on expected income, as in (4), while the bliss point for civil rights only depends on social progressiveness,  $q^\psi = \psi$ . The socially optimal tax rate  $\tau^0$  is still given by (5), while the socially optimal civil rights policy is  $q^0 = 0$ .

## 4.1 Politically Relevant Social Identities

The identity of type  $(\psi, \varepsilon)$  can be defined either based on income, or on culture. As before, the upper and the lower class are separated by the income threshold  $\hat{\varepsilon}$ . Cultural groups are defined by a threshold  $\hat{\psi}$ , so that a voter is socially conservative  $SC \equiv \{\psi|\psi < \hat{\psi}\}$  or socially progressive,  $SP \equiv \{\psi|\psi \geq \hat{\psi}\}$ . To simplify, we assume that groups are divided at the mean  $\hat{\psi} = \hat{\varepsilon} = 0$ . In economic conflict, the majority continues to be lower class,  $\pi_L > \pi_U$ . In cultural conflict, it is socially conservative,  $\pi_{SC} > \pi_{SP}$ .

Figure 2 below illustrates the space of possible types. On the vertical axis,  $\hat{\psi}$  separates socially conservatives (bottom) from socially progressive (top). On the horizontal axis,  $\hat{\varepsilon}$  divides the upper and lower classes.

Figure 2

We assume that an individual can have one of two identities: economic or cultural. For instance, a type in the lower-left corner of Figure 2 is poor ( $\varepsilon < 0$ ) and socially conservative ( $\psi < 0$ ). He can thus identify either with the lower class, or with the socially conservatives, but not with both at the same time (nor with even more granular groups).

This simplifying assumption is realistic. It reflects many instances of social conflict, in which relevant categories for political action are few, broad, and historically determined.

---

<sup>12</sup>We omit for simplicity the additive constant  $k\text{Var}(\psi)$  that obtains when computing  $\mathbb{E}(q - \tilde{\psi})^2$ .

Lipset and Rokkan (1967) argue that much political change in recent history can be accounted for by switching identity and conflict across such broad groups. Salient examples are "rich/elite vs. poor/masses", "traditional/religious vs. progressive/secular", "center vs. periphery" etc. In our two dimensional setting, the natural groups are defined along the class and cultural divides. A fuller understanding of the emergence and structure of identitarian groups, and the role of politicians in creating them, is an important problem for future work.<sup>13</sup>

## 4.2 Meta Contrast Similarity and Identification

How does a voter select the dimension of identification? According to the meta-contrast principle discussed in Section 2, identity is formed so as to optimize the tradeoff between: i) maximizing similarity between oneself and the in-group, and ii) maximizing contrast between the in-group and the out-group. Since we consider political identities, policy preferences are the natural measure of similarity. Thus, we assume that the individual identifies with the group, class or culture based, that is closer to him in terms of overall policy preferences, but also that maximizes the divergence of interests with the outgroup.

We capture conflict between a voter  $(\varepsilon, \psi)$  and his ingroup by the welfare loss that the voter experiences when moving from his bliss point policy  $(\tau^\varepsilon, q^\psi)$  to the policy  $(\tau^G, q^G)$  preferred by the average type in his ingroup  $G$ :

$$W^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) - W^{\varepsilon\psi}(\tau^G, q^G). \quad (15)$$

Similarly, intergroup conflict is captured by the loss borne by the average ingroup type when moving to the bliss point of the average outgroup member,  $(\tau^{\bar{G}}, q^{\bar{G}})$ :

$$W^G(\tau^G, q^G) - W^G(\tau^{\bar{G}}, q^{\bar{G}}) \quad (16)$$

Implicitly, ingroups and outgroups are measured by the average types in each group. Note that these conflicts are defined on the basis of rational bliss points. Footnote 15 below shows that little changes if instead they are computed from the stereotyped bliss points.<sup>14</sup>

Combining these criteria, we define the relative distance between  $(\varepsilon, \psi)$  and ingroup  $G$  as:

$$\Delta^{\varepsilon\psi}(G) = [W^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) - W^{\varepsilon\psi}(\tau^G, q^G)] - \lambda [W^G(\tau^G, q^G) - W^G(\tau^{\bar{G}}, q^{\bar{G}})]. \quad (17)$$

---

<sup>13</sup>The preeminence of broad political groups may be in part due to the fact that they render the "us versus them" nature of conflict more salient for the largest part of society, creating strong identification. In line with this intuition, footnote 16. Shows that identification with broad rather than granular groups is precisely one implication of our model of meta contrast similarity.

<sup>14</sup>Consistent with the idea that social categories and their representations are historically given, the average types of income classes or cultural groups do not depend on which voters identify with them. For instance, for a voter identifies with the lower class, the ingroups are all poor people and the outgroups are all rich people, regardless on how these people are actually identified.

The voter feels similar to his ingroups  $G$ , namely  $\Delta^{\varepsilon\psi}(G)$  is small, when his conflict with them is small, relative to the overall conflict between ingroups and outgroups. The role of intergroup conflict is intuitive. When class conflict is strong, a lower middle class individual perceives himself as more similar to the poor, because of contrast with the very rich. Likewise, when cultural or religious conflict is strong, many voters may feel socially conservative in contrast with the strongly progressive ones. Parameter  $\lambda \geq 0$  captures the importance of cross group conflict relative to ingroup congruence as a driver of identification.

To make the model tractable, we approximate relative distance by the expression:

$$\Delta^{\varepsilon\psi}(G) \simeq |\varepsilon - \bar{\varepsilon}_G| + \alpha |\psi - \bar{\psi}_G| - \lambda (|\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}}| + \alpha |\bar{\psi}_G - \bar{\psi}_{\bar{G}}|). \quad (18)$$

The appendix shows that (18) is a shorthand for a second order approximation of the right hand side of (17).<sup>15</sup> The relative welfare importance of cultural vs economic conflict is given by  $\alpha = \kappa\nu\varphi$ , where  $\kappa$  is the importance of civil rights,  $\nu$  is the marginal utility of public consumption and  $\varphi$  captures the severity of the tax distortions. If civil rights become more important ( $\kappa$  rises), or if income differences become a less important source of disagreement (because tax distortions  $\varphi$  or the value  $\nu$  of the public good rise), then  $\alpha$  rises and voters feel stronger similarity along culture than along class.<sup>16</sup>

Individual  $(\varepsilon, \psi)$  identifies with his economic class or cultural group to minimize relative distance  $\Delta^{\varepsilon\psi}(G)$ . As we show below, identification depends on the: i) welfare relevance of conflict, ii) contrast across opposing groups, and iii) individual features.

**Proposition 3** *There exist two thresholds  $\underline{\alpha}$  and  $\bar{\alpha}$  for the relative welfare importance of cultural versus economic conflict, with*

$$\underline{\alpha} \equiv \frac{\lambda |\bar{\varepsilon}_U - \bar{\varepsilon}_L| - \bar{\varepsilon}_L}{\lambda |\bar{\psi}_{SP} - \bar{\psi}_{SC}| + \bar{\psi}_{SP}} < \bar{\alpha} \equiv \frac{\lambda |\bar{\varepsilon}_U - \bar{\varepsilon}_L| + \bar{\varepsilon}_U}{\lambda |\bar{\psi}_{SP} - \bar{\psi}_{SC}| - \bar{\psi}_{SC}}, \quad (19)$$

---

<sup>15</sup>To obtain (18), we take a second order approximation of (15) around  $(\tau^\varepsilon, q^\psi)$ , and of (16) around  $(\tau^G, q^G)$ , and replace quadratic distance with absolute value. Little would change if we define the conflict of interest using the stereotyped (rather than rational) bliss points. In this case, equation (18) becomes:

$$\Delta^{\varepsilon\psi} \simeq |\varepsilon - \varepsilon_G| + \alpha |\psi - \psi_G| - \lambda (1 + 2\theta) (|\varepsilon_G - \varepsilon_{\bar{G}}| + \alpha |\psi_G - \psi_{\bar{G}}|),$$

so that the meta contrast term becomes more important when  $\theta$  is higher.

<sup>16</sup>By gauging Equation (17) one can see that higher  $\lambda$  also favors identification with broad rather than narrow groups. For instance, if  $\alpha$  is so large that cultural clash is strong, high  $\lambda$  makes it more fitting for a poor and socially conservative voter to identify with the broad socially conservative group, even though the latter includes some rich individuals. This broad group is in particular favored also over a more granular group of socially conservative and lower class voters if identification with such group is available. Intuitively, I am most congruent with a group consisting only of similar types, but cross-group contrast is increased by including in the in-group more extreme members than myself along the most salient issue. Thus, our restriction that the most granular groups are not available for identification is natural when cross-group contrast has a strong weight in determining the salience of alternative social categories.

and a threshold  $\infty > \hat{\lambda} > \min(\pi_U, \pi_{SP})$  for the relevance of group contrast such that:

i) If  $\lambda > \hat{\lambda}$ , all types  $(\varepsilon, \psi)$  identify with their class for  $\alpha < \underline{\alpha}$ , while they identify with their cultural group for  $\alpha > \bar{\alpha}$

ii) If  $\lambda < \min(\pi_U, \pi_{SP})$ , identification differs across types. If  $\alpha < \underline{\alpha}$ , types  $(\varepsilon, \psi)$  with high or low income identify with their class; only types with income close to  $\hat{\varepsilon} = 0$  identify with their cultural group. If  $\alpha > \bar{\alpha}$ , socially progressive and socially conservative types identify with their cultural group; only types with culture close to  $\hat{\psi} = 0$  identify with their class.

The pattern of identification depends on the relative importance  $\alpha$  of cultural vs income conflict. If conflict about civil rights is more important than redistributive conflict,  $\alpha$  is high, voters predominantly identify along the cultural dimension. Viceversa, if income conflict is more relevant, they predominantly identify based on income.

Critically, this effect is shaped by intergroup contrast, which determines the thresholds  $\underline{\alpha}$  and  $\bar{\alpha}$ . When income inequality  $|\bar{\varepsilon}_U - \bar{\varepsilon}_L|$  between classes is large,  $\underline{\alpha}$  and  $\bar{\alpha}$  are also large and it is easy for class identity to prevail. Cultural polarization may be significant, but individuals feel similar to the members of their income class, whichever their culture. When instead cultural contrast  $|\bar{\psi}_{SP} - \bar{\psi}_{SC}|$  is strong, stemming for instance from large opposing clusters in education or religion, the thresholds  $\underline{\alpha}$  and  $\bar{\alpha}$  are small and culture based identification is favored. Income inequality may be acute, but individuals perceive themselves to be similar to members of their cultural group, whichever their income.

Historical discussions of class cohesion echo these themes. In a 1851 letter, Friedrich Engels wrote that US workers were not drawn to class struggle because of the "rapidly growing prosperity of the country, which makes bourgeois conditions look like a beau ideal to them..." (Howe, 1977, p.105). In our model, prospects of social mobility for the lower class may be high enough that  $|\bar{\varepsilon}_U - \bar{\varepsilon}_L|$  is low. In another letter in 1893, Engels also blamed deep ethnic cleavages within the US working class. This point was also made by the labor historian Selig Perlman "The massive waves of immigration led to deep ethnic cleavages within the working class...It has therefore been extremely hard to achieve a unified class consciousness within the American working class" (Howe, 1977, p.111). This is captured by the case in which cultural clash  $|\bar{\psi}_{SP} - \bar{\psi}_{SC}|$  is high. American workers felt too dissimilar from fellows coming from different cultural and ethnic backgrounds.

Proposition 3 sheds light on what brings about changes in social and political identities. Suppose that the importance of cultural conflict  $\alpha$  increases, due for instance to an inflow of immigrants. This increases conflict among voters who wish to accommodate to the inflow, and those who oppose it. This growing conflict exerts two effects. First, it reduces the cohesiveness of income classes. Indeed, the distance between socially conservative and socially progressive class members becomes more relevant, enhancing within class heterogeneity. Second, it enhances the cohesiveness of cultural groups: individuals perceive greater similarity

with cultural ingroups in contrast to outgroups. If this effects swamps their dissimilarity in economic interests, identity switches to culture.

The strength of the latter effect depends on the weight of contrast  $\lambda$ . In case i), such weight is large and so everybody identifies along income or culture. In case ii) the weight of contrast is lower, so different individuals identify with different groups by weighing also their individual interests. Figure 3 illustrates equilibrium identification in both cases.

Figure 3

The upper row refers to the case  $\lambda > \hat{\lambda}$ , where everyone identifies along the same dimension: income (if  $\alpha \leq \underline{\alpha}$ ) or culture (if  $\alpha \geq \bar{\alpha}$ ). The lower row depicts the case  $\lambda < \min(\pi_U, \pi_{SP})$ , where different individuals identify along different dimensions. If  $\alpha \leq \underline{\alpha}$ , the area of class identification is larger, if  $\alpha \geq \bar{\alpha}$  cultural identification prevails. Note that voters with extreme income and culture are more "flexible": they feel intensely about both issues, so they readily switch identity to reduce dissonance along the most relevant dimension.<sup>17</sup>

### 4.3 Beliefs

For a given social identity, beliefs overweight the stereotypical features of the group with which the individual identifies. The only difference with Section 3 is that now identification influences beliefs in two dimensions: income prospects,  $\tilde{\varepsilon}$ , and preferences over civil rights,  $\tilde{\psi}$ . As a result, when identity changes, beliefs in both dimensions are affected.

To characterize beliefs, consider a type  $(\varepsilon, \psi)$ . His stereotyped beliefs are still formed according to Equation (8) in the previous section. We have:

**Proposition 4** *Suppose that  $f(\tilde{\varepsilon}|\varepsilon)$  and  $z(\tilde{\psi}|\psi)$  are both Gaussian. Then, the beliefs about income and civil rights of a type  $(\varepsilon, \psi)$  identified with group  $G$  are equal to  $\varepsilon_G^\theta$  and  $\psi_G^\theta$ , where:*

- 1) *If  $(\varepsilon, \psi)$  identifies with his class  $G = L, U$ , his perception of income is distorted,  $\varepsilon_G^\theta = \varepsilon + \theta(\bar{\varepsilon}_G - \bar{\varepsilon}_G)$ , while his perception of civil rights is non-distorted  $\psi_G^\theta = \psi$ .*
- 2) *If  $(\varepsilon, \psi)$  identifies with his cultural group  $G = SC, SP$ , his perception of income is non-distorted,  $\varepsilon_G^\theta = \varepsilon$ , while his perception of civil rights is distorted  $\psi_G^\theta = \psi + \theta(\bar{\psi}_G - \bar{\psi}_G)$ .*

When identity changes from income to culture, the entire beliefs system of voters changes. When individuals identify with their income class, they stereotype class differences in their beliefs about social mobility but not their cultural differences. Due to zero correlation between

---

<sup>17</sup>To determine how many voters identify along income or culture, we need an assumption on the distribution of  $(\varepsilon, \psi)$ . The appendix shows that, if both variables have a piecewise uniform density like in Figure 4 below, then the majority of voters is identified along income for  $\alpha \leq \underline{\alpha}$  and along culture for  $\alpha \geq \bar{\alpha}$ . Thus, even if some people stick to their original identities, a sufficiently large increase in the importance of cultural conflict  $\alpha$  turns the identity of the majority of voters from income to culture.

$\varepsilon$  and  $\psi$ , the distribution of  $\psi$  is identical across classes, so no stereotyped cultural differences arise. Likewise, when individuals identify with their cultural group, social progressiveness is distorted but beliefs about income are not. Indeed, social mobility is the same across socially conservatives and progressives.

Specifically, take a poor and socially conservative voter,  $\varepsilon < 0$   $\psi < 0$ . When this voter identifies with the lower class he underestimates social mobility  $\varepsilon_L^\theta = \varepsilon + \theta(\bar{\varepsilon}_L - \bar{\varepsilon}_U) < \varepsilon$ . His views on civil rights are undistorted  $\psi_L^\theta = \psi$ . When cultural conflict becomes strong, the same type switches to a socially conservative identity. He no longer perceives himself as a member of an economically disadvantaged group, his views on social mobility improve  $\varepsilon_{SC}^\theta = \varepsilon$ , while his social conservatism increases  $\psi_{SC}^\theta = \psi + \theta(\bar{\psi}_{SC} - \bar{\psi}_{SP}) < \psi$ . As a result, the voter demands less redistribution but more constraints on civil rights.

The assumption of zero correlation helps us illustrate our results in the starkest way, but in Section 5 we show that results are similar when the traits are positively, though not too strongly, correlated. In this case, lower class identification also leads to some excess cultural conservatism, but the latter is enhanced when direct cultural identification takes place.

Switches in prevailing identities exert far reaching effects on conflict and on polarization. When many voters change their identity, society changes as follows.

**Corollary 3** *When  $\alpha$  rises from below  $\underline{\alpha}$  to above  $\bar{\alpha}$ , many voters switch identity from class to culture. Thus, beliefs over civil rights become more polarized, while polarization in attitudes towards redistribution dampens. The opposite occurs if  $\alpha$  switches in the reverse direction.*

Some recent findings suggest that economic or social change may affect polarization in ways related to this result. Using ANES data, Bordalo et al. (2019) find that actual and perceived polarization between liberals vs conservatives on domestic issues increased after the end of the cold war, and it temporarily subsided after the September 11 attacks. These changes are more pronounced on the issues that are more stereotypical of the liberal vs conservative divide, in the precise sense of the likelihood ratio in (1). One explanation for this finding is that the end of the cold war increased the salience of domestic divisions, leading to a liberal versus conservative divide. By contrast, the September 11th attacks reduced the salience of such domestic divisions, creating the reverse effect.

The evidence in Alesina et al. (2019) and Alesina et al. (2018a) is also consistent with our theoretical results. Using survey data, these papers show that respondents are less in favor of redistribution if they are exposed to larger inflows of immigrants, or if they are primed to think about immigrants. Similarly, Tabellini (2018) shows that US cities that received more immigrants between 1910 and 1930 became politically more conservative and anti-immigrants, and reduced local government spending. One interpretation of these facts is that individuals are averse to redistributing towards strangers. But an equally plausible mechanism is that



these treatments increase the salience of cultural identities relative to economic class. Socially conservative and low-middle class voters become more focused on their native identities. Thus, they discount the benefit from redistribution because they no longer perceive themselves as typical members of low-income groups. If this is true, our model predicts that the same respondents should become more socially conservative also in other dimensions.

#### 4.4 Equilibrium Policy

We now study how identity and belief distortions feed into equilibrium policy. Denote by  $W^{\varepsilon\psi\theta}(\tau, q|G)$  the expected utility of type  $(\varepsilon, \psi)$  when he identifies with  $G$ , so that his beliefs about income and culture are distorted by stereotypes,  $\varepsilon_G^\theta$  and  $\psi_G^\theta$ .<sup>18</sup>

Let  $\phi_G(\alpha)$  denote the fraction of the population that identifies with group  $G$ . It is a function of  $\alpha$  because of endogenous identities. Then the equilibrium policy satisfies:

$$(\tau^*, q^*) = \arg \max_{\tau, q} \sum_{G=L, U, SC, SP} \phi_G(\alpha) \int W^{\varepsilon\psi\theta}(\tau, q|G) dH(\psi, \varepsilon). \quad (20)$$

The first order conditions of the problem imply:

$$\tau^* = \tau^\circ + \frac{\theta(\bar{\varepsilon}_U - \bar{\varepsilon}_L)[\phi_L(\alpha) - \phi_U(\alpha)]}{\nu\varphi}, \quad (21)$$

$$q^* = q^\circ - \theta(\bar{\psi}_{SP} - \bar{\psi}_{SC})[\phi_{SC}(\alpha) - \phi_{SP}(\alpha)]. \quad (22)$$

where  $\tau^\circ$  and  $q^\circ$  denote the socially optimal policies.

Stereotypes,  $\theta > 0$ , distorts equilibrium policy. Taxation is excessive relative to the social optimum if those identified with the lower class outnumber those identified with the upper class ( $\phi_L(\alpha) - \phi_U(\alpha) > 0$ ). Civil rights are too restricted if the identified social conservatives exceed the identified socially progressives ( $\phi_{SC}(\alpha) - \phi_{SP}(\alpha) > 0$ ). This implies that policy distortions depend on parameter  $\alpha$  because it determines the prevalent identities.

To do comparative statics, we must specify the distribution of voter types  $(\varepsilon, \psi)$  and compute how shocks affect the identification shares  $\phi_G(\alpha)$ . For tractability, we assume that each trait  $x = \varepsilon, \psi$  is distributed according to the same "piecewise uniform" distribution displayed in Figure 4 below.

Figure 4

This distribution has mean zero, and  $\bar{x} > \underline{x} > 0$  ensures that it is left skewed, so that poor and socially conservative voters are in majority. The appendix then proves that if someone identifies with his class, the lower class identified voters outnumber the upper class

<sup>18</sup>To obtain  $W^{\varepsilon\psi\theta}(\tau, q|G)$ , just replace in (14) the true  $\varepsilon$  and  $\psi$  with their stereotyped versions.

identified ones. And similarly, if someone identifies with his cultural group, then conservative identification dominates progressive identification.

These results are important to characterize equilibrium policies:

**Proposition 5** *Equilibrium taxation is weakly excessive and civil rights are weakly underprovided,  $\tau^* \geq \tau^0$  and  $q^* \leq q^0$ , with at least one strict inequality. There are two cases:*

*i) If  $\lambda > \hat{\lambda}$ , all voters identify along either class or culture. Thus, only taxation is distorted if  $\alpha < \underline{\alpha}$ , while only civil rights are distorted if  $\alpha > \bar{\alpha}$ .*

*ii) If  $\lambda < \min(\pi_U, \pi_{SP})$ , identification is mixed, so both taxes and civil rights are distorted. Moreover, policy distortions are higher in the prevalent dimension of identification.*

*In both cases, a change from  $\alpha < \underline{\alpha}$  to  $\alpha > \bar{\alpha}$  reduces both  $\tau^*$  and  $q^*$ , so that taxes become less distorted and civil rights more distorted.*

Policy is distorted by stereotypes. Lower class-identified voters exaggerate the benefit of redistribution, socially conservative-identified voters exaggerate the risks of liberal policies. The upper class and socially progressive voters entertain the opposite belief distortions, but they are a minority. Thus, redistribution is excessive, civil rights too limited.

This result can shed light on why certain non economic shocks can influence the demand for redistribution. An increase in the importance of cultural conflict  $\alpha$ , driven for instance by an upward trend in education or by a large inflow of immigrants, changes prevailing identities from class to culture, promoting more restrictive civil rights and less redistributive policies. In a rational model, the effect of the same social change would be less drastic, but also qualitatively different. Insofar as income inequality or tax distortions do not change, demand for redistribution by rational voters would be stable. This is not true in our model, in which cultural conflict causes demand for redistribution to drop. In fact, in our model such demand may decrease even if income inequality has meanwhile increased, provided of course that it stays low enough that class conflict does not become salient.

## 4.5 Discussion

Our model captures a chain reaction whereby, by influencing beliefs, changes in identity alter the effects of economic shocks:

$$\text{economic and social change} \Rightarrow \text{group identities} \Rightarrow \text{beliefs} \Rightarrow \text{policies}$$

Suppose that class based identification was initially prevalent. As the importance of cultural conflict  $\alpha$  rises, or if cultural contrast between groups becomes starker, for a while nothing happens. But if  $\alpha$  keeps rising relative to the thresholds defined above, social identities switch and become based on culture. When this occurs, beliefs and policies react accordingly.

Given plausible assumptions on the distribution of  $\varepsilon$  and  $\psi$ , redistribution and its distortions fall, while distortions from too limited civil rights increase. Thus, sustained changes in the socioeconomic environment can eventually induce sudden and sharp political change.

This mechanism can shed light on why the growing importance of cultural conflict in the US and other Western democracies has occurred in tandem with a drop in the demand for redistribution, despite increasing income inequality. Modern societies have become much more liberal and more permissive than in the 1950s and 1960s, partly due to the influence of an increasing mass of educated and progressive voters (cf. Fukuyama 2018, Goodhart 2017). More traditional social strata did not alter their value system to the same extent, and resent the new status quo. This can be viewed as increasing the importance of cultural policy  $\alpha$  relative to taxes. In recent years, a contributing factor may have also been increased mobility of capital, which have likely increased tax distortions  $\varphi$  and hence  $\alpha$ . According to our theory, these transformations eventually trigger identity switches, exacerbating conflict over immigration and civil rights, while income related redistributive conflict is dampened. In section 6 we offer some evidence consistent with this interpretation.

In our model, politicians only care about winning the election. In a richer setting with partisan politicians who also care about policy outcomes (as in Persson and Tabellini 2000, ch. 5), identity switches have additional implications. Suppose that there are two politicians: the left wing candidate likes redistribution and socially progressive policies (i.e. he is close to voters located in the upper left corner of Figure 2), while the right wing candidate likes limited government and conservative policies (he is close to voters located in the lower right corner). If prevalent identification is on class, as in panel A of Figure 3, then the left wing candidate predominantly draws the support of lower class-identified voters, while the right wing candidate is predominantly voted by those identified as upper class. Now suppose that candidate positions remain fixed, but the importance of cultural conflict  $\alpha$  rises to the point where identification switches to culture, as in panel B of Figure 3. Social groups, then, reshuffle. Culturally progressive and upper class voters are attracted by the left wing candidate, while socially conservative and lower class voters switch to the right wing candidate. Rational voters would react similarly to a rise in  $\alpha$  (or to shifting candidate positions). But endogenous identity and stereotyping amplify the effect of a rise in  $\alpha$  and make it more sudden. As discussed by Piketty (2018), this sorting of voters across parties according to their cultural views is exactly what has taken place in the US, UK and France.

## 5 Non Conventional Effects of Economic Shocks

In the model of Section 4, adverse economic shocks make redistributive conflict more salient, favoring class identification and increasing the demand for redistribution by the less well off.

Yet, the evidence discussed in the Introduction points in the opposite direction: in many recent instances, economic distress has failed to harness or has even decreased the demand for redistribution. We now show that our model generates such non-conventional responses when economic shocks are significantly correlated with other carriers of identity.

We illustrate this mechanism with two examples. Subsection 5.1 considers shocks like skill biased technical change, that increase the correlation between income and education, and hence also make income and social progressiveness more strongly positively correlated. We show that, by rendering cultural groups more economically homogeneous, such shocks can: i) cause identities to switch from class to culture, thereby ii) reducing the demand for redistribution.

Subsections 5.2 and 5.3 add a new dimension of economic conflict, between winners and losers from globalization. As globalization becomes more important, identity switches from class to a new nationalist vs cosmopolitan dimension. As a result, the losers from globalization become more protectionist and demand less redistribution. If the cleavage over globalization is correlated with the cultural divide, as suggested by survey evidence, then this identity switch has far reaching effects on beliefs. In particular, a purely economic shock can also amplify conflict over non-economic dimensions, like immigration or civil rights. This prediction is broadly consistent with existing empirical results (Autor et al. 2017, Colantone and Stanig 2017). Section 6 offers more specific evidence in its support.

## 5.1 Economic Shocks and Correlated Traits

Consider the two dimensional model of Section 4, but assume that individual income  $\varepsilon$  and social progressiveness  $\psi$  are correlated in the population. This means that economic classes differ not only in their income but also in their average social progressiveness, and cultural groups have different mean incomes. To obtain simple closed form solutions, we assume that  $\varepsilon$  and  $\psi$  are jointly normally distributed, with mean 0, the same variance and correlation coefficient  $\rho$ . This implies that the average social progressiveness of the lower and upper classes is given by  $\bar{\psi}_L = \rho\bar{\varepsilon}_L$  and  $\bar{\psi}_U = \rho\bar{\varepsilon}_U$  respectively. Similarly, the average income of socially progressive and conservative voters is given by  $\bar{\varepsilon}_{SP} = \rho\bar{\psi}_{SP}$  and  $\bar{\varepsilon}_{SC} = \rho\bar{\psi}_{SC}$  respectively.

We plausibly assume that income contrast is larger among economic classes than cultural groups, and the reverse is true for cultural contrast. Formally,  $|\bar{\varepsilon}_L - \bar{\varepsilon}_U| > |\bar{\varepsilon}_{SC} - \bar{\varepsilon}_{SP}|$  and  $|\bar{\psi}_{SC} - \bar{\psi}_{SP}| > |\bar{\psi}_L - \bar{\psi}_U|$ . This condition is satisfied if the correlation coefficient is not too close to 1 in absolute value, namely:

$$\frac{1}{|\rho|} > \frac{|\bar{\varepsilon}_L - \bar{\varepsilon}_U|}{|\bar{\psi}_{SC} - \bar{\psi}_{SP}|} > |\rho|, \quad (23)$$

which we assume throughout.<sup>19</sup>

Finally, we assume that  $\lambda \rightarrow \infty$ , so that group contrast is the unique driver of identification for every individual. This assumption greatly simplifies the analysis of the correlated case.

### 5.1.1 Identities

By Equation (23), identification is determined as follows:

**Proposition 6** *(i) Suppose that (23) holds. There is a threshold*

$$\hat{\alpha} \equiv \frac{|\bar{\varepsilon}_L - \bar{\varepsilon}_U| - |\rho| |\bar{\psi}_{SC} - \bar{\psi}_{SP}|}{|\bar{\psi}_{SC} - \bar{\psi}_{SP}| - |\rho| |\bar{\varepsilon}_L - \bar{\varepsilon}_U|} > 0 \quad (24)$$

*for the relative importance of cultural conflict, such that everybody identifies with their economic class if  $\alpha < \hat{\alpha}$ , and everyone identifies with their cultural group if  $\alpha > \hat{\alpha}$ .*

*(ii) As income and social progressiveness become more correlated,  $|\rho|$  increases, people are more likely to identify along the dimension where group contrast is starker. Formally,*

$$\frac{\partial \hat{\alpha}}{\partial |\rho|} \begin{matrix} \geq \\ \leq \end{matrix} 0 \text{ as } |\bar{\varepsilon}_L - \bar{\varepsilon}_U| \begin{matrix} \geq \\ \leq \end{matrix} |\bar{\psi}_{SC} - \bar{\psi}_{SP}|.$$

Part (i) says that, as in Proposition 3, people identify with their class if redistribution is important relative to culture, namely if  $\alpha$  is low. Part (ii) says that higher correlation between  $\varepsilon$  and  $\psi$  favors cultural identification, namely it reduces  $\hat{\alpha}$ , if and only if cultural contrast  $|\bar{\psi}_{SC} - \bar{\psi}_{SP}|$  is larger than class contrast  $|\bar{\varepsilon}_L - \bar{\varepsilon}_U|$ .

To see why, suppose that income and social progressiveness become more positively correlated. This makes cultural groups more economically homogeneous, and economic groups more culturally homogeneous. Which of the two identities prevails? If group contrast is stronger in culture than in income, higher  $\rho$  induces cultural identification. The reason is that cultural groups are more efficient carriers of social conflict, and their greater economic homogeneity makes them more cohesive. If instead contrast is higher in income than in culture, then higher correlation reinforces class identity.

This result can shed light on the political consequences of recent economic shocks. In the US and other advanced democracies, phenomena like skill biased technological change have predominantly hurt workers with low education (Autor 2019). Since education is one of the main determinants of attitudes towards civil rights and immigration, these shocks can be seen as having raised  $\rho$ : socially conservative voters are now poorer, while socially progressives are richer. If, due to historical reasons, cultural conflict is starker than class conflict, these

---

<sup>19</sup>In our proofs we allow  $\rho$  to be negative, to show how the model can also be applied to settings where this assumption is more realistic or convenient.

economic shocks cause poor and socially conservative voters to identify with their cultural group rather than with their economic class. We now consider the effect of these changes in policy demand and in equilibrium policies.

### 5.1.2 Beliefs

Under correlated traits, beliefs are distorted in both dimensions. If  $\rho > 0$ , the poor are less socially progressive than the rich,  $\bar{\psi}_L < \bar{\psi}_U$ . Thus, a voter identified with the lower class exaggerates his downward income risk (cf. Proposition 4), but he also exaggerates his social conservatism, since  $\psi_L^\theta = \psi + \theta (\bar{\psi}_L - \bar{\psi}_U) = \psi + \rho\theta (\bar{\varepsilon}_L - \bar{\varepsilon}_U) < \psi$ .

If civil rights become more important and  $\alpha$  rises above  $\hat{\alpha}$ , the same voter abandons class identification and identifies as a socially conservative. His belief distortions remain directionally the same, but their relative magnitude changes in a crucial way. His exaggeration of social conservatism increases to  $\psi_{SC}^\theta = \psi + \theta (\bar{\psi}_{SC} - \bar{\psi}_{SP}) < \psi_L^\theta$ , where the last inequality follows from Equation (23). By contrast, his exaggeration of downward income risk falls, since  $\varepsilon_{SC}^\theta = \varepsilon + \theta (\bar{\varepsilon}_{SC} - \bar{\varepsilon}_{SP}) = \varepsilon + \rho\theta (\bar{\psi}_{SC} - \bar{\psi}_{SP}) > \varepsilon_L^\theta$ . As a result, shocks hurting unskilled workers cause them to demand less redistribution.

By extending this reasoning to all voters types, we obtain the following result.

**Corollary 4** *i) If (23) holds, then beliefs about income (civil rights) are more distorted when income (culture) is the dimension of identification. ii) The distortions in beliefs about income and civil rights have the same sign if and only if  $\varepsilon$  and  $\psi$  are positively correlated.*

The intuition for this result is the same as in Subsection 4.3. A switch from class to cultural identity brings  $L$  and  $U$  voters in the same group. This mixing reduces income contrast across groups, dampening income stereotypes and belief distortions. The reverse happens with regard to beliefs over civil rights, where stereotypes are enhanced by greater contrast between cultural groups. The Appendix characterizes the political equilibrium and shows that it remains qualitatively similar to Proposition 5 above.

Thus, adverse economic shocks that hit specific segments of the population - say less educated and socially conservative voters - can paradoxically move voters' beliefs and policy demands to the right rather than to the left, despite an increase in inequality.<sup>20</sup> This happens because the new poor feel a closer affinity with their socially conservative cultural group rather than with their economic class. Goodhart (2017) and Frank (2005) provide anecdotal evidence supporting this mechanism.

---

<sup>20</sup>Of course, skilled bias technical change may have arguably increase also overall inequality and thus class contrast  $|\bar{\varepsilon}_L - \bar{\varepsilon}_U|$ . Our mechanism is still valid, however, provided that the latter effect is sufficiently weak relative to the change in correlation  $\rho$ .

## 5.2 Globalization

Neglect for now cultural conflict, and consider a small open economy with a continuum of individuals of size 1 and two traded goods:  $x$  is the exported good, and  $m$  is the imported good. The utility of a generic voter of type  $i$  (to be defined below) is:

$$w^i = x^i + U(m^i) + \nu g, \quad (25)$$

where  $x^i$  and  $m^i$  denote his consumption of the exported and imported good respectively,  $U(\cdot)$  is a concave quadratic utility function,  $\nu g$  is the utility of public consumption  $g$ , where  $\nu > 1$  and large. The price of the exported good is 1, the price of the imported good is  $p$ . Thus, type  $i$ 's budget constraint (based on expected income  $Y^i$ ) is:

$$Y^i = x^i + pm^i.$$

There are two sources of income. First, a stochastic and taxable income  $(1 + \varepsilon)$  from the export sector with individual specific average  $(1 + \varepsilon)$ . Second, a non-taxable income from a specific factor that can be employed in either sector with individual-specific probabilities that are ex-ante uncertain. The income is 1 if employed in the export sector, and  $p$  if employed in the import competing sector. The probability of employment in the import sector is  $\sigma(1 - \tilde{\eta})$ , where  $\sigma \in (0, 1)$  captures exposure to foreign competition, while  $\tilde{\eta}$  is a random variable with individual specific mean  $\eta$ . An individual with higher  $\eta$  is more employable in the export sector, perhaps because he is more mobile or skilled. Higher  $\eta$  thus corresponds to higher benefits from globalization.<sup>21</sup>

There are two policies. First, the income tax  $\tau$  with quadratic distortions as before. Second, an ad valorem tariff  $t$ , so that the domestic price of imports is  $p = (1 + t)p^*$ , where  $p^*$  is the world price. Revenue from both instruments is used to finance  $g$ .

A voter's type is summarized by the  $(\varepsilon, \eta)$  vector, and his expected income is:

$$Y^{\varepsilon\eta}(\tau, t) = (1 + \varepsilon)(1 - \tau) + \sigma(1 - \eta) [p^*(1 + t) - 1] + 1$$

The first term is taxable income net of taxes. The remaining terms denote expected income from the specific input. The income tax  $\tau$  redistributes between individuals with high vs low expected income  $\varepsilon$ , while the tariff redistributes between individuals with high vs low benefits from trade openness,  $\eta$ . If  $p < 1$ , increased exposure to import competition (a higher  $\sigma$ ) causes the income of this voter to drop, the more so the lower is  $\eta$ . In this precise sense,

---

<sup>21</sup>To rely on normality, which yields the exact distortion in beliefs of Equation (9), we assume that the variance is sufficiently low that  $\sigma(1 - \tilde{\eta})$  is between zero and one with very large probability. We could also specify a bounded distribution for  $\tilde{\eta}$ , such as a beta, and obtain very similar results.

globalization, is an income shock correlated with a potential carrier of identity: exposure to foreign competition.

The appendix solves for the equilibrium. Because individuals differ only in their expected income and are risk neutral, their indirect utility function is separable in  $\varepsilon$  and  $\eta$ . The tax rate  $\tau^\varepsilon$  rationally preferred by type  $\varepsilon$  is still given by (4), the rationally optimal tariff  $t^\eta$  decreases in export sector employability  $\eta$ .

In this model, identification can either be class based, with the usual groups  $L \equiv \{\varepsilon | \varepsilon < \widehat{\varepsilon}\}$  and  $U \equiv \{\varepsilon | \varepsilon > \widehat{\varepsilon}\}$ , or trade-based with a group of nationalists  $N \equiv \{\eta | \eta < \widehat{\eta}\}$  and a group of cosmopolitans  $C \equiv \{\eta | \eta > \widehat{\eta}\}$ . The distance function  $\Delta^{\varepsilon\eta}(G)$  that determines identification is still approximated by equation (18), except that  $\eta$  replaces  $\psi$  and the relative weight  $\alpha$  is replaced by a new parameter  $\gamma$  capturing the importance of trade policy relative to income taxes. A higher value of  $\gamma$  makes identification on the trade dimension more likely. As shown in the appendix, this happens (among other things) if: 1) individuals are more exposed to import competition ( $\sigma$  increases); 2) income tax distortions  $\varphi$  are higher.

Before adding cultural conflict, we can see that this two dimensional model with trade is isomorphic to the model of Section 4. As a result, all results obtained there apply here, except that the interpretation of the shocks, beliefs, and policies is different. If individuals become more exposed to import competition ( $\sigma$  rises), the prevailing identities switch to nationalists vs cosmopolitans. Rich and poor opponents of globalization unite against rich and poor cosmopolitans. That is, because the income drop is correlated with the group of losers from trade, class identity wanes. As a result, the demand for income redistribution drops and, if nationalists outnumber cosmopolitans, the demand for protectionism rises. Equilibrium policy becomes less redistributive and more protectionist.

### 5.3 Three Dimensional Conflict

We add cultural conflict by simply adding the loss function  $-\frac{\kappa}{2} (q - \tilde{\psi})^2$  to the utility function in (25). Now there are three policy instruments: income tax  $\tau$ , import tariff  $t$ , and civil rights  $q$ . A voter's type is pinned down by the vector  $(\varepsilon, \eta, \psi)$  reporting his expected income, exposure to trade, and culture. Identification can occur along any of these dimensions. The distance function is now a direct extension of (18):

$$\Delta^{\varepsilon\psi\eta}(G) \simeq |\varepsilon - \bar{\varepsilon}_G| + \alpha |\psi - \bar{\psi}_G| + \gamma |\eta - \bar{\eta}_G| - \lambda (|\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}}| + \alpha |\bar{\psi}_G - \bar{\psi}_{\bar{G}}| + \gamma |\bar{\eta}_G - \bar{\eta}_{\bar{G}}|). \quad (26)$$

Distances along culture  $\psi$  are weighed by  $\alpha = \kappa\nu\varphi$  as before. Distances along the openness trait  $\eta$  are weighted by parameter  $\gamma$ . Globalization (higher  $\sigma$ ) increases the welfare relevance of trade  $\gamma$ , favoring trade-based identification.

As in section 5.1, we assume that  $\lambda \rightarrow \infty$ , so that all voters identify along either income,



openness, or culture. We also assume that there is an identical majority of poor, nationalists and conservatives,  $\hat{\varepsilon} = \hat{\eta} = \hat{\psi} = z > 0$ . This assumption is relaxed in the appendix.

The joint distribution of  $(\varepsilon, \psi, \eta)$  is Gaussian around the zero mean  $(0, 0, 0)$  with unitary variances and with the following correlation structure: income ( $\varepsilon$ ) is uncorrelated with the other traits, while social progressiveness ( $\psi$ ) and openness to trade ( $\eta$ ) are positively associated, with correlation coefficient  $\rho > 0$ . The results are qualitatively similar if all traits are correlated (as we show in the appendix). What is key is that openness to trade and social progressiveness are more positively correlated with each other than with income. This is consistent with survey data, as shown in Appendix Table A1, which reports the correlation matrix for attitudes towards government spending, trade protection, immigration and abortion in the 2000 ANES survey, and in a similar Appendix Table A2 on France. One body of work attributes this correlation to a psychological predisposition to "closedness" (Johnston et al. 2017), which is reflected in aversion to both international and cultural openness. Another body of work relates it to a belief in "communal values" (Graham et al. 2009, Enke 2019), which places a special emphasis in defending local communities and their customs.

### 5.3.1 Identity and Beliefs

For the purpose of determining beliefs and policies, the key question is whether voters choose to identify with their income, trade, or cultural group. The answer depends, once again, on the importance of civil rights relative to taxation,  $\alpha$ , and on the importance of trade protection relative to taxation,  $\gamma$ . The Appendix proves the following result.

**Proposition 7** *If the importance of trade and culture relative to taxes is low,  $\max(\gamma\rho + \alpha, \gamma + \alpha\rho) < 1$ , then identification is class-based. If trade is important,  $\gamma > \max(\alpha, 1 - \alpha\rho)$ , then identification is trade-based. Finally, if culture is important,  $\alpha > \max(\gamma, 1 - \gamma\rho)$ , then identification is culture-based.*

When redistribution is important, society is divided into upper and lower classes. Economic or social change can upset this equilibrium. In particular, globalization hits less mobile or less educated losers from trade ( $\eta < 0$ ). By doing so, it renders tariff policy more welfare relevant for them, namely  $\gamma$  increases, favoring their nationalistic identification.

Critically, the change in prevalent identities now exerts pervasive effects. If voters identify with their economic class, belief distortions only concern the redistributive domain. The reason is that income is uncorrelated with the other traits and hence economic classes only differ in expected income. If instead identities switch to nationalism vs cosmopolitanism, then belief distortions concern both trade *and* civil rights, since traits are correlated in these two dimensions. As a result, society does not just witness a change in the dimension along which preferences are most polarized. Overall polarization increases in a correlated way across issues.

To see this, suppose that a lower class voter abandons class identification and acquires a nationalistic identity. Going through the same steps as in Subsection 5.1, this voter becomes: i) less concerned about social mobility (since  $\varepsilon_N^\theta = \varepsilon > \varepsilon_L^\theta$ ); ii) more concerned about exposure to trade (since  $\eta_N^\theta < \eta_L^\theta = \eta$ ); and iii) more socially conservative (since  $\psi_N^\theta < \psi_L^\theta = \psi$ ). As a result, he demands less redistribution, more trade protection but also more conservative policies towards civil rights or immigrants. A similar outcome arises if identification switches to culture, rather than to nationalism. Thus, positive correlation between economic nationalism and social conservatism implies that trade shocks (or cultural changes) exacerbate conflict over several dimensions at once.

The analysis of all cases yields the following result.

**Corollary 5** *Suppose that an increase in  $\alpha$  or  $\gamma$  or  $\rho$  changes identification from class to trade, or from class to culture. Then, conflict over tax policy is dampened, while conflict over both trade and civil rights policy is enhanced.*

As shown in the Appendix, in the overall political equilibrium the end result of these identity switches is that trade policy becomes distorted, and so does civil rights policy.

## 5.4 Discussion

In many countries, trade and technology shocks have contributed to stagnant wages and increasing inequality. One may have expected losers from these shocks to swell the ranks of leftwing parties, but this has not occurred. Our model offers an explanation: these shocks have reduced the cohesiveness of the lower class. Socially conservative poor voters, who traditionally identified with the left despite their social conservatism, are now attracted by nationalism or by culturally conservative platform, because they appeal to both their trade preferences and their cultural views, and viceversa for voters with opposite political features. Greater income inequality between nationalists and cosmopolitans, or between socially conservatives and socially progressives, renders this effect stronger by making these alternative cleavages even more efficient carriers of conflict. By weakening the cohesion of income groups, these shocks favor a common nationalistic and socially conservative response by economic losers concentrated in less educated strata and in declining regions. As a result, there is a reduction in the demand of classical income redistribution, a support for protectionist policies, and strong aversion to immigrants and more generally to culturally progressive policies.

This mechanism can also explain why in the US partisan views have become more correlated across policy dimensions (e.g., Gentzkow 2017). The growing cleavage between Democrats and Republicans over basic cultural traits such as "universalism" vs. "communalism" is an instance of this phenomenon (Enke 2018). When class identities were dominant, these

cultural traits shaped individual political preferences, but in a latent way. As conflict over the cultural dimension becomes more important, it triggers the adoption of new cultural identities. These, in turn, increase polarization across many issues because, by its very nature, culture has broad implications. The growing divide over globalization is a complementary mechanism. Conflict between winners and losers from international trade shapes identities based on geography or sector of employment. These dimensions are correlated with local values, and the new identities influence beliefs across several domains.

In the next Section we look at some of these predictions in more detail.

## 6 Some Evidence

This section explores the empirical validity of some of the theoretical predictions outlined above, using survey data for the US and France. These two countries have different political systems. The US has two parties, and each of them represents more than one social or political group. In France there are several political parties, that can be matched with groups defined on specific issues, or on different dimensions of political conflict (like left vs right, as well as nationalism vs globalism). These differences allow us to explore the data in different ways. Our aim is not to estimate precise causal effects, but to show that the main predictions of the theory are consistent with observed correlations in the data.

### 6.1 Evidence on the US

Here we consider two implications of the theory: those associated with a rise in cultural conflict (the parameter  $\alpha$ ), and with a rise in trade conflict (the parameter  $\gamma$ ). Cultural conflict has risen over time, but we cannot exploit any exogenous source of variation. We thus discuss whether our theory can explain observed time patterns in survey data. In the case of trade conflict, instead, there are exogenous changes in imports exposure, associated with China joining the WTO. Following Autor (2017), we thus compare changes in the attitudes of individuals more or less exposed to those changes.

#### 6.1.1 Cultural conflict

As discussed in Subsection 4.5, there is much anecdotal evidence that the relevance of cultural conflict related to race, immigration and civil rights has risen over time. This is reflected in survey data. Figure 5 plots the percentage of respondents in a repeated Pew survey that mention different issues as one of the three most important problems (other than employment, the state of the economy and national security issues) facing the US, between 2001 and 2018. Since 2013, race and immigration are mentioned with the highest frequency. 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.<sup>22</sup> Similar trends are observed in several European countries.

Figure 5

### 6.1.2 Changing patterns of polarization

Propositions 3, 6 and 7 imply that, as the relevance of cultural conflict increases (as  $\alpha$  rises), more voters identify based on culture, abandoning income based identification. Then, Proposition 4 and Corollary 3 and 5 imply that, when identities switch in this way, new stereotypes are formed that cause social groups to become more polarized on cultural issues, and less polarized on redistributive policies.

We do not observe social identities, but much anecdotal evidence is consistent with a switch of prevailing identities in the direction suggested by the model. Frank (2004) beautifully describes the "cultural backlash" in Kansas, a state that was firmly Democrat in the past, where socially conservative and low income voters have become obsessed with civil rights, leading them to see politics through issues such as immigration or race, and away from redistribution. Similarly, Goodhart (2017) and Fukuyama (2018) emphasize how political conflict in the UK and the US has shifted from traditional left vs right economic conflict to a new cleavage concerning cultural issues.

Has this resulted in a greater polarization in cultural issues and a reduction of polarization in redistributive issues? Survey data are broadly consistent with this idea. Figure 6 depicts the proportion of ANES respondents with opposite extreme views on immigration and on the size of government between 1998 and 2016. Attitudes on government spending are measured by views on whether the government should provide more or fewer services; attitudes towards immigrants are measured by views on the desired level of immigrants. Both variables have been rescaled to lie between 0 and 1, and higher values denote more liberal attitudes. To remove the effect of changes in fundamentals, we have estimated the residuals after conditioning on income and education. Individuals with extreme views are defined as those whose estimated residuals fall in the top and bottom 5% of the pooled distribution (i.e. pooling all waves between 1998 and 2016). Results are very similar if the extreme ranges are defined as the top and bottom 10% of the pooled distribution. The frequency of extremists on immigration rises over time and almost doubles between 1998 and 2016. This happens because both extreme opponents and supporters of immigration have risen. Extremism on government spending, instead, falls until 2012, with extreme conservatives rising and extreme liberals declining, and

---

<sup>22</sup>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).

in 2016 it returns towards its original level.

Figure 6

Of course, Figure 1 on the changing stance of working class vs. upper/middle class members on redistribution and immigration is also consistent with our model. When identity switches from economic to cultural, demand for redistribution by poorer strata drops relative to the upper class and disagreement over immigration intensifies.<sup>23</sup>

The average pattern in Figure 1 could reflect changes in the composition of social classes in terms of education, age or other individual features. To control for possibly confounding factors, we have estimated multiple OLS regressions, one per wave, where the dependent variables are those depicted in Figure 1. Throughout we control for age, age-squared and dummy variables for gender and education. Figure 7 plots the estimated coefficients of dummy variables for whether the respondent belongs to the working class (on the left hand side of each panel) or upper-middle and upper class (on the right hand side), together with 95% confidence intervals; thus, the omitted categories are the lower-middle and middle classes. Working class individuals tend to favor higher government spending and oppose immigration, while the opposite is true for upper-middle and upper class individuals, relative to the middle classes. But the important fact is that, since 2008, social class has become less correlated with preferences on the size of government, and more strongly correlated with attitudes towards immigration. Note that the switching dimension of political conflict reflects changes in both economic classes: working class individuals now demand less redistribution and more control of immigration, while upper middle / upper class individuals have become less opposed to redistribution and more in favor of immigration. This symmetry is consistent with our theory, while it is harder to reconcile with the idea that reduced demand for redistribution is simply due to an exceptional inflow of immigrants accompanied by a consequent reluctance to redistribute in favor of strangers (cf. Alesina et al. 2019)

Figure 7

This phenomenon is not confined to the US. Evans and Mellon (2016) document how, in the British Social Attitudes survey of 2015, UK working class respondents hold balanced views on questions about inequality and redistribution, but they are predominantly conservative on issues like immigration, the death penalty, homosexuality and morality.

---

<sup>23</sup>We can interpret a self defined social class in the ANES survey in two ways. First, as an income group in our model. In this case, the theory predicts the pattern in Figure 1 if income and social progressiveness are positively correlated. Alternatively, we can interpret a self defined social class in the survey as a bundle of correlated individual traits such as income, occupation and education. In this case, self defined working class individuals are mostly located in the lower-left quadrant of Figure 2 (i.e. they are predominantly poor and socially conservative), while upper middle class are mostly in the upper-left quadrant; as identity switches from income to culture, again the theory predicts the pattern in Figure 1.

Theories based on rational voters have a hard time explaining these changing patterns of divergence. Increased importance of immigration and civil rights may change how people vote, but there is no reason why it should affect how policy preferences are related to social classes, or why extremism should rise on immigration more than on other policy issues. Belief distortions associated with endogenous identities can instead explain these patterns.

### 6.1.3 Exposure to import competition

One of the main results of section 5 is that trade shocks can induce political reactions in several correlated policy dimensions. As those hurt by the trade shock switch to nationalistic identities, they demand more trade protection and less income based redistribution, but they also turn anti-immigrants and socially conservative (see Proposition 7 and Corollary 5). We now ask whether the effects of increased exposure to imports from China are consistent with this prediction. Unlike in the previous subsections, we go beyond mere correlations. We implement the IV strategy of Autor et al. (2017), but rather than focusing on political polarization and an aggregate measures of conservative ideology, as they do, we study the effect on particular attitudes on which our theory has specific predictions.<sup>24</sup>

We exploit the CCES survey, a large scale survey conducted in the US over the internet between 2006 and 2016. The advantage over ANES data is that here about 36 000 individuals were interviewed every year on average. For a subset of 8300 respondents there is also a panel dimension over the period 2010-14. Import exposure varies at the commuter zone (CZ) level, and we have about 60 respondents per CZ and year on average (15 in the panel analysis). We focus on four indicators: whether the respondent prefers to cut domestic spending or raise taxes in order to reduce the budget deficit, aversion to immigrants (measured by the first principal component of 2 variables on immigrants), how important is the issue of abortion (not available in the panel), and how the respondent voted in the closest election.<sup>25</sup>

We estimate two sets of regressions. First, a two-period repeated cross section, where individual attitudes are sampled in the first (2006-8) and last (2016) year of our sample - the initial year varies by question. Second, a panel regression in first difference, between the first (2010) and last (2014) years of the panel. Our treatment variable is the change in US imports from China, that we instrument with the change in European imports from China, as in Autor

---

<sup>24</sup>The main outcome variables in Autor et al. (2017) are different voting variables (in the elections and in the US Congress). They also look at indicators of conservative ideology from survey data, obtained by aggregating 10 questions in the Pew (2014) survey; the Pew survey lacks the panel dimension and has about 25 respondents per CZ.

<sup>25</sup>The question on abortion only asks how important is this issue, and not whether the respondent is in favor or against. But year-by-year correlations with a related question on the acceptability of abortion suggests that those who regard abortion as a more important issue are also more opposed to it. Unfortunately the formulation of this second question on abortion changed over time, making it impossible to use it in cross years comparisons. Votes refer to presidential elections in the cross section and to state Senate in the panel.

et al. (2017). US imports from China grew particularly fast before the start of our sample period. In the repeated cross section we thus take the change in imports exposure between the year 2000 and the last year of our sample period (2016). This amounts to assuming that the full effect of increased imports on identity is not instantaneous, but entails some lag. Since the panel starts four years later than the cross section, to preserve symmetry with the cross sectional regressions, in the panel we take the change in imports between 2004 and 2014.

The specification and the estimation method are as in Autor et al. (2017), and are described in the Appendix. The specification allows us to study how individual attitudes have changed over time within CZ as a result of increased imports exposure, both in the repeated cross section and in the panel. Table 1 reports the estimates for two specifications, with and without covariates for the CZ. The first two columns refer to repeated cross sections, the others to panel data. The variable of interest is the change in exposure in the CZ (as in Autor et al. 2017, in the repeated cross section this variable is interacted with a dummy that equals 1 in the last period, so as to estimate the effect on attitudes in the last period).

Table 1

Residents of CZ more exposed to an acceleration of imports from China became more willing to accept cuts in domestic public spending, as predicted by the theory if relatively poor respondents abandoned class identification. They also became more averse to immigration and more attentive to the issue of abortion, as predicted given that protectionist attitudes are positively correlated with sentiments against immigrants and abortion (cf. Table A1 in the Appendix). Finally, they became more likely to vote Republican (no effect is found on votes for Democrats). These results are in line with those of Autor et al. (2017), but they clarify which attitudinal dimensions account for the shift towards a more conservative ideology.<sup>26</sup>

An implication of our theory is that the effect of identity switches on beliefs is stronger if stereotyping is more pronounced (if  $\theta$  is larger). We do not observe this behavioral parameter, but social media may have enhanced the strength of stereotypes, because they make it easier to convey caricatures and emotional content. CCES data also contain information on how frequently respondents use blogs and listen to online news (all respondents use internet, since the survey was carried out online). In Table 2 we estimate a similar set of regressions on the 2010-2014 longitudinal sample, but we interact the change in imports exposure with two

---

<sup>26</sup>According to our panel estimates, an acceleration in exposure by one standard deviation increases the change in willingness to cut spending by 2%, and increases the change in aversion to immigrants by 8%, both relative to average attitudes in the first year. The change in the probability of voting Republican is 3.6 percentage points higher, if the change in exposure rises by one standard deviation. The magnitude of the estimated effects in the repeated cross sections is smaller: a one standard deviation increase in imports exposure changes attitudes by about 1.5%-2.5% of the sample mean in the first period, and increases the probability of voting Republican by about 2 percentage points.

dummy variables for whether the respondent uses blogs and online news respectively. The dependent variable is the change in a dummy variable for supporting a tougher immigration policy and cuts to domestic public spending (the latter is defined as having a willingness to cut domestic public spending above a given threshold; in columns 1 and 2 the threshold corresponds to the 2010 median of the same variable, in columns 3 and 4 to the 67th percentile). Respondents living in countries more exposed to import competition are more likely to become anti-immigrant and opponents of redistribution, confirming the results of Table 1. But the effect of trade shocks is 66% larger for blog users, while it is smaller for those who follow online news. The amplification effect of blog use and the dampening effect of online news are just a correlation and can have more than one interpretation, but they are suggestive that exaggeration due to stereotyping is one of the mechanisms behind the political effects of trade shocks.

Table 2

## 6.2 Evidence on France

France is an ideal testing ground for our theory, because there was a clear shift in the dimensions of political conflicts between 2013 and 2017. This emerges clearly using data from the Dynamob survey, a representative panel of French citizens between 2013 and 2017.

We consider 9 questions on three sets of policy issues: (i) the traditional left vs right economic cleavage, namely attitudes towards redistribution and protection of workers' rights; (ii) civil rights (role of women and rights of homosexual); (iii) the cosmopolitan vs nationalist cleavage, namely attitudes towards immigrants, globalization and European integration - the appendix provides more detail. On each set of issues we have extracted the first polychoric principal component, estimated the residuals by conditioning on the respondent's income and education, and the resulting variables have been normalized across the two years so that they all have the same variance. These variables thus describe individual attitudes on issues (i)-(iii), after removing the component correlated with income and education.

Figure 8 illustrates the scatter plots of two of these variables. The vertical axis measures attitudes towards globalization and immigration (higher values correspond to more open attitudes), the horizontal axis attitudes on redistribution (higher values correspond to more right wing attitudes). Each dot corresponds to an individual. The tones indicate how respondents were split between two clusters estimated from all three variables: in 2013 on the left hand panel, in 2017 on the right hand panel (more individuals were interviewed in 2017 than in 2013). The change in the dimension of political conflict is striking. In 2013 respondents were split between left and right, in the traditional dimension of economic conflict over redistribution and the role of the state in the domestic economy. In 2017, the cleavage concerned



globalization and immigration.

Figure 8

The criterion for assigning observations to a cluster is to minimize within group variance, in the space of the three variables described above.<sup>27</sup> This clustering exercise thus tells us that, between 2013 and 2017, individual views on issues (i)-(iii) became more distant between opponents and supporters of international openness, and more similar in the traditional left vs right dimension of political conflict. According to our theory, this should be associated with a switch in the dimension of identification, from left / right to nationalism / cosmopolitanism. Identities are not observed, but they are likely correlated with how people voted, which we observe. Indeed, the two clusters track closely how the respondent voted in the first and second rounds of presidential elections. Figure 9 reports the vote shares within each cluster. The two vertical clusters of 2013 largely correspond to how votes were split between left and right wing candidates in the 2012 Presidential election. The two horizontal clusters of 2017 instead correspond to how votes were split in 2017 between Le Pen and Macron, two candidates that are hard to pin down on the left vs right divide, but that differ instead in their positions over globalization and immigration.<sup>28</sup>

Figure 9

The theory discussed in Section 5 also predicts that identity switches lead to a change in policy preferences. Someone who identified with the right in 2012, and switched to globalism in 2017, should become more pro-globalization and less anti-redistribution. To the extent that progressive views on civil rights are more correlated with support for globalization than with attitudes towards redistribution, he should also become more socially progressive. Likewise, someone who abandoned his left wing identity and became a nationalist should become more anti-globalization, more socially conservative and less pro-redistribution.<sup>29</sup>

These predictions can be tested using the panel dimension of the Dynamob survey, assuming that identities are correlated with how people voted in the first round of the presidential elections (a subset of about 470 individuals were interviewed both in 2013 and 2017, and were asked how they voted in the previous presidential election). Specifically, we assume that those who voted left or right in the first round identified with the left ( $L$ ) and right ( $R$ ) wing group respectively. Similarly, those who cast their first round vote for Le Pen or Macron identified with the nationalist ( $N$ ) and cosmopolitan ( $C$ ) group respectively. Voters who in the first

---

<sup>27</sup>Clusters are defined by applying Ward's minimum variance method on the normalized residuals of the three first principal components described above. The appendix provides more details.

<sup>28</sup>Note that the clusters depicted in Figure 8 were identified without exploiting the voting information. The vote shares in Figure 9 do not sum to 100% because some survey respondents abstained or voted for minor candidates not classifiable on the left / right or nationalist / globalist dimensions.

<sup>29</sup>Table A2 in the appendix shows that, as assumed in Section 5 also the attitudes of French respondents towards civil rights are more correlated with attitudes toward globalization than toward redistribution

round abstained or voted for parties not clearly positioned on the left vs right axis, or globalist vs nationalist dimension, are taken to be not politically identified on these dimensions.<sup>30</sup>

We estimate the following regression, on different samples:

$$\Delta y_i = \alpha + \beta_C C_i + \beta_N N_i + \gamma y_{i0} + X_i' \delta + FEs + u_i \quad (27)$$

The first sample includes all respondents for which we have data both in 2013 and 2017. Here, the dependent variable,  $\Delta y_i$ , is the change in attitudes towards (i) globalization and immigration and (ii) civil rights, between 2013 and 2017. Attitudes are measured by the first principal components described above, rescaled to lie between 0 and 1 (higher values denoting a more open or socially progressive policy). The variables of interest are the dummy variables  $C$  and  $N$  for whether in 2017 individual  $i$  voted cosmopolitan and nationalist respectively. The variable  $y_{i0}$  denotes his initial attitudes in 2013,  $X$  is a vector of individual covariates measured in the initial period (education, gender, age, immigrant status, region and rural area) or differenced (income and employment status),  $u$  is the error term and  $FEs$  are dummy variables that have to be included to estimate the effect of acquiring a cosmopolitan or nationalist identity, specifically a dummy variable for voting  $N$  both in 2012 and 2017, and for voting  $N$  in 2012 (nobody switched from  $N$  to  $C$ ). Thus, the coefficients  $\beta_C$  and  $\beta_N$  refer to those who voted  $C$  and  $N$  *for the first time* in 2017 respectively (i.e., under our assumption who acquired a new cosmopolitan or nationalist identity). Hence, we expect  $\beta_{Gl} > 0 > \beta_N$ . All votes refer to the first round. The appendix provides full details on all the variables.

Table 3 reports the results, with and without the individual covariates  $X$ . As expected, those who voted Le Pen for the first time in 2017 became more opposed to international openness and more socially conservative, while those who turned to Macron in the first round of 2017 became more in favor of globalization (but the estimated coefficient is only significant in the more parsimonious specification) and did not change their views on civil rights. The point estimates imply that voting for Le Pen for the first time in 2017 is associated with a drop of attitudes in favor of openness by about 25% of average initial attitudes amongst non-Le Pen voters in 2012, relative to the non-Le Pen voters in 2012 who did not vote for either Le Pen or Macron in 2017. The negative effect of voting le Pen on support for civil rights is about -11% of average initial attitudes, computed in the same way.

Table 3

---

<sup>30</sup>Both in 2012 and 2017, there was more than one candidate on the left and on the right at the first round elections (the main candidates on the left and right were Hollande and Sarkozy respectively, but there were other minor candidates). Le Pen was a candidate in both 2012 and 2017, whereas Macron only in 2017. Thus, implicitly we assume that nobody identified as a globalist in 2012, a likely exaggeration. See the appendix for how we classify presidential candidates on the left and right dimensions.

In the second set of estimates, the dependent variable is the change in attitudes over redistribution, measured by the first principal component described above and rescaled to lie between 0 and 1, with higher values indicating a more right wing policy. Here the theory has predictions only for those who initially voted either left or right and then switched dimension of identification to the cosmopolitan vs nationalist axis (views on redistribution should not be affected by the acquisition of a nationalist or cosmopolitan identity, for those who were not previously identified on the left / right dimension). Thus, here we impose two sample restrictions: first we only consider those who in 2012 voted Left and in 2017 voted either Left, or Right or Macron or Le Pen (we call this the 2012 Left sample), and then only those who in 2012 voted Right and in 2017 again voted either Left, or Right or Macron or Le Pen (the 2012 Right sample). In both samples we then estimate the same specification as in (27) above, except that  $FEs$  only includes dummy variable for those who in 2017 voted right (in the 2012 Left sample) or who in 2017 voted Left (in the 2012 Right sample). Thus, the omitted category is those who did not change their vote between 2012, and kept voting Left (in 2012 Left sample) or Right (in the 2012 Right sample). The coefficients of interest have the following expected signs:  $\beta_C, \beta_N > 0$  in the 2012 Left sample, and  $\beta_C, \beta_N < 0$  in the 2012 Right sample (recall that a higher value of the dependent variable means that respondent became more opposed to redistribution).

Table 4 reports the estimates, again with and without the covariates. All coefficients of interest have the expected sign, and they are statistically significant for those who switched from Left to Macron, and for those who switched from Right to Le Pen. The point estimates imply that switching from Left to Macron increases opposition to redistribution by about 36% of the initial average attitudes of those who voted Left in 2012, relative to those who voted Left in 2012 and did not switch. Such change is half as large as the one experienced by those who switched from Left to Right, in line with expectations. The effect of switching from Right to Le Pen, evaluated in a similar manner, implies that opposition to redistribution falls by 27%. Note that here the estimated coefficient on those switching from Right to Left is not statistically significant, but only five individuals fall in this category.

Table 4

All in all, therefore, the evidence from vote switchers in the first round of French presidential elections supports several predictions of theory. Of course, these estimated coefficients only capture correlations in the data, and cannot be interpreted as causal effects of identity changes. Individuals may have changed how they voted because they rationally changed opinions, or both voting and attitudes could reflect relevant omitted variables.

## 7 Concluding Remarks

Identity theory provides a rich and promising framework. We have shown how it can explain systematic distortions in political beliefs, actual and perceived polarization, causes and consequences of changing political cleavages, the effects of trade or technology shocks.

The general idea is that political conflict builds on a set of latent social groupings, characterized along economic and cultural traits, and representing demands that are more or less correlated across different issues. A well known grouping, of course, is income or wealth based. But ascriptive groups have also played a role historically, for instance emphasizing culture, geography, or race. As political cleavages change, voter switch their identification from their income class, to their cultural, geographical, or racial group. Crucially, while the switch may be driven by the political issue of the day, it influences beliefs across the board, because different social groups cut society in clusters of interests along many domains.

Here we explored some key implications of this approach, but much more remains to be done. An important issue concerns the role of the media, which are often held responsible for incorrect information and extremism. New digital media, such as Twitter or Facebook, may enhance stereotypical thinking because they focus attention on simple and forceful messages at the expenses of more nuanced policy debates. Disintermediation of the traditional media may favor leaders that reach out to voters with emotional and symbolic messages that appeal to their identities. At the same time, identities also form in spontaneous interactions with neighbors or friends sharing similar problems. Exploring the mechanisms for the diffusion of identities is an important topic for future empirical research.

In our model opportunistic politicians simply adapt to voters' demands. But identity politics also matters on the supply side. One natural question is about persuasion. Which dimensions of conflict ought to be emphasized by vote maximizing politicians? More importantly, how can voters be induced to identify with political leaders, and which features of politicians make them more appealing to voters? 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. Sometimes political leaders may even 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.

We think that our demand side approach may be useful to understand the supply side of the process. For instance, which messages by leaders or parties will be particularly persuasive and when? When are existing parties able to exploit new social cleavages, and when will they be dismantled? Social psychology suggests that stereotypical types exert more influence. This implies that, when polarization is strong, the most successful politician is likely to come from the tails, not the middle. This may help explain why successful populist politicians often look

similar to the unskilled and unexperienced labor market outsiders that voted them in office (see in particular Dal Bo et al. (2018) on Sweden).

Yet another important set of questions concern political identification and the evolution of party systems. Under what circumstances does a political party represents a single identity group, and when instead it acts as an ensemble of heterogeneous social identities? In the mixed identification regime of Proposition 2, it seems that different parties may specialize on "pockets" of voters identified, and hence polarized, along different issues. 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 get the opposite groups in each dimension of identification. Why is this so, and could it imply that, under some assumptions, policies are extreme in both dimensions?

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

## 8 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 and Edward Glaeser (2005), *Fighting Poverty in the US and Europe: A World of Difference*, Oxford UP, Oxford.

Alesina, Alberto, Elie Murard and Hillel Rapoport (2019) "Immigration and Preferences for Redistribution in Europe" Harvard, mimeo

Alesina, Alberto, Stefanie Stantcheva, and Armando Miano (2018), "Immigration and Redistribution", *mimeo*.

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso (2018), "Intergenerational Mobility and Preferences for Redistribution", *mimeo*.

Anelli, Massimo, Italo Colantone, and Piero Stanig (2018), "We Were The Robots: Automation and Voting Behavior in Western Europe", *mimeo*.

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

Autor, David, David Dorn and Gordon Hanson (2013), "The China Syndrome: Local Labor Market Effects of Import Competition in the United States", *American Economic Review*, 103(6), 2121-68.

Autor, David, David Dorn, Gordon Hanson and Kaveh Majlesi (2017), "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure", *MIT Working Paper*.

Bénabou, Roland, and Jean Tirole (2006), Belief in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, 121(2) (2006), 699-746

Bénabou, Roland, and Jean Tirole (2011), "Identity, Morals, and Taboos: Beliefs as Assets", *The Quarterly Journal of Economics*, Volume 126, Issue 2, 805–855.

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 (2018), "Beliefs about Gender", *American Economic Review*, forthcoming.

Bordalo, Pedro, Marco Tabellini and David Yang (2019) "Stereotypes and Politics", mimeo, Harvard

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes (1960), *The American Voter*, The University of Chicago Press.

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

Colantone, Italo, and Piero Stanig (2018a), "The Economic Roots of the Cultural Backlash: Global Competition and Attitudes in Western Europe", *mimeo*

Colantone, Italo, and Piero Stanig (2018b), "Global Competition and Brexit", *American Political Science Review*, 112(2).

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 (2018), "Economic Losers and Political Winners: Sweden's Radical Right", *mimeo*.

Diermeier, Daniel, and Christopher Li (2018), "Partisan Affect and Elite Polarization", *American Political Science Review*, 1-5.

- Enke, Ben (2018), "Moral Values and Voting", *mimeo*.
- Enke, Ben (2019), "Kinship, Cooperation, and the Evolution of Moral System", *Quarterly Journal of Economics*, 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 (2005), *"What's the Matter with Kansas? - How Conservatives Won the Heart of America"*, Picador, New York
- Fryer, Roland, and Matthew Jackson (2008), "A Categorical Model of Cognition and Biased Decision-Making", *B.E. Journal of Theoretical Economics*, 8 (2008), 1–42
- 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", *mimeo*.
- Gidron, Norman, and Peter A. Hall (2017), "The politics of social status: economic and cultural roots of the populist right", *British Journal of Sociology*, 68 (S1), S57.S84.
- Gimpelson, Vladimir, and Daniel Treisman (2018), "Misperceiving inequality", *Economics & Politics*, Vol. 30 (1), 27-54.
- Glaeser, Edward and Bryce Ward (2006) "Myths and Realities of American Political Geography", *Journal of Economic Perspectives*, 20(2), 119–144
- 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"*, Penguin Books
- Graham, Jesse, Brian Nosek, and Jonathan Haidt (2009), "Liberals and Conservatives Rely on Different Sets of Moral Foundations," *Journal of Personality and Social Psychology*, 96 (5), 1029.
- Guiso, Luigi, Helios Herrera, Massimo Morelli, and Tommaso Sonno (2017), "Demand and Supply of Populism", *EIEF Working Papers Series 1703*, Einaudi Institute for Economics and Finance (EIEF).

Grossman, Gene M., and Elhanan Helpman (2018), "Identity Politics and Trade Policy", *mimeo*.

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 Abrans, 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.

Howe, Irving (1977), "Socialism and America", Harcourt Brace.

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. (2014), "Climate-Science Communication and the Measurement Problem", *Advances in Pol. Psych.*, 36, 1-43.

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.

Layman, Geoffrey C., Thomas M. Carsey, and Juliana M. Horowitz (2006), "Party Polarization in American Politics: Characteristics, Causes, and Consequences", *Annual Review of Political Science*, 9: 83-110.

Lipset, Seymour Martin and Stein Rokkan (1967) Cleavage structures, party systems, and voter alignments: an introduction, in Lipset S M & Rokkan S, eds. *Party systems and voter alignments: cross-national perspectives*. Free Press, New York. p. 1-64

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

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 Schleifer (2004), "Persuasion in Politics", *American Economic Review Papers and Proceedings*, 94(2), 435-439.



Oakes, Penelope J. (1987), "The Salience 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*.

Pierce, Justin R., and Peter Schott (2012), "A concordance between ten-digit US Harmonized System codes and SIC/NAICS product classes and industries", *Journal of Economic and Social Measurement*, 37, 1-2, 61-96.

Rosch, Eleanor (1978), "Principles of Categorization", In E. Rosch, and B.B. Lloyd (Eds.), *Cognition and Categorization*, Hillsdale, NJ: Erlbaum, 27-48.

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.

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

Tabellini, Marco (2018), "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration", Harvard Business School, mimeo

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.

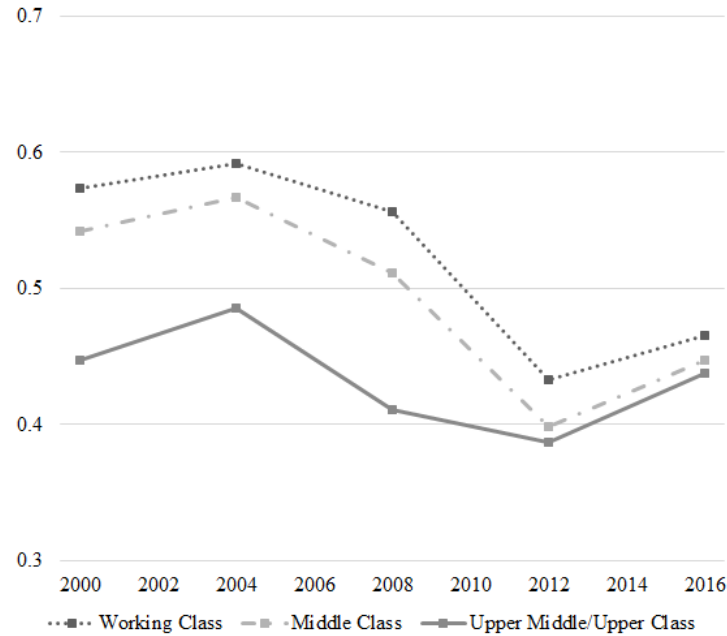
Tajfel, Henri, M. Billig, R. Bundy, and Claude Flament (1971), "Social Categorization and Intergroup Behavior", *European Journal of Social Psychology*, 1, 149-178.

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.

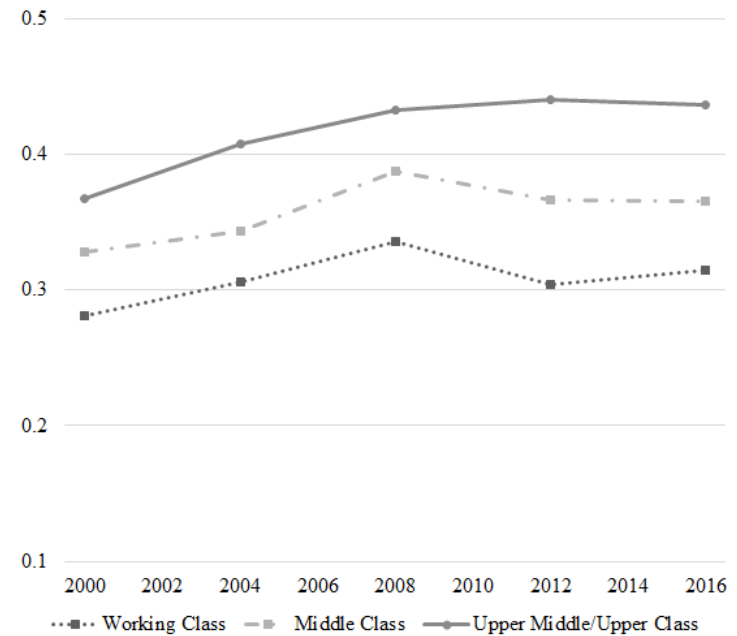
Tversky, Amos (1977), "Features of Similarity", *Psychological Review*, 84, 327-352.

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

**Figure 1: Social Class and Policy Preferences**



**(a) Desired Size of Government**



**(b) Desired Immigration Levels**

Notes: Desired size of government is measured as the answer to the following question: “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. Desired immigration levels are measured using the following question: “Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?”. The resulting five-point variable is recoded so that higher values reflect more liberal views on immigration. Both variables are rescaled between 0 and 1. The sample consists of all white individuals aged 18 or more. Source: ANES Time Series Studies.

Figure 2: Primitive Groups Defined on Income and Culture

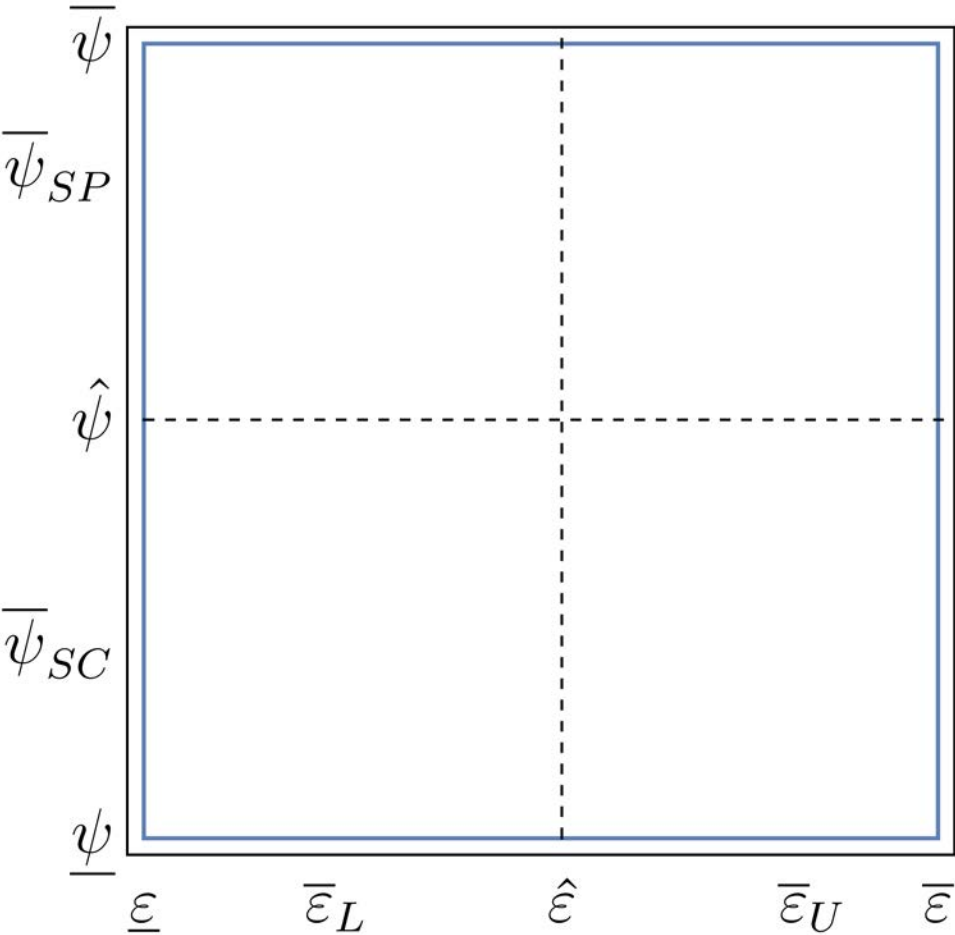
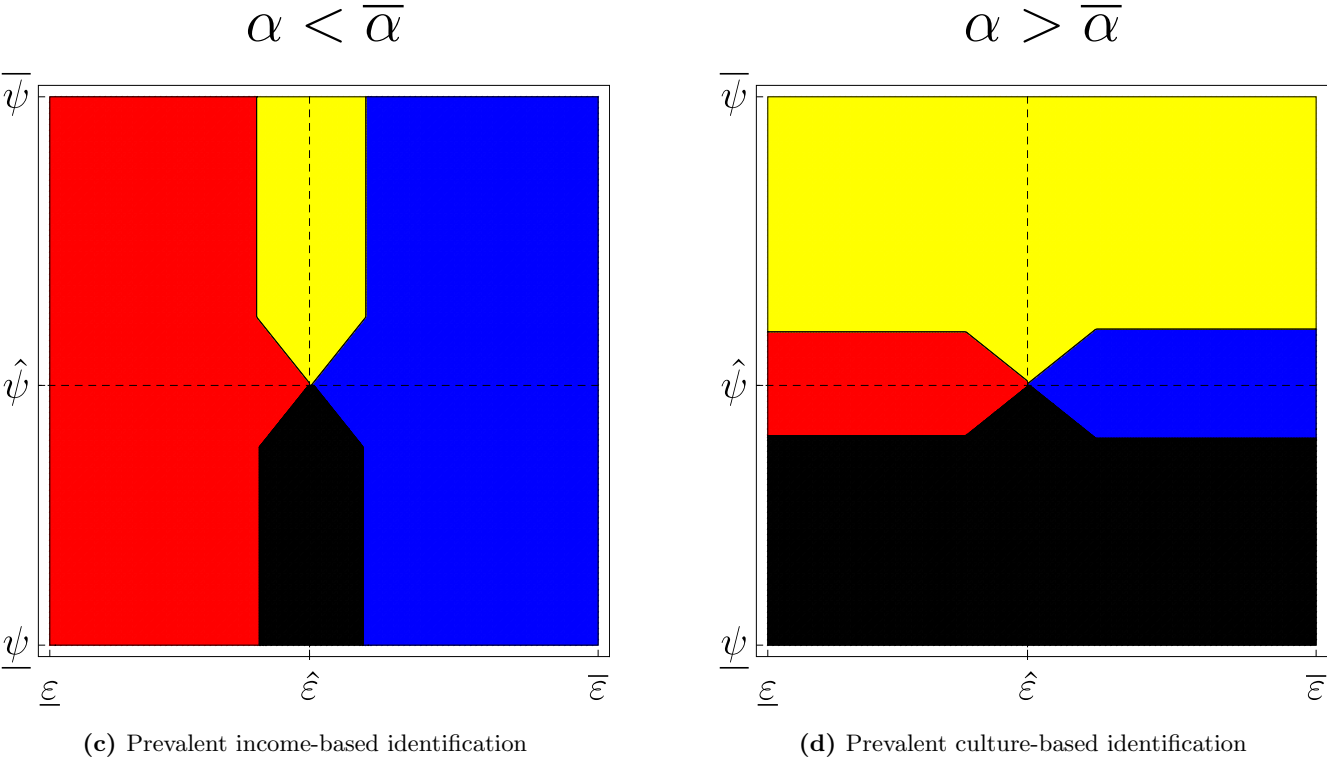
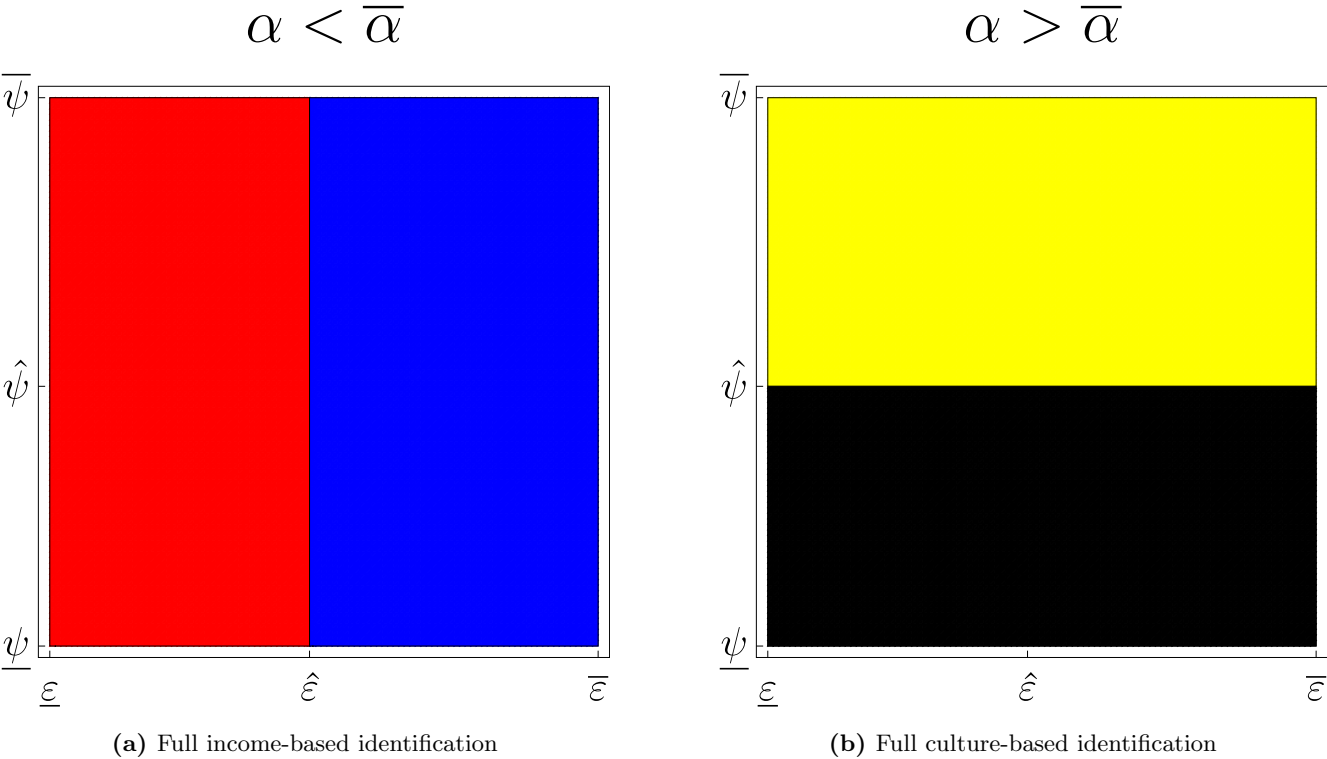


Figure 3: Identification



■  $U$  ■  $L$  ■  $SP$  ■  $SC$

Figure 4: Primitive Income Groups

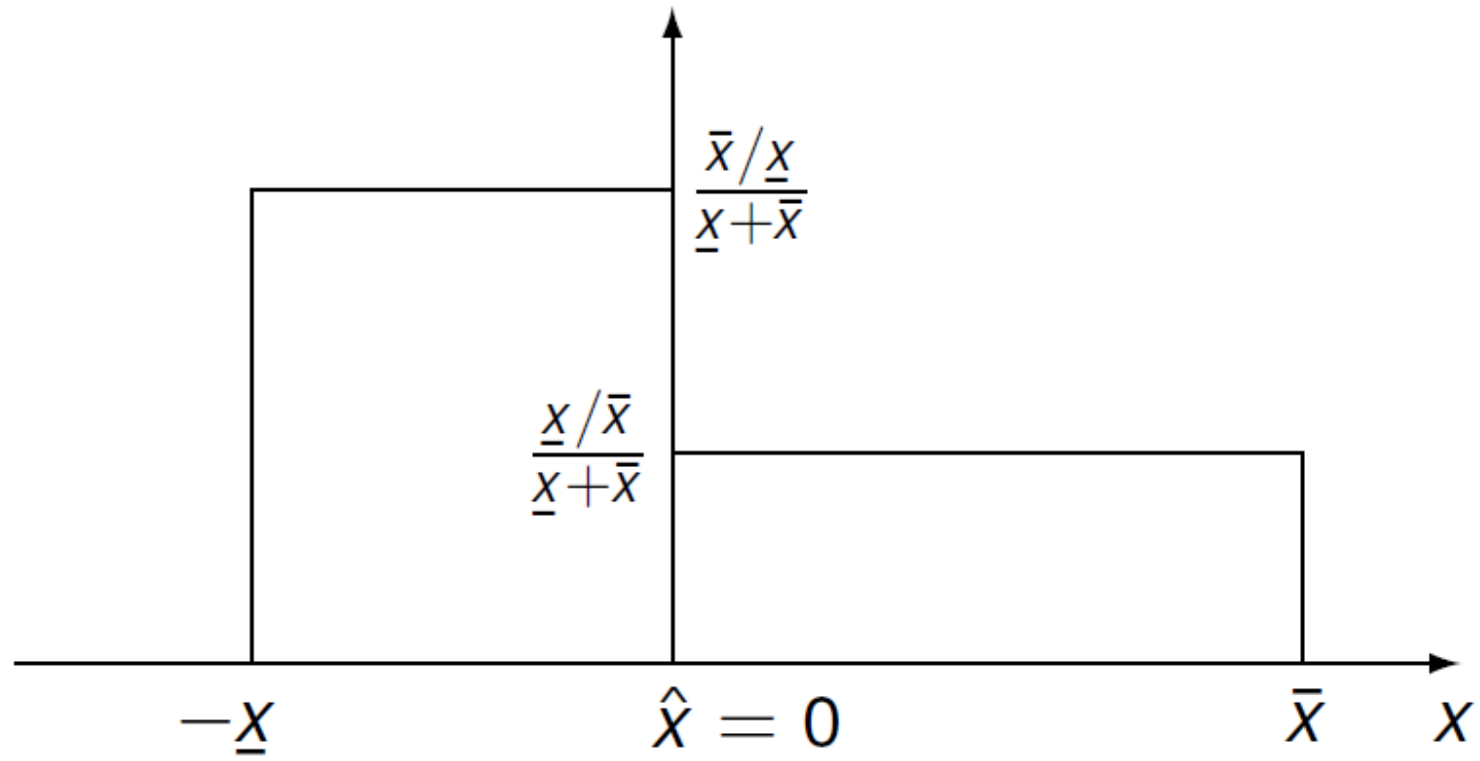
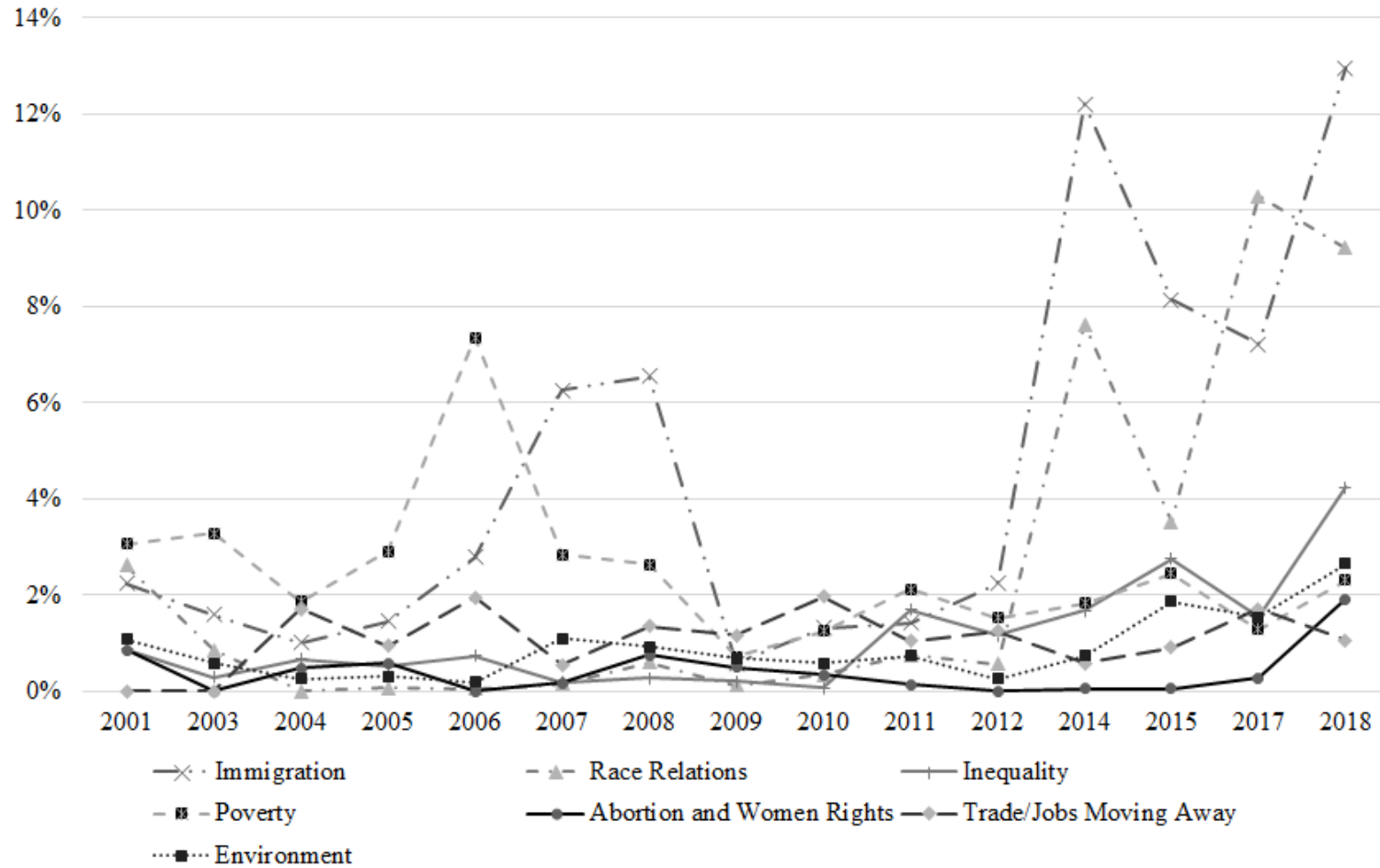
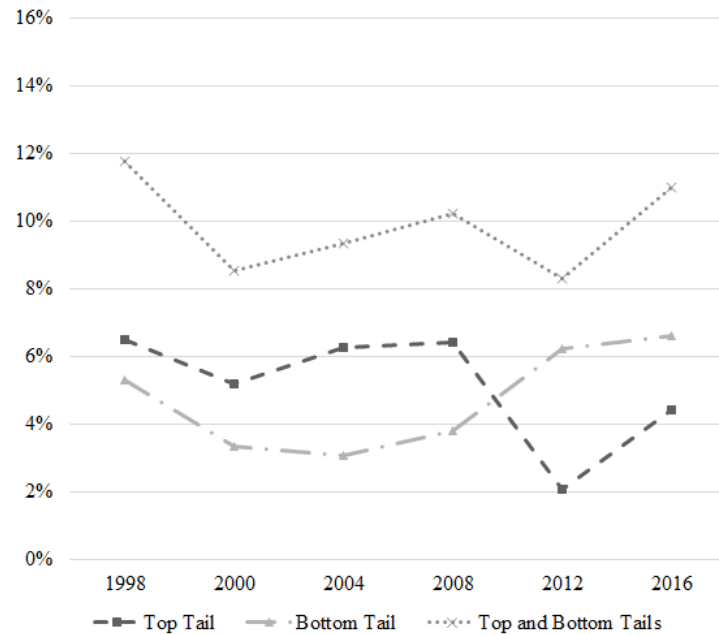


Figure 5: Most Important Problem Facing The Country

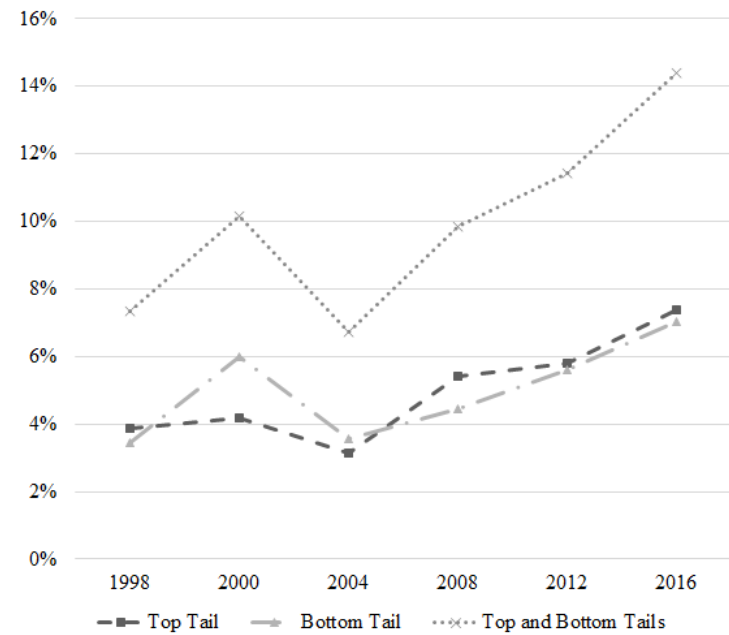


Notes: The graph shows the share of respondents mentioning the selected issues among the top three most important problems facing the US. Source: Pew Research Center public data.

**Figure 6: Extremism in the Population**



**(a) Desired Size of Government**



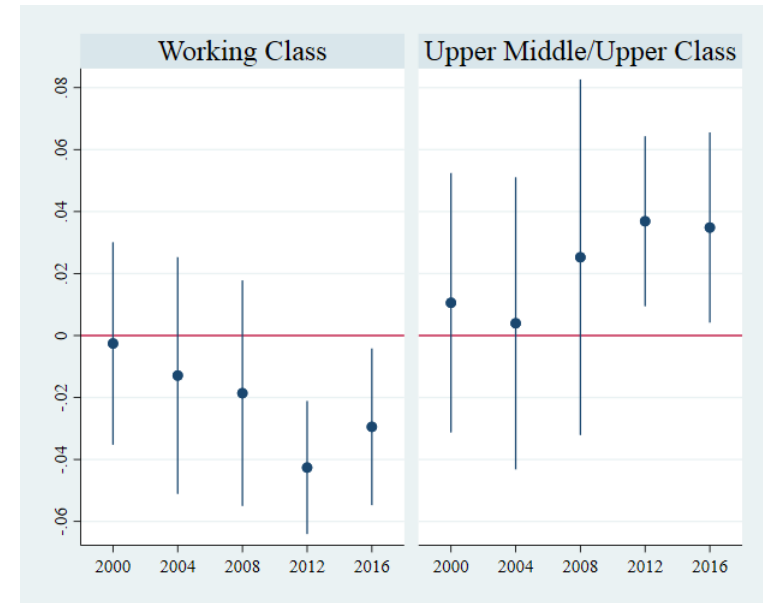
**(b) Desired Immigration Levels**

Notes: Desired size of government is measured as the answer to the following question: “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. Desired immigration levels are measured using the following question: “Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?”. The resulting five-point variable is recoded so that higher values reflect more liberal views on immigration. Both variables are rescaled between 0 and 1. The figure is constructed as follows. We have estimated the residuals of these two variables, by conditioning on income and education. For each residual we have computed the ranges corresponding to the top and bottom 5% in the pooled distribution that combines all waves between 1998 and 2016. Individuals holding extreme views are those whose estimated residuals are included in these extreme ranges. Both pictures look very similar if extremism is defined on a range corresponding to top and bottom 10% on the pooled distribution. The sample consists of all individuals aged 18 or more. Source: ANES Time Series Studies.

**Figure 7: Social Class and Policy Preferences**



**(a) Desired Size of Government**

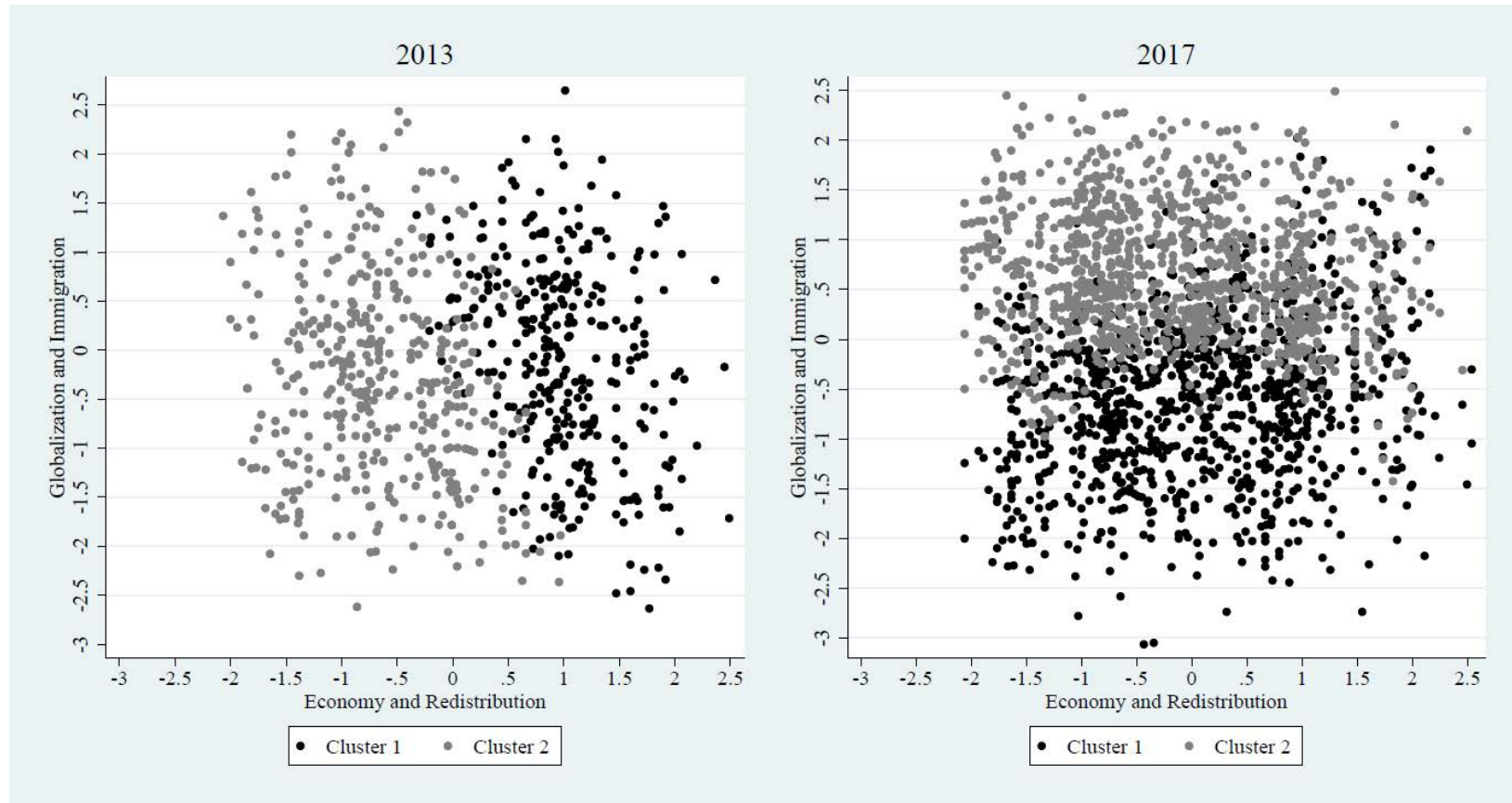


**(b) Desired Immigration Levels**

Notes: Desired size of government is measured as the answer to the following question: "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' willingness to redistribute. Desired immigration levels are measured using the following question: "Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?". The resulting five-point variable is recoded so that higher values reflect more liberal views on immigration. Both variables are rescaled between 0 and 1. Coefficients are obtained from OLS regressions on dummy variables for whether respondents identify as working class or upper middle/upper class. The omitted category are those identifying as lower middle/middle class. Controls include age, age-squared and dummies for gender and education. 5% confidence intervals displayed. The analysis on immigration is restricted to white individuals. Source: ANES Time Series Studies.

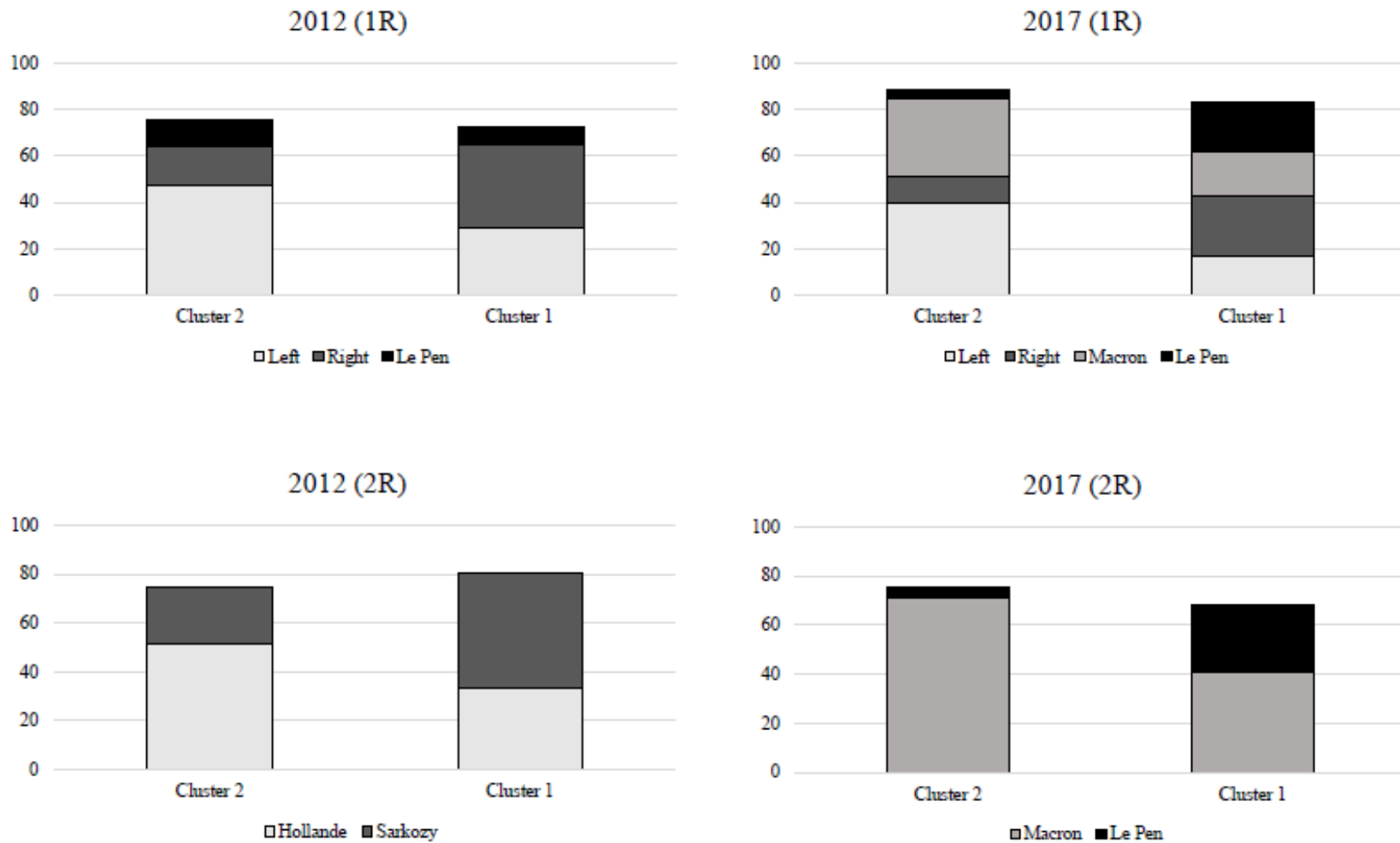


**Figure 8: Changing Dimension of Political Conflict:  
France in 2013 and 2017, Cluster Analysis**



Notes: The vertical axes measure attitudes towards immigration, globalization and European integration (higher values correspond to more open attitudes), the horizontal axes attitudes on the role of the government in regulating the economy and on redistribution (higher values correspond to more right-wing attitudes), for two samples drawn from the French adult population in 2013 and 2017. These measures were constructed by first extracting the first polychoric principal component from two sets of questions, one set for each of these two dimensions of political conflict, and then estimating the residuals after conditioning on education and income. Each marker corresponds to an individual. The colors indicate how respondents were split between two clusters, individuated applying Ward's method on the above-mentioned residuals and on a third variable, obtained extracting the first principal component of a set of questions on civil rights (on gender roles and homosexuals), and then conditioning on income and education. The clustering algorithm is run separately for each year, after standardizing the input variables. See the appendix for the list of questions, the data sources and more information on how the variables were treated before the analysis. Source: Dynamob

Figure 9: Composition of Votes in the Two Clusters



Notes: The four panels show the vote shares received by the different candidates in the French presidential elections of 2012 and 2017, within each of the clusters illustrated in Figure 8. The stacked vote shares do not add up to 100% because some of the respondents in each clusters abstained or voted for candidates that could not be classified on the left / right or nationalist / globalist dimension. Source: Dynamob

**Table 1: Exposure to Imports from China and Attitudes**

	Cross Section		Panel	
	A.	Cut Domestic Spending (0-100)		
$\Delta$ CZ Exposure * Second Period	0.512 (0.363)	1.086** (0.525)		
$\Delta$ CZ Exposure			0.703 (0.938)	3.503** (1.771)
Observations	72,712	72,712	8,296	8,296
F on Excluded Instruments	63.5	25.7	80.31	42.7
	B.	Migrant Aversion (PC)		
$\Delta$ CZ Exposure * Second Period	0.010** (0.004)	0.014 (0.010)		
$\Delta$ CZ Exposure			0.076*** (0.023)	0.120** (0.054)
Observations	73,484	73,484	9,451	9,451
F on Excluded Instruments	75.17	31.01	65.11	42.1
	C.	Abortion Importance		
$\Delta$ CZ Exposure * Second Period	0.013*** (0.005)	0.019** (0.009)		
Observations	48,871	48,871		
F on Excluded Instruments	72.99	27.29		
	D.	Republican Vote		
$\Delta$ CZ Exposure * Second Period	0.0146* (0.008)	0.0197* (0.012)		
$\Delta$ CZ Exposure			0.054*** (0.0181)	0.097*** (0.0278)
Observations	77,558	77,558	6,673	6,673
F on Excluded Instruments	61.14	24.53	69.73	46.39
CZ Controls	NO	YES	NO	YES

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . For each commuter zone (CZ), the change in exposure refers to the period between 2000 and 2016 in the cross section and between 2004 and 2014 in the panel. In the cross section, the dependent variables are measured in the following pairs of years: 2006 and 2016 (Panel A and C); 2007 and 2016 (Panel B); 2008 and 2016 (Panel D). In the panel, all dependent variables are first differenced and refer to 2010-2014. Republican vote refers to presidential elections in the cross section and state senate elections in the panel. All specifications include demographic controls for gender, a quadratic of age, educational attainment and race (in the cross section they are interacted with the second period dummy, while in the panel they refer to the first period); the cross section also includes the CZ mean of the dependent variable in the initial period interacted with a dummy variable for the second period, while in the panel we include the first period level of the dependent variable; finally, the panel also includes a dummy variable for those who changed CZ between 2010 and 2014, alone and interacted with the change in imports exposure. Fixed effects for CZ and for the second period (referred to as “Second Period” in the table) are included in the cross sections. CZ controls refer to year 2000 and include the manufacturing share in CZ employment, the offshorability and routine task indexes of Autor and Dorn (2013), and the county-level republican vote share interacted with a dummy for Republican victory in that county. Standard errors are clustered at CZ level. Estimation is by 2SLS. Source: CCES.

**Table 2: Interactions with Blog Use**

	Against Migrants and Above Median Cuts		Against Migrants and Top Third Cuts	
$\Delta$ CZ Exposure	0.0394*	0.118***	0.0333	0.0999***
	(0.0237)	(0.0366)	(0.0224)	(0.0357)
$\Delta$ CZ Exposure * Blog	0.0886***	0.0775**	0.0673*	0.0572*
	(0.0318)	(0.0313)	(0.0346)	(0.0342)
$\Delta$ CZ Exposure * Online News	-0.0841***	-0.0819***	-0.0514*	-0.0498*
	(0.0313)	(0.0311)	(0.0277)	(0.0276)
CZ Controls	NO	YES	NO	YES
Observations	8,262	8,262	8,262	8,262
F on Excluded Instruments	43.14	21.94	43.20	21.97

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Each specification includes controls for blog and online newspaper use. Refer to Table 1 for a detailed description of the sample and other controls. The dependent variable is a dummy equal to one if the respondent supports a tougher immigration policy and (i) reports a value above the first-year median of the “tax vs cuts” 100-point scale (first two columns); or (ii) reports a value higher than two thirds of the answers of the first year (last two columns). In all specifications, the dependent variables are differenced. Standard errors clustered are clustered at CZ level. Estimation is by 2SLS. Source: CCES.

**Table 3: Switching Identities, Openness and Culture**

	$\Delta$ Globalization and Immigration		$\Delta$ Cultural Progressiveness	
C 17	0.0318** (0.0147)	0.0161 (0.0153)	-0.0111 (0.0213)	-0.00919 (0.0210)
N 17	-0.127*** (0.0237)	-0.121*** (0.0260)	-0.0506 (0.0322)	-0.0678** (0.0343)
N12	-0.0402 (0.0327)	-0.0779* (0.0398)	-0.0661 (0.0456)	0.0489 (0.0537)
N12 * N17	0.0164 (0.0456)	0.0457 (0.0539)	0.0663 (0.0650)	-0.0401 (0.0736)
Individual Controls	NO	YES	NO	YES
Observations	469	398	472	401
R-squared	0.294	0.333	0.309	0.368

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Standard errors in parentheses. The dependent variable is the change in attitudes between 2013 and 2017, in a panel of respondents. All specifications include the level of the dependent variable in the first year and the intercept. Individual controls are education, income and dummy variables for gender, age, immigrant status, employment status, NUTS 1 region and rural area. Income and employment are differenced. Estimation is by OLS. Source: Dynamob.

**Table 4: Switching Identities and Economy**

		Switching from Left		Switching from Right	
C 17		0.105*** (0.0304)	0.137*** (0.0367)	-0.0112 (0.0419)	-0.0432 (0.0467)
N 17		0.00207 (0.0675)	0.0730 (0.0749)	-0.249*** (0.0555)	-0.176*** (0.0659)
R 17		0.206*** (0.0722)	0.260*** (0.0753)		
L 17				-0.0743 (0.0933)	-0.128 (0.0960)
Sample					
	2012	L		R	
	2017	L, R, C, N		L, R, C, N	
Individual Controls		NO	YES	NO	YES
Observations		159	146	120	107
R-squared		0.418	0.519	0.335	0.521

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Standard errors in parentheses. The dependent variable is the change in attitudes towards the economy and redistribution between 2013 and 2017, in a panel of respondents. All specifications include the level of the dependent variable in the first year and the intercept. Individual controls are education, income and dummy variables for gender, age, immigrant status, employment status, NUTS 1 region and rural area. Income and employment are differenced. Estimation is by OLS. Source: Dynamob.

**Table A1: Correlation Between Policy Issues  
US**

	Redistribution	Globalization	Immigration	Abortion
Redistribution	1			
Globalization	-0.0692 0.1550	1		
Immigration	-0.0035 0.9192	0.2154 0.0000	1	
Abortion	0.0900 0.0092	0.1608 0.0000	0.0157 0.5179	1

Notes: Data refer to year 2000. The labels “Redistribution”, “Globalization”, “Immigration” and “Abortion” correspond respectively to the variables “Desired Size of Government”, “Trade Openness”, “Desired Immigration Levels” and “Abortion Policy” described in the appendix. Higher values denote more liberal and open attitudes on all issues (more left-wing when it comes to redistribution). Source: ANES Time Series Studies.

**Table A2: Correlation Between Policy Issues  
France**

	Redistribution	Globalization	Culture
Redistribution	1		
Globalization	-0.0383 0.0347	1	
Culture	0.1019 0.0000	0.2597 0.0000	1

Notes: Correlations are computed pooling together the samples of 2013 and 2017. The labels “Redistribution”, “Globalization” and “Culture” correspond respectively to the polychoric principal components “Economy and Redistribution”, “Immigration and Globalization” and “Social Progressiveness” described in the appendix. The sign of the component “Economy and Redistribution” has been reversed so that higher values denote more left-wing attitudes. Source: Dynamob.

## 9 Appendix

### 9.0.1 Proofs

**Proof of Proposition 1.** By adapting Equation (2), we model the stereotyped belief that type  $\varepsilon$  identified with group  $G$  has about himself as:

$$f^\theta(\tilde{\varepsilon}|\varepsilon, G) = f(\tilde{\varepsilon}|\varepsilon) \left[ \frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})} \right]^\theta Z. \quad (28)$$

As a result, the distorted perception of social mobility is equal to:

$$\begin{aligned} \varepsilon_G^\theta &= \int \tilde{\varepsilon} f^\theta(\tilde{\varepsilon}|\varepsilon, G) d\tilde{\varepsilon} = \int \tilde{\varepsilon} \left[ \frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})} \right]^\theta f(\tilde{\varepsilon}|\varepsilon) Z d\tilde{\varepsilon}, \\ &= \varepsilon + Z \cdot \text{cov} \left( \tilde{\varepsilon}, \left[ \frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})} \right]^\theta \right), \end{aligned} \quad (29)$$

where the second line follows because by definition  $\mathbb{E} \left\{ \left[ \frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})} \right]^\theta Z | \varepsilon \right\} = 1$  (i.e.  $Z$  is set so that  $f^\theta(\tilde{\varepsilon}|\varepsilon, G)$  integrates to one). Hold  $\varepsilon$  fixed. When  $\frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})}$  is increasing in  $\tilde{\varepsilon}$ , we have that  $\varepsilon_G^\theta > \varepsilon$ . When instead  $\frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})}$  is decreasing in  $\tilde{\varepsilon}$ , we have that  $\varepsilon_G^\theta < \varepsilon$ . We now show that the former case holds for  $G = U$ , while the latter for  $G = L$ . The cdf of income  $\tilde{\varepsilon}$  in group  $G$  satisfies:

$$F(\tilde{\varepsilon}|G) \propto \int_{\varepsilon \in G} F(\tilde{\varepsilon}|\varepsilon) h_G(\varepsilon) d\varepsilon,$$

Where  $h_G(\varepsilon)$  is the distribution of types in group  $G$ . It is fine to define the cdf and pdf up to a proportionality term because this is absorbed by the normalizing constant  $Z$ . By differentiating we find that the pdf is proportional to:

$$f(\tilde{\varepsilon}|G) \propto \int_{\varepsilon \in G} f(\tilde{\varepsilon}|\varepsilon) h_G(\varepsilon) d\varepsilon.$$

The likelihood ratio is thus equal to:

$$\frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})} \propto \frac{\int_{\varepsilon \in G} f(\tilde{\varepsilon}|\varepsilon) h_G(\varepsilon) d\varepsilon}{\int_{\varepsilon \in \overline{G}} f(\tilde{\varepsilon}|\varepsilon) h_{\overline{G}}(\varepsilon) d\varepsilon} = \frac{\int_{\varepsilon \in G} \frac{f(\tilde{\varepsilon}|\varepsilon)}{f(\tilde{\varepsilon}|\hat{\varepsilon})} h_G(\varepsilon) d\varepsilon}{\int_{\varepsilon \in \overline{G}} \frac{f(\tilde{\varepsilon}|\varepsilon)}{f(\tilde{\varepsilon}|\hat{\varepsilon})} h_{\overline{G}}(\varepsilon) d\varepsilon}.$$

The rightmost way of writing the likelihood ratio is key. By assumption,  $f(\tilde{\varepsilon}|\varepsilon)$  satisfies the MLRP. As a result,  $\frac{f(\tilde{\varepsilon}|\varepsilon)}{f(\tilde{\varepsilon}|\hat{\varepsilon})}$  increases in  $\tilde{\varepsilon}$  for  $\varepsilon > \hat{\varepsilon}$  and decreases in  $\tilde{\varepsilon}$  for  $\varepsilon < \hat{\varepsilon}$ . This



implies that  $\frac{f(\tilde{\varepsilon}|U)}{f(\tilde{\varepsilon}|L)}$  is increasing while  $\frac{f(\tilde{\varepsilon}|L)}{f(\tilde{\varepsilon}|U)}$  is decreasing. As a result,  $\varepsilon_U^\theta > \varepsilon$  while  $\varepsilon_L^\theta < \varepsilon$ .

To prove property i), note that by differentiating the first row of (29), and by denoting  $\frac{f(\tilde{\varepsilon}|G)}{f(\tilde{\varepsilon}|\overline{G})} = L_G(\tilde{\varepsilon})$  we find:

$$\begin{aligned} \frac{\partial \varepsilon_G^\theta}{\partial \theta} &\propto \int \tilde{\varepsilon} \left\{ \ln L_G(\tilde{\varepsilon}) - \frac{\mathbb{E} \left[ \ln L_G(\tilde{\varepsilon}) \cdot L_G(\tilde{\varepsilon})^\theta | \varepsilon \right]}{\mathbb{E} \left[ L_G(\tilde{\varepsilon})^\theta | \varepsilon \right]} \right\} \frac{L_G(\tilde{\varepsilon})^\theta}{\mathbb{E} \left[ L_G(\tilde{\varepsilon})^\theta | \varepsilon \right]} f(\tilde{\varepsilon} | \varepsilon) d\tilde{\varepsilon} \\ &= \text{cov} \left( \tilde{\varepsilon}, \left\{ \ln L_G(\tilde{\varepsilon}) - \frac{\mathbb{E} \left[ \ln L_G(\tilde{\varepsilon}) \cdot L_G(\tilde{\varepsilon})^\theta | \varepsilon \right]}{\mathbb{E} \left[ L_G(\tilde{\varepsilon})^\theta | \varepsilon \right]} \right\} \frac{L_G(\tilde{\varepsilon})^\theta}{\mathbb{E} \left[ L_G(\tilde{\varepsilon})^\theta | \varepsilon \right]} \right). \end{aligned}$$

It is immediate to see that the covariance is positive if  $L_G(\tilde{\varepsilon})$  is increasing and negative if  $L_G(\tilde{\varepsilon})$  is decreasing, which proves that  $\frac{\partial \varepsilon_U^\theta}{\partial \theta} > 0 > \frac{\partial \varepsilon_L^\theta}{\partial \theta}$ .

To prove property ii), denote by  $f'(\tilde{\varepsilon}|G)$  and by  $f'(\tilde{\varepsilon}|\overline{G})$  the new, more polarized, income densities of all individuals except for the (measure zero) atom whose density continues to be  $f(\tilde{\varepsilon}|\varepsilon^*)$ , where  $\varepsilon^*$  is the voter's true average income. Note that by definition the new likelihood ratio can be written as:

$$\begin{aligned} \frac{f'(\tilde{\varepsilon}|G)}{f'(\tilde{\varepsilon}|\overline{G})} &\propto \frac{\int_{\varepsilon \in G} k_G(\tilde{\varepsilon}) f(\tilde{\varepsilon}|\varepsilon) h_G(\varepsilon) d\varepsilon}{\int_{\varepsilon \in \overline{G}} k_{\overline{G}}(\tilde{\varepsilon}) f(\tilde{\varepsilon}|\varepsilon) h_{\overline{G}}(\varepsilon) d\varepsilon} = \\ &= M_G(\tilde{\varepsilon}) L_G(\tilde{\varepsilon}), \end{aligned}$$

where  $M_G(\tilde{\varepsilon}) = k_G(\tilde{\varepsilon})/k_{\overline{G}}(\tilde{\varepsilon})$  is increasing if and only if  $L_G(\tilde{\varepsilon})$  is increasing. By replacing the new likelihood ratio and the new density in Equation (29), one obtains that beliefs of  $\varepsilon^*$  are more optimistic under the more polarized distributions if and only if:

$$\begin{aligned} \varepsilon_G^\theta &= \int \tilde{\varepsilon} f'^\theta(\tilde{\varepsilon} | \varepsilon^*, G) d\tilde{\varepsilon} = \int \tilde{\varepsilon} \frac{M_G(\tilde{\varepsilon})^\theta L_G(\tilde{\varepsilon})^\theta}{\mathbb{E} \left[ M_G(\tilde{\varepsilon})^\theta L_G(\tilde{\varepsilon})^\theta \right]} f(\tilde{\varepsilon} | \varepsilon^*) d\tilde{\varepsilon} > \\ &\int \tilde{\varepsilon} \frac{L_G(\tilde{\varepsilon})^\theta}{\mathbb{E} \left[ L_G(\tilde{\varepsilon})^\theta \right]} f(\tilde{\varepsilon} | \varepsilon^*) d\tilde{\varepsilon} \end{aligned}$$

which is equivalent to:

$$\int \tilde{\varepsilon} L_G(\tilde{\varepsilon})^\theta \left\{ M_G(\tilde{\varepsilon})^\theta - \frac{\mathbb{E} \left[ M_G(\tilde{\varepsilon})^\theta \cdot L_G(\tilde{\varepsilon})^\theta | \varepsilon \right]}{\mathbb{E} \left[ L_G(\tilde{\varepsilon})^\theta | \varepsilon \right]} \right\} f(\tilde{\varepsilon} | \varepsilon^*) d\tilde{\varepsilon} > 0,$$

which is true if and only if  $L_G(\tilde{\varepsilon})$  (and thus  $M_G(\tilde{\varepsilon})$ ) is increasing in  $\tilde{\varepsilon}$ . As a result, for  $G = U$  individuals become more optimistic under the more polarized distribution, while for

$G = L$  they become more pessimistic. ■

**Proof of Corollary 1.** This is obtained by plugging normal densities into Equation (8), which implies that  $f^\theta(\tilde{\varepsilon}|\varepsilon, G)$  is normal with variance  $\sigma_{\tilde{\varepsilon}}^2$  and mean  $\varepsilon + \theta(\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})$ . ■

**Proof of Corollary 2.** Denote the distributions of policy preferences in the two groups as  $f(\tau|Left)$  and  $f(\tau|Right)$ . Here we mechanically apply Equation (2). In light of the proof of Proposition 1, to show that perceptions of disagreement about taxes are exaggerated we only need to show that  $f(\tau|Left)/f(\tau|Right)$  satisfies MLRP, where  $f(\tau|G)$  is the true distribution of policy preferences (distorted by stereotypes) in group  $G$ . Denote by  $\alpha_G$  the share of  $L$ -identified voters in party  $G = Left, Right$ . We then have the cdf:

$$F(\tau(\varepsilon)|G) = \begin{cases} (1 - \alpha_G) H(\varepsilon) & \text{for } \tau(\varepsilon) \leq \tau(\hat{\varepsilon}) \\ \alpha_G[H(\varepsilon) - \pi_U] + (1 - \alpha_G) \pi_U & \text{for } \tau(\varepsilon) > \tau(\hat{\varepsilon}) \end{cases},$$

where the mapping  $\tau(\varepsilon)$  takes into account also stereotypes about own social mobility. The pdf is equal to:

$$f(\tau(\varepsilon)|G) = \begin{cases} (1 - \alpha_G) h(\varepsilon) & \text{for } \tau(\varepsilon) \leq \tau(\hat{\varepsilon}) \\ \alpha_G h(\varepsilon) & \text{for } \tau(\varepsilon) > \tau(\hat{\varepsilon}) \end{cases}.$$

The likelihood ratio is then equal to:

$$\frac{f(\tau(\varepsilon)|Left)}{f(\tau(\varepsilon)|Right)} = \begin{cases} \frac{1 - \alpha_{Left}}{1 - \alpha_{Right}} & \text{for } \tau(\varepsilon) \leq \tau(\hat{\varepsilon}) \\ \frac{\alpha_{Left}}{\alpha_{Right}} & \text{for } \tau(\varepsilon) > \tau(\hat{\varepsilon}) \end{cases},$$

which satisfies MLRP because  $\alpha_{Left} > \alpha_{Right}$ . As a result, in analogy with the proof of Proposition 1, the perceived average position of *Left* is more redistributionist than in reality and the perceived average position of *Right* is more laissez faire [?] than in reality, creating exaggerated perceived polarization. ■

**Proof of Proposition 2.** By inspection of Equation (13). ■

**Approximate relative distance function.** Here we derive the piecewise linear approximation of  $\Delta_d^{\varepsilon\psi}$  in (18) from equation (17). Define the bliss point of types  $\psi$  by  $q^\psi = \psi \equiv Q(\psi)$ . Recall that  $\tau^G = T(\bar{\varepsilon}_G) = \frac{v - (1 + \bar{\varepsilon}_G)}{\varphi v}$  and  $q^G = \bar{\psi}^G \equiv Q(\bar{\psi}^G)$ , so that  $W^{\varepsilon\psi}(\tau^G, q^G) = W^{\varepsilon\psi}(T(\bar{\varepsilon}_G), Q(\bar{\psi}_G))$ . Taking a second order approximation of  $W^{\varepsilon\psi}(\tau^G, q^G)$  with respect to  $\bar{\varepsilon}_G$  and  $\bar{\psi}_G$  at the point  $\bar{\varepsilon}_G = \varepsilon$  and  $\bar{\psi}_G = \psi$ , and recalling that by (14)  $W_{\tau q}^{\varepsilon\psi} = 0$  and that the optimality conditions imply  $W_{\tau}^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) = W_q^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) = 0$ , we have:

$$W^{\varepsilon\psi}(\tau^G, q^G) \simeq W^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) + \frac{1}{2} W_{\tau\tau}^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) (T_\varepsilon)^2 (\bar{\varepsilon}_G - \varepsilon)^2 + \frac{1}{2} W_{qq}^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) (Q_\psi)^2 (\bar{\psi}_G - \psi)^2$$

where  $T_\varepsilon = -1/\varphi v < 0$  and  $Q_\psi = 1$ . Taking a second order approximation of  $W^G(\tau^G, q^G)$  with respect to mean income and social progressiveness in  $G$ ,  $(\bar{\varepsilon}_G, \bar{\psi}_G)$ , at the point  $(\bar{\varepsilon}_G, \bar{\psi}_G)$ ,

we have:

$$\begin{aligned} \Delta_d^{\varepsilon\psi} \simeq & -\left[\frac{1}{2}W_{\tau\tau}^{\varepsilon\psi}(\tau^\varepsilon, q^\psi)(T_\varepsilon)^2(\bar{\varepsilon}_G - \varepsilon)^2 + \frac{1}{2}W_{qq}^{\varepsilon\psi}(\tau^\varepsilon, q^\psi)(\bar{\psi}_G - \psi)^2\right] + \\ & + \lambda\left[\frac{1}{2}W_{\tau\tau}^G(\tau^G, q^G)(T_{\varepsilon_G})^2(\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})^2 + \frac{1}{2}W_{qq}^G(\tau^G, q^G)(Q_{\psi_G})^2(\bar{\psi}_G - \bar{\psi}_{\bar{G}})^2\right] \end{aligned} \quad (30)$$

Given the quadratic functional form of the utility function in (14), we have:  $\frac{1}{2}W_{\tau\tau}^{\varepsilon\psi} = -\nu\varphi$  and  $\frac{1}{2}W_{qq}^{\varepsilon\psi} = -\kappa$ . Thus,  $\Delta_d^{\varepsilon\psi}$  can be rewritten as:

$$\Delta_d^{\varepsilon\psi} \simeq \frac{1}{\nu\varphi} \left\{ [(\bar{\varepsilon}_G - \varepsilon)^2 + \alpha(\bar{\psi}_G - \psi)^2] - \lambda[(\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}})^2 + \alpha(\bar{\psi}_G - \bar{\psi}_{\bar{G}})^2] \right\}$$

where  $\alpha = \kappa\nu\varphi$ . Finally, to preserve linearity, we replace the square brackets with absolute values. Moreover, since identification corresponds to a choice of dimension  $d$  that minimizes  $\Delta_d^{\varepsilon\psi}$ , and since the term  $1/\nu\varphi > 0$  does not depend on  $d$  (and hence does not affect identity choice), without loss of generality we can omit this term and write the approximation to  $\Delta_d^{\varepsilon\psi}$  as in the RHS of (18) in the text. ■

**Proof of Proposition 3.** In the following proof, to ease notation, we denote average types using  $(\varepsilon_G, \psi_G)$  rather than  $(\bar{\varepsilon}_G, \bar{\psi}_G)$ . We summarize the two dimensional groups by the parameters:  $\varepsilon_U, \varepsilon_L = -\chi_\varepsilon\varepsilon_U, \psi_{SP}, \psi_{SC} = -\chi_\psi\psi_{SP}$ , with  $\chi_\varepsilon, \chi_\psi < 1$ . Indeed, the assumption that the mean type is zero implies that  $\pi_U\varepsilon_U + \pi_L\varepsilon_L = 0$  and  $\pi_{SC}\psi_{SC} + \pi_{SP}\psi_{SP} = 0$ , where  $\pi_G$  denotes the population shares of different groups. This implies that  $\varepsilon_L = -\frac{\pi_U}{\pi_L}\varepsilon_U$  and  $\psi_{SC} = -\frac{\pi_{SP}}{\pi_{SC}}\psi_{SP}$ , where  $\chi_\varepsilon = \frac{\pi_U}{\pi_L}$  and  $\chi_\psi = \frac{\pi_{SP}}{\pi_{SC}}$ . Here  $\chi_\varepsilon, \chi_\psi < 1$  follows from the assumption that the poor outnumber the rich and the socially conservatives outnumber the socially progressive. To analyze identification, we must consider all possible cases in the  $(\varepsilon, \psi)$  space.

Case 1. Consider upper class and socially progressive types, namely  $C_1 \equiv \{(\varepsilon, \psi) | \varepsilon > 0, \psi > 0\}$ . A type  $(\varepsilon, \psi)$  from this set identifies along class lines if and only if:

$$\begin{aligned} \Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} \leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} & \Leftrightarrow |\varepsilon - \varepsilon_U| + \alpha|\psi| - \lambda\varepsilon_U(1 + \chi_\varepsilon) \leq |\varepsilon| + \alpha|\psi - \psi_{SP}| - \lambda\alpha\psi_{SP}(1 + \chi_\psi), \\ & |\varepsilon - \varepsilon_U| - \alpha|\psi - \psi_{SP}| + \alpha\psi - \varepsilon \leq \lambda[\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] \equiv \lambda MC. \end{aligned}$$

The term on the right is the meta contrast ratio, which - as we saw before - becomes uniquely dispositive for  $\lambda \rightarrow \infty$ . When  $\lambda$  is finite, the meta contrast ratio entails a general tendency for all types  $(\varepsilon, \psi)$  to identify according to class for  $\varepsilon_U(1 + \chi_\varepsilon) > \alpha\psi_{SP}(1 + \chi_\psi)$  and according to values otherwise.

There are four subcases, each associated with a different consideration of the term on the left.

Case 1.A.  $\{(\varepsilon, \psi) \in C_1 \mid \varepsilon > \varepsilon_U, \psi > \psi_{SP}\}$ . For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\varepsilon}^{\varepsilon\psi} &\leq \Delta_{\psi}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_U + \alpha\psi_{SP} \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\ \Leftrightarrow \alpha &\leq \alpha_1 \equiv \frac{\varepsilon_U}{\psi_{SP}} \frac{\lambda(1 + \chi_{\varepsilon}) + 1}{\lambda(1 + \chi_{\psi}) + 1} . \end{aligned} \quad (31)$$

Case 1.B.  $\{(\varepsilon, \psi) \in C_1 \mid \varepsilon > \varepsilon_U, \psi < \psi_{SP}\}$ . For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\varepsilon}^{\varepsilon\psi} &\leq \Delta_{\psi}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_U - \alpha\psi_{SP} + 2\alpha\psi \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\ \Leftrightarrow \psi &\leq \psi_1 \equiv \psi_{SP} \left[ 1 + \frac{\lambda(1 + \chi_{\psi}) + 1}{2} \frac{\alpha_1 - \alpha}{\alpha} \right] . \end{aligned} \quad (32)$$

Case 1.C.  $\{(\varepsilon, \psi) \in C_1 \mid \varepsilon < \varepsilon_U, \psi > \psi_{SP}\}$ . For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\varepsilon}^{\varepsilon\psi} &\leq \Delta_{\psi}^{\varepsilon\psi} \Leftrightarrow \varepsilon_U + \alpha\psi_{SP} - 2\varepsilon \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\ \Leftrightarrow \varepsilon &\geq \varepsilon_1 \equiv \varepsilon_U \left[ 1 - \frac{\lambda(1 + \chi_{\varepsilon}) + 1}{2} \frac{\alpha_1 - \alpha}{\alpha_1} \right] . \end{aligned} \quad (33)$$

Case 1.D.  $\{(\varepsilon, \psi) \in C_1 \mid \varepsilon < \varepsilon_U, \psi < \psi_{SP}\}$ . For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\varepsilon}^{\varepsilon\psi} &\leq \Delta_{\psi}^{\varepsilon\psi} \Leftrightarrow \varepsilon_U - \alpha\psi_{SP} + 2\alpha\psi - 2\varepsilon \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\ \Leftrightarrow \varepsilon &\geq \alpha(\psi - \psi_{SP}) + \varepsilon_1 . \end{aligned} \quad (34)$$

Case 2. Consider rich and socially regressive types, namely  $C_2 \equiv \{(\varepsilon, \psi) \mid \varepsilon > 0, \psi < 0\}$ . A type  $(\varepsilon, \psi)$  from this set identifies along class lines if and only if:

$$\begin{aligned} \Delta_{\varepsilon}^{\varepsilon\psi} &\leq \Delta_{\psi}^{\varepsilon\psi} \Leftrightarrow |\varepsilon - \varepsilon_U| + \alpha|\psi| - \lambda\varepsilon_U(1 + \chi_{\varepsilon}) \leq |\varepsilon| + \alpha|\psi - \psi_{SC}| - \lambda\alpha\psi_{SP}(1 + \chi_{\psi}), \\ |\varepsilon - \varepsilon_U| - \alpha|\psi - \psi_{SC}| - \alpha\psi - \varepsilon &\leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \equiv \lambda MC . \end{aligned}$$

Once again, the meta contrast ratio entails a general identification tendency for all types  $(\varepsilon, \psi)$ .

There are again four subcases, each associated with a different consideration of the term on the left.

Case 2.A.  $\{(\varepsilon, \psi) \in C_2 \mid \varepsilon > \varepsilon_U, \psi < \psi_{SC}\}$ . For these types, class identity prevails provided:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_U + \alpha\chi_{\psi}\psi_{SC} \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \\
&\Leftrightarrow \alpha \leq \alpha_2 \equiv \frac{\varepsilon_U}{\psi_{SP}} \frac{\lambda(1 + \chi_{\varepsilon}) + 1}{\lambda(1 + \chi_{\psi}) + \chi_{\psi}},
\end{aligned} \tag{35}$$

where  $\alpha_2 > \alpha_1$  because  $\chi_{\psi} < 1$ . This implies that if types in 1.A are identified along class, then types in 2.A are also identified along class. Likewise, if types in 2.A are identified along values, then types in 1.A are also identified along values. To simplify the analysis, then, we will restrict to the more extreme cases  $\alpha < \alpha_1$  versus  $\alpha > \alpha_2$ .

Case 2.B.  $\{(\varepsilon, \psi) \in C_2 \mid \varepsilon > \varepsilon_U, \psi > \psi_{SC}\}$ . For these types, class identity prevails provided:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_U - \alpha\chi_{\psi}\psi_{SP} - 2\alpha\psi \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \\
&\Leftrightarrow \psi \geq \psi_2 \equiv -\psi_{SP} \left[ \chi_{\psi} + \frac{\lambda(1 + \chi_{\psi}) + \chi_{\psi}}{2} \frac{\alpha_2 - \alpha}{\alpha} \right].
\end{aligned} \tag{36}$$

Case 2.C.  $\{(\varepsilon, \psi) \in C_2 \mid \varepsilon < \varepsilon_U, \psi < \psi_{SC}\}$ . For these types, class identity prevails provided:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_U - \alpha\psi_{SC} - 2\varepsilon \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \\
&\Leftrightarrow \varepsilon \geq \varepsilon_2 \equiv \varepsilon_U \left[ 1 - \frac{\lambda(1 + \chi_{\varepsilon}) + 1}{2} \frac{\alpha_2 - \alpha}{\alpha} \right],
\end{aligned} \tag{37}$$

which is isomorphic to (33), but  $\varepsilon_2$  is lower than  $\varepsilon_1$  because  $\alpha_2$  is larger than  $\alpha_1$ .

Case 2.D.  $\{(\varepsilon, \psi) \in C_2 \mid \varepsilon < \varepsilon_U, \psi > \psi_{SC}\}$ . For these types, class identity prevails provided:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_U + \alpha\chi_{\psi}\psi_{SP} - 2\alpha(\psi - \psi_{SC}) - 2\varepsilon \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \\
&\Leftrightarrow \varepsilon \geq -\alpha(\psi - \psi_{SC}) + \varepsilon_2.
\end{aligned} \tag{38}$$

Case 3. Consider lower class and socially progressive types, namely  $C_3 \equiv \{(\varepsilon, \psi) \mid \varepsilon < 0, \psi > 0\}$ . A type  $(\varepsilon, \psi)$  from this set identifies along class lines if and only if:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow |\varepsilon - \varepsilon_L| + \alpha|\psi| - \lambda\varepsilon_U(1 + \chi_{\varepsilon}) \leq |\varepsilon| + \alpha|\psi - \psi_{SP}| - \lambda\alpha\psi_{SP}(1 + \chi_{\psi}), \\
&|\varepsilon - \varepsilon_L| - \alpha|\psi - \psi_{SP}| + \alpha\psi + \varepsilon \leq \lambda [\varepsilon_U(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \equiv \lambda MC.
\end{aligned}$$

There are four subcases, each associated with a different consideration of the term on the

left.

Case 3.A.  $\{(\varepsilon, \psi) \in C_3 \mid \varepsilon < \varepsilon_L, \psi > \psi_{SP}\}$ . For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_L + \alpha\psi_{SP} \leq \lambda [\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] \\ &\Leftrightarrow \alpha \leq \alpha_3 \equiv \frac{\varepsilon_U}{\psi_{SP}} \frac{\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon}{\lambda(1 + \chi_\psi) + 1}, \end{aligned} \quad (39)$$

where  $\alpha_3 < \alpha_1 < \alpha_2$ . Thus, the most stringent condition obtained so far for class identification is  $\alpha \leq \alpha_3$ . The most stringent for social progressive identification is still  $\alpha > \alpha_2$ .

Case 3.B.  $\{(\varepsilon, \psi) \in C_3 \mid \varepsilon < \varepsilon_L, \psi < \psi_{SP}\}$ . For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_L - \alpha\psi_{SP} + 2\alpha\psi \leq \lambda [\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] \\ &\Leftrightarrow \psi \leq \psi_3 \equiv \psi_{SP} \left[ 1 + \frac{\lambda(1 + \chi_\psi) + 1}{2} \frac{\alpha_3 - \alpha}{\alpha} \right]. \end{aligned} \quad (40)$$

Case 3.C.  $\{(\varepsilon, \psi) \in C_3 \mid \varepsilon > \varepsilon_L, \psi > \psi_{SP}\}$ . For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow 2\varepsilon \leq \lambda [\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] - \alpha\psi_{SP} - \chi_\varepsilon\varepsilon_U \\ &\Leftrightarrow \varepsilon \leq \varepsilon_3 \equiv -\varepsilon_U \left[ \chi_\varepsilon - \frac{\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon}{2} \frac{\alpha_3 - \alpha}{\alpha_3} \right]. \end{aligned} \quad (41)$$

Case 3.D.  $\{(\varepsilon, \psi) \in C_3 \mid \varepsilon > \varepsilon_L, \psi < \psi_{SP}\}$ . For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_L - \alpha\psi_{SP} + 2\alpha\psi + 2\varepsilon \leq \lambda [\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] \\ &\Leftrightarrow \varepsilon \leq -\alpha(\psi - \psi_{SP}) + \varepsilon_3. \end{aligned} \quad (42)$$

Case 4. Consider poor and socially conservative types, namely  $C_4 \equiv \{(\varepsilon, \psi) \mid \varepsilon < 0, \psi < 0\}$ . A type  $(\varepsilon, \psi)$  from this set identifies along class lines if and only if:

$$\begin{aligned} \Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow |\varepsilon - \varepsilon_L| + \alpha|\psi| - \lambda\varepsilon_U(1 + \chi_\varepsilon) \leq |\varepsilon| + \alpha|\psi - \psi_{SC}| - \lambda\alpha\psi_{SP}(1 + \chi_\psi), \\ &|\varepsilon - \varepsilon_L| - \alpha|\psi - \psi_{SC}| - \alpha\psi + \varepsilon \leq \lambda [\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] \equiv \lambda MC. \end{aligned}$$

Case 4.A.  $\{(\varepsilon, \psi) \in C_4 \mid \varepsilon < \varepsilon_L, \psi < \psi_{SC}\}$ . For these types, class identity prevails provided:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_L - \alpha\psi_{SC} \leq \lambda [\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] . \\
&\Leftrightarrow \alpha \leq \alpha_4 \equiv \frac{\varepsilon_U}{\psi_{SP}} \frac{\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon}{\lambda(1 + \chi_\psi) + \chi_\psi},
\end{aligned} \tag{43}$$

where  $\alpha_3 < \alpha_4 < \alpha_2$ . Thus, the most stringent condition obtained so far for income identification is still  $\alpha \leq \alpha_3$ . The most stringent condition for social progressive identification is still  $\alpha > \alpha_2$ .

Case 4.B.  $\{(\varepsilon, \psi) \in C_4 \mid \varepsilon < \varepsilon_L, \psi > \psi_{SC}\}$ . For these types, class identity prevails provided:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\chi_\varepsilon\varepsilon_U + \alpha\psi_{SC} - 2\alpha\psi \leq \lambda [\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] . \\
&\Leftrightarrow \psi \geq \psi_4 \equiv -\psi_{SP} \left\{ \chi_\psi + \frac{\lambda(1 + \chi_\psi) + \chi_\psi}{2} \frac{\alpha_4 - \alpha}{\alpha} \right\}.
\end{aligned} \tag{44}$$

Case 4.C.  $\{(\varepsilon, \psi) \in C_4 \mid \varepsilon > \varepsilon_L, \psi < \psi_{SC}\}$ . For these types, class identity prevails provided:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_L - \alpha\psi_{SC} + 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] . \\
&\Leftrightarrow \varepsilon \leq \varepsilon_4 \equiv -\varepsilon_U \left\{ \chi_\varepsilon - \left( \frac{\alpha_4 - \alpha}{\alpha_4} \right) \frac{\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon}{2} \right\}.
\end{aligned} \tag{45}$$

Case 4.D.  $\{(\varepsilon, \psi) \in C_4 \mid \varepsilon > \varepsilon_L, \psi > \psi_{SC}\}$ . For these types, class identity prevails provided:

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_P + \alpha\psi_{SC} - 2\alpha\psi + 2\varepsilon \leq \lambda [\varepsilon_U(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] . \\
&\Leftrightarrow \varepsilon \leq \alpha(\psi - \psi_{SC}) + \varepsilon_4.
\end{aligned} \tag{46}$$

*Patterns of Identification.* The prevalent direction of identification is driven by the location of  $\alpha$  relative to the thresholds  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ . When  $\alpha < \underline{\alpha} \equiv \min(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ , class identification  $d = \tilde{\varepsilon}$  is prevalent. When  $\alpha > \bar{\alpha} \equiv \max(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ , cultural identification  $d = \tilde{\psi}$  is prevalent. To see this, consider these two cases one at the time.

Case  $\alpha < \underline{\alpha}$ . Inspection of the thresholds reveals that  $\underline{\alpha} = \alpha_3$ . As a result, in Cases 1-4 all types located in regions A and B are identified with their class (upper or lower),  $d = \tilde{\varepsilon}$ . Identification in regions C and D is instead ambiguous. A necessary and sufficient condition to have some types identified along the values dimension  $d = \tilde{\psi}$  in each of these regions is that:

$$\min(\varepsilon_1, \varepsilon_2, -\varepsilon_3, -\varepsilon_4) > 0. \tag{47}$$

By inspecting thresholds (33), (37), (41) and (45) it is immediate to see, given our assumption  $\chi_\varepsilon, \chi_\psi < 1$ , that condition (47) is implied by  $\min(\varepsilon_2, -\varepsilon_4) > 0$ , which is in turn implied by:

$$\lambda < \frac{\chi_\varepsilon}{1 + \chi_\varepsilon}. \quad (48)$$

Condition (48) is therefore sufficient to guarantee that some types feature  $d = \tilde{\psi}$  when  $\alpha < \underline{\alpha}$ .

On the other hand, when  $\alpha < \underline{\alpha}$  identification universally occurs along the class dimension, namely  $d = \tilde{\varepsilon}$  for all types provided:

$$\max(\varepsilon_1, \varepsilon_2, -\varepsilon_3, -\varepsilon_4) < 0, \quad (49)$$

which is implied by:

$$\frac{1 - \lambda(1 + \chi_\varepsilon)}{2} + \frac{\lambda(1 + \chi_\varepsilon) + 1}{2} \frac{\alpha}{\alpha_1} < 0. \quad (50)$$

By substituting  $\alpha_1$  in the above expression, the condition becomes:

$$-\lambda \left[ (1 + \chi_\varepsilon) - \alpha \frac{\psi_{SP}}{\varepsilon_U} (1 + \chi_\psi) \right] + 1 + \alpha < 0, \quad (51)$$

which holds provided  $\lambda > \hat{\lambda}_0$ , where  $\hat{\lambda}_0$  is a suitable threshold. Indeed, the expression in square brackets is positive for  $\alpha < \underline{\alpha}$  because  $\chi_\varepsilon, \chi_\psi < 1$ .

Case  $\alpha > \bar{\alpha}$ . Inspection of the thresholds reveals that  $\bar{\alpha} = \alpha_2$ . It is now easy to see that in Cases 1-4 all types located in regions A and C are identified along culture (socially progressive or regressive),  $d = \tilde{\psi}$ . Identification in regions B and D is instead ambiguous. A necessary and sufficient condition to have some types identified along the income dimension  $d = \tilde{\varepsilon}$  in these two regions is that:

$$\min(\psi_1, -\psi_2, \psi_3, -\psi_4) > 0. \quad (52)$$

By inspecting thresholds (32), (36), (40) and (44) one can see, given our assumption  $\chi_\varepsilon, \chi_\psi < 1$ , that condition (52) is implied by  $\min(\psi_3, -\psi_4) > 0$ , which is in turn implied by:

$$\lambda < \frac{\chi_\psi}{1 + \chi_\psi}. \quad (53)$$

Condition (53) is therefore sufficient to guarantee that some types feature  $d = \tilde{\varepsilon}$  when  $\alpha > \bar{\alpha}$ .

On the other hand, when  $\alpha > \bar{\alpha}$  identification universally occurs along the values dimen-



sion, namely  $d = \tilde{\psi}$  for all types provided:

$$\max(\psi_1, -\psi_2, \psi_3, -\psi_4) < 0, \quad (54)$$

which is implied by  $\max(\psi_1, -\psi_2) < 0$ . It is immediate to check that, because in this regime  $\alpha > \alpha_2$ , a sufficient condition for universal  $d = \tilde{\psi}$  identification is given by:

$$\frac{1 - \lambda(1 + \chi_\psi)}{2} + \frac{\lambda(1 + \chi_\psi) + 1}{2} \frac{\alpha_1}{\alpha} < 0. \quad (55)$$

By plugging the expression for  $\alpha_1$  this becomes:

$$-\lambda \left[ (1 + \chi_\psi) - \frac{\varepsilon_U}{\psi_{SP}} \frac{(1 + \chi_\varepsilon)}{\alpha} \right] + \frac{1}{2} \left( 1 + \frac{\varepsilon_U}{\alpha \psi_{SP}} \right) < 0, \quad (56)$$

which holds when  $\lambda > \hat{\lambda}_1 > 0$ , where  $\hat{\lambda}_1$  is a suitable threshold. Indeed, when  $\alpha > \bar{\alpha}$  the expression in square brackets is positive for  $\chi_\varepsilon, \chi_\psi < 1$ .

By conditions (48), (50), (53) and (55) we therefore conclude that across the two regimes  $\alpha < \underline{\alpha}$ ,  $\alpha > \bar{\alpha}$  identification switches from being fully  $d = \tilde{\varepsilon}$  to being fully  $d = \tilde{\psi}$  provided:

$$\lambda > \hat{\lambda} \equiv \max(\hat{\lambda}_0, \hat{\lambda}_1).$$

On the other hand, a change across the two regimes  $\alpha < \underline{\alpha}$ ,  $\alpha > \bar{\alpha}$  preserves some  $d = \tilde{\varepsilon}$  types in all regions B and D and some  $d = \tilde{\psi}$  types in all regions C and D provided:

$$\lambda < \min \left( \frac{\chi_\varepsilon}{1 + \chi_\varepsilon}, \frac{\chi_\psi}{1 + \chi_\psi} \right).$$

After rearranging  $\underline{\alpha}$  and  $\bar{\alpha}$  it is easy to see that they can be expressed as in Equation (19). Which proves the proposition. ■

**Proof of Proposition 4.** Consider a type  $(\varepsilon, \psi)$  identified with group  $G$ . Denote by  $z(\tilde{\psi}|\psi)$  his uncertainty about civil rights and by  $f(\tilde{\varepsilon}|\varepsilon)$  his uncertainty about future income. Then, when these variables are normal, it is immediate to obtain the expressions in the proposition by plugging normal densities in Equation (8).

We can also show that the same qualitative result obtains if one extends the general formulation of Equation (28), as in Proposition 1. First, even in this case beliefs are distorted only along the dimensions of identification. This is due to the fact that  $\varepsilon$  and  $\psi$  are uncorrelated. As a result, the likelihood ratios for the dimension along which identification is not defined are constant at 1, namely  $z(\tilde{\psi}|U) = z(\tilde{\psi}|L)$  and  $f(\tilde{\varepsilon}|SP) = f(\tilde{\varepsilon}|SC)$ . Second, identification with  $L$  (resp.  $U$ ) makes one too pessimistic (resp. optimistic) about social mobility and identification with  $SC$  (resp.  $SP$ ) makes one too averse (resp. open) to civil rights, because

MLRP of  $z(\tilde{\psi}|\psi)$  and  $f(\tilde{\varepsilon}|\varepsilon)$  yields the same characterization of Proposition 1. ■

**Proof of Corollary 3.** Equilibrium beliefs for a type identified with  $G$  are equal to:

$$\varepsilon_G^\theta = \varepsilon + \theta [\bar{\varepsilon}_G - \bar{\varepsilon}_{\bar{G}}], \quad (57)$$

$$\psi_G^\theta = \psi + \theta [\bar{\psi}_G - \bar{\psi}_{\bar{G}}]. \quad (58)$$

When identification occurs along classes only beliefs about income are distorted because - given zero correlation -  $\bar{\psi}_U = \bar{\psi}_L$ . Likewise, when identification is culture based, only beliefs about civil rights are distorted, because  $\bar{\varepsilon}_{SP} = \bar{\varepsilon}_{SC}$ . Thus, identity switches increase polarization in the dimension of identification and reduce it in the other. ■

**Proof of Proposition 5.** To characterize policy distortions, we first need to characterize the equilibrium identification patterns and then the resulting policy distortions as  $\alpha$  varies. To indicate the equilibrium share of voters identified with group  $G$  we use  $\phi_G$ , keeping in mind that this is an endogenous object that depends on  $\alpha$ . In what follows, to ease notation, we continue to denote average types using  $(\varepsilon_G, \psi_G)$  rather than  $(\bar{\varepsilon}_G, \bar{\psi}_G)$ .

**Equilibrium Identification.** Denote by  $l_x$  and  $h_x$  the low and high density for  $x = \varepsilon, \psi$ . Specifically, using the density function in Figure 3 we have  $h_x = \frac{\bar{x}}{\bar{x} + \underline{x}} \frac{1}{\underline{x}}$  and  $l = \frac{\underline{x}}{\bar{x} + \underline{x}} \frac{1}{\bar{x}}$ , so that:

$$\frac{l_x}{h_x} = \chi_x^2. \quad (59)$$

In the text we assume that  $\chi_\varepsilon^2 = \chi_\psi^2$ , but here we generalize the proof to the more general case in which this restriction is relaxed.

We first compute which dimension of identification is prevalent. When  $\alpha < \underline{\alpha}$ , income identity prevails. In Figure 2, the size of class identified agents is larger than  $((\bar{\varepsilon} - \varepsilon_U) + (\varepsilon_L + \underline{\varepsilon}))(\bar{\psi} + \underline{\psi})$ . The size of culture identified agents is smaller than  $(\varepsilon_U - \varepsilon_L)(\bar{\psi} + \underline{\psi})$ . Thus, it is sufficient to show that:

$$((\bar{\varepsilon} - \varepsilon_U) + (\varepsilon_L + \underline{\varepsilon}))(\bar{\psi} + \underline{\psi}) > (\varepsilon_U - \varepsilon_L)(\bar{\psi} + \underline{\psi}).$$

Given that  $\varepsilon_U = \frac{\bar{\varepsilon}\chi_\varepsilon}{2(1+\chi_\varepsilon)}$ ,  $\underline{\varepsilon} = \chi_\varepsilon\bar{\varepsilon}$ ,  $\varepsilon_L = -\chi_\varepsilon\varepsilon_U$  and plugging in such expressions, the previous inequality is simply verified. By the same argument, when  $\alpha > \bar{\alpha}$  culture based identification is prevalent.

Consider the case  $\alpha < \underline{\alpha}$ , in which class identification is prevalent. We need to compute and compare population shares  $(\phi_L, \phi_U, \phi_{SC}, \phi_{SP})$  and check that  $\phi_L \geq \phi_U$  and that  $\phi_{SC} \geq \phi_{SP}$ .

We start with the  $L - U$  cleavage by computing  $(\phi_L, \phi_U)$ . To do so, we focus on two sub-cases. We first consider socially conservative types, which entails computing  $\phi_L(SC)$ ,  $\phi_U(SC)$ . We next consider socially progressive types, which entails computing  $\phi_L(SP)$ ,  $\phi_U(SP)$ .

*Class-identified Socially Conservative types, SC.* Here  $\phi_L(SC) = (\varepsilon_L + \underline{\varepsilon})(0 + \underline{\psi})h_\varepsilon h_\psi + A_L(SC)$ , and  $\phi_U(SC) = (\bar{\varepsilon} - \varepsilon_U)(0 + \underline{\psi})l_\varepsilon h_\psi + A_U(SC)$ , where  $A_L(SC)$  are the  $L$ -identified in

quadrant 4.D while  $A_U(SC)$  are the  $U$ -identified in quadrant 2.D. We first show that, leaving quadrant D aside, the lower class are more numerous than the upper class:

$$(\underline{\varepsilon} + \varepsilon_4)(0 + \underline{\psi})h_\varepsilon h_\psi - (\bar{\varepsilon} - \varepsilon_2)(0 + \underline{\psi})l_\varepsilon h_\psi > 0$$

By plugging in the relevant expressions for thresholds and densities, we find that this is equivalent to:

$$\begin{aligned} \underline{\psi}h_\varepsilon h_\psi [(\underline{\varepsilon} + \varepsilon_4) - (\bar{\varepsilon} - \varepsilon_2)\chi_\varepsilon^2] &> 0, \\ \frac{1}{2}\underline{\psi}h_\varepsilon h_\psi \varepsilon_U(1 - \chi_\varepsilon) \left[ \begin{array}{c} 3\chi_\varepsilon + \lambda(1 + \chi_\varepsilon)^2 \\ -\frac{\alpha}{\alpha_4}(\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon)(1 + \chi_\varepsilon) \end{array} \right] &> 0. \end{aligned}$$

In light of the definitions of  $\alpha_2$  and  $\alpha_4$ , a sufficient condition for the above inequality to hold when  $\alpha < \underline{\alpha}$  is that it holds at  $\alpha < \underline{\alpha} = \alpha_3$ , which is equivalent to:

$$\frac{1}{2}\underline{\psi}h_\varepsilon h_\psi \varepsilon_U(1 - \chi_\varepsilon) \left[ \begin{array}{c} 3\chi_\varepsilon + \lambda(1 + \chi_\varepsilon)^2 \\ -\frac{\lambda(1 + \chi_\psi) + \chi_\psi}{\lambda(1 + \chi_\psi) + 1}(\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon)(1 + \chi_\varepsilon) \end{array} \right] > 0,$$

which is easily verified because  $\chi_\psi, \chi_\varepsilon < 1$ .

Compare now the regions  $A_L(SC)$  and  $A_U(SC)$ . In light of the previous result, for the  $L$ -identified  $SC$  types to outnumber the  $U$ -identified  $SC$  types, it is sufficient to show that  $A_L(SC) - A_U(SC) > 0$ . If  $A_L(SC)$  and  $A_U(SC)$  are triangles, which occurs when  $\alpha$  is low enough that  $\hat{\varepsilon}_4 = \alpha\chi_\psi\psi_{SP} + \varepsilon_4 < 0$  and  $\hat{\varepsilon}_2 = -\alpha\chi_\psi\psi_{SP} + \varepsilon_2 > 0$  we have that  $2A_L(SC) = \alpha(\chi_\psi\psi_{SP})^2 h_\varepsilon h_\psi > 2A_U(SC) = \alpha(\chi_\psi\psi_{SP})^2 l_\varepsilon h_\psi$ . When  $A_L(SC)$  is a trapezoid and  $A_U(SC)$  is a triangle, we have that:

$$A_L(SC) > A_U(SC) \Leftrightarrow \left(2\chi_\psi\psi_{SP} + \frac{\varepsilon_4}{\alpha}\right)(-\varepsilon_4) > \alpha\chi_\psi^2\psi_{SP}^2\chi_\varepsilon^2. \quad (60)$$

This condition is fulfilled because  $-\varepsilon_4 > \chi_\varepsilon\varepsilon_2 > \alpha\chi_\psi\psi_{SP}\chi_\varepsilon$ , and because - after some algebra - one can show that when  $A_L(SC)$  is a trapezoid, namely when  $\hat{\varepsilon}_4 > 0$  or  $\chi_\psi\psi_{SP} > -\varepsilon_4/\alpha$ , we also have that:

$$2\chi_\psi\psi_{SP} + \frac{\varepsilon_4}{\alpha} > \chi_\psi\psi_{SP} > \chi_\varepsilon\chi_\psi\psi_{SP},$$

so that (60) is fulfilled.

Of course, it is a fortiori true that  $A_L(SC) > A_U(SC)$  when they are both trapezoids, because the area  $A_U(SC)$  is largest when - for the same parameter values - this set is a triangle. No other case needs to be considered because when  $A_U(SC)$  is a trapezoid,  $A_L(SC)$

is a trapezoid also.

*Class-identified Socially Progressive types, SP.* Here  $\pi_L(SP) = (\varepsilon_3 + \underline{\varepsilon})\bar{\psi}h_\varepsilon l_\psi + A_L(SP)$ , and  $\pi_U(SP) = (\bar{\varepsilon} - \varepsilon_1)\bar{\psi}l_\varepsilon l_\psi + A_U(SP)$ , where  $A_L(SP)$  are the lower class-identified in quadrant 3.D while  $A_U(SP)$  are the upper class-identified in quadrant 1.D. Once again, we first show that, leaving quadrant D aside, the  $L$  are indeed more numerous than the  $U$ :

$$(\varepsilon_3 + \underline{\varepsilon})\bar{\psi}h_\varepsilon l_\psi - (\bar{\varepsilon} - \varepsilon_1)\bar{\psi}l_\varepsilon l_\psi > 0.$$

By plugging in the relevant expressions for thresholds and densities, this is equivalent to:

$$\begin{aligned} \bar{\psi}h_\varepsilon l_\psi [(\varepsilon_3 + \underline{\varepsilon}) - (\bar{\varepsilon} - \varepsilon_1)\chi_\varepsilon^2] &> 0 \\ \frac{1}{2}\bar{\psi}h_\varepsilon l_\psi \varepsilon_U (1 - \chi_\varepsilon) \left[ \frac{3\chi_\varepsilon + \lambda(1 + \chi_\varepsilon)^2 - \frac{\alpha}{\alpha_3} [\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon] (1 + \chi_\varepsilon)}{\alpha} \right] &> 0, \end{aligned}$$

which is fulfilled for  $\alpha \leq \underline{\alpha}$  because  $\chi_\varepsilon < 1$ .

Compare now the regions  $A_L(SP)$  and  $A_U(SP)$ . In light of the previous result, for the  $L$ -identified  $SP$  types to outnumber the  $U$ -identified  $SP$  types, it is sufficient to show that  $A_L(SP) - A_U(SP) > 0$ . When both  $A_L(SC)$  and  $A_U(SC)$  are triangles, namely when  $\alpha$  is low enough that  $\hat{\varepsilon}_3 = \alpha\psi_{SP} + \varepsilon_3 < 0$  and  $\hat{\varepsilon}_1 = -\alpha\psi_{SP} + \varepsilon_1 > 0$  we have that  $2A_L(SP) = \alpha(\psi_{SP})^2 h_\varepsilon h_\psi > 2A_U(SP) = \alpha(\psi_{SP})^2 l_\varepsilon h_\psi$ . When  $A_L(SP)$  is a trapezoid and  $A_U(SP)$  is a triangle, we have that:

$$A_L(SP) > A_U(SP) \Leftrightarrow \left(2\psi_{SP} + \frac{\varepsilon_3}{\alpha}\right)(-\varepsilon_3) > \alpha\psi_{SP}^2 \chi_\varepsilon^2. \quad (61)$$

This condition is fulfilled because  $-\varepsilon_3 > \varepsilon_1 > \alpha\chi_{SP}$ , and because - after some algebra - one can show that when  $A_L(SP)$  is a trapezoid, namely  $\hat{\varepsilon}_3 > 0$  we have  $\varepsilon_3/\alpha > -\psi_{SP}$ , so that (61) is fulfilled. It is a fortiori true that  $A_L(SP) > A_U(SP)$  when they are both trapezoids, because the area of  $A_U(SP)$  is largest when it is a triangle. No other case needs to be considered because when  $A_U(SP)$  is a trapezoid,  $A_L(SP)$  is a trapezoid also.

*Class Identified: Taking Stock.* In sum, for  $\alpha < \underline{\alpha}$  the poor are in majority among the class identified voters, namely  $\phi_L(\alpha) \geq \phi_U(\alpha)$ .

We now move on to establish, in the same parametric case  $\alpha < \underline{\alpha}$ , that among the culture identified voters, the socially conservatives dominate the social progressives, namely  $\phi_{SC}(\alpha) \geq \phi_{SP}(\alpha)$ . To compute whether socially conservative or socially progressive identification is more prevalent, we again focus on two subcases. The first is within the lower class, which entails computing  $\phi_{SC}(L)$ ,  $\phi_{SP}(L)$ . The second is within the upper class, which entails computing  $\phi_{SC}(U)$ ,  $\phi_{SP}(U)$ .

*Culture identified Lower Class types, P.* Here  $\phi_{SC}(L) = (0 - \varepsilon_4)(\psi_{SC} + \underline{\psi})h_\varepsilon h_\psi + A_{SC}(L)$ , and  $\phi_{SP}(L) = -\varepsilon_3(\overline{\psi} - \psi_{SP})h_\varepsilon l_\psi + A_{SP}(L)$ , where  $A_{SC}(L)$  are the socially conservatives in quadrant 4.D while  $A_{SP}(L)$  are the socially progressive in quadrant 3.D. We first set aside quadrants D and show that socially conservatives are a majority in the other quadrants:

$$-\varepsilon_4(\psi_{SC} + \underline{\psi})h_\varepsilon h_\psi + \varepsilon_3(\overline{\psi} - \psi_{SP})h_\varepsilon l_\psi > 0.$$

By plugging in the relevant expressions for thresholds and densities, we find that this is equivalent to:

$$\begin{aligned} & -\varepsilon_4(\psi_{SC} + \underline{\psi})h_\varepsilon h_\psi + \varepsilon_3(\overline{\psi} - \psi_{SP})h_\varepsilon l_\psi > 0. \\ h_\varepsilon h_\psi \psi_{SP} \chi_\psi \varepsilon_U & \left[ \frac{\chi_\varepsilon - \lambda(1 + \chi_\varepsilon)}{2} (1 - \chi_\psi) + \frac{\alpha}{\alpha_3} \frac{\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon}{2} \frac{\lambda(1 + \chi_\psi)(1 - \chi_\psi)}{\lambda(1 + \chi_\psi) + 1} \right] > 0, \end{aligned}$$

which is easily verified because  $\chi_\psi, \chi_\varepsilon < 1$ . Compare now the regions  $A_{SC}(L)$  and  $A_{SP}(L)$ . In light of the previous result, for the socially conservative identified  $L$  types to outnumber the socially progressive-identified  $L$  types, it is sufficient to show that  $A_{SC}(L) - A_{SP}(L) > 0$ . Consider first the case in which both  $A_{SC}(L)$  and  $A_{SP}(L)$  are triangles. This occurs when  $\alpha$  is sufficiently large that  $\widehat{\varepsilon}_4 = \alpha\psi_{SP} + \varepsilon_4 > 0$  and  $\widehat{\varepsilon}_3 = \alpha\psi_{SP} + \varepsilon_3 > 0$ . In this case we have:

$$A_{SC}(L) - A_{SP}(L) > 0 \Leftrightarrow \frac{\varepsilon_4^2}{\alpha} > \frac{\varepsilon_3^2}{\alpha} \chi_\psi^2,$$

which holds because  $-\varepsilon_4 > -\varepsilon_3 \chi_\psi$ .

Consider next the case in which  $A_{SC}(P)$  is a trapezoid and  $A_{SP}(P)$  is a triangle, which occurs when  $\alpha$  is intermediate so that  $\widehat{\varepsilon}_4 < 0$  and  $\widehat{\varepsilon}_3 > 0$ . In this case, we have that:

$$A_{SC}(L) - A_{SP}(L) > 0 \Leftrightarrow (-2\varepsilon_4 - \alpha\chi_\psi\psi_{SC}) \chi_\psi \psi_{SP} > \frac{\varepsilon_3^2}{\alpha} \chi_\psi^2,$$

which also holds because  $-\varepsilon_4 > \alpha\chi_\psi\psi_{SC}$  and because  $\alpha\psi_{SC} > -\varepsilon_3$ . This implies that  $A_{SC}(L) - A_{SP}(L) > 0$  is a fortiori true when  $A_{SP}(L)$  is a trapezoid, because the latter area is small than in the case the set is a triangle. It is easy to check that it cannot be that  $A_{SC}(L)$  is a triangle and  $A_{SP}(L)$  a trapezoid. We thus conclude that in all cases  $A_{SC}(L) > A_{SP}(L)$  and hence  $\pi_{SC}(L) > \pi_{SP}(L)$ .

*Culture identified Upper class types, U.* Here  $\phi_{SC}(U) = \varepsilon_2(\underline{\psi} + \psi_{SC})l_\varepsilon h_\psi + A_{SC}(U)$ , and  $\phi_{SP}(U) = \varepsilon_1(\overline{\psi} - \psi_{SP})l_\varepsilon l_\psi + A_{SC}(U)$ , where  $A_{SC}(U)$  are the socially conservatives in quadrant 2.D while  $A_{SP}(U)$  are the socially progressives in quadrant 1.D. We first show that, leaving

quadrant D aside, the socially conservatives outnumber the socially progressive:

$$\varepsilon_2(\underline{\psi} + \psi_{SC})l_\varepsilon h_\psi - \varepsilon_1(\bar{\psi} - \psi_{SP})l_\varepsilon l_\psi > 0.$$

By plugging in the relevant expressions for thresholds and densities, we find that this is equivalent to:

$$l_\varepsilon h_\psi \chi_\psi \psi_{SP} \varepsilon_U \left[ \frac{1 - \lambda(1 + \chi_\varepsilon)}{2} (1 - \chi_\psi) + \frac{\alpha}{\alpha_2} \frac{\lambda(1 + \chi_\varepsilon) + 1}{2} \frac{\lambda(1 + \chi_\psi) (1 - \chi_\psi)}{\lambda(1 + \chi_\psi) + \chi_\psi} \right] > 0.$$

Which is indeed fulfilled for  $\alpha \leq \underline{\alpha}$  because  $\chi_\psi < 1$ .

Compare now the regions  $A_{SC}(U)$  and  $A_{SP}(U)$ . In light of the previous result, for the socially conservative -identified  $U$  types to outnumber the socially progressive-identified  $U$  types, it is sufficient to show that  $A_{SC}(U) - A_{SP}(U) > 0$ . Consider first the case in which both  $A_{SC}(U)$  and  $A_{SP}(U)$  are triangles. This occurs when  $\alpha$  is sufficiently large that  $\hat{\varepsilon}_2 = \varepsilon_2 - \alpha \chi_\psi \psi_{SP} < 0$  and  $\hat{\varepsilon}_1 = \varepsilon_1 - \alpha \psi_{SP} < 0$ . In this case we have:

$$A_{SC}(U) - A_{SP}(U) > 0 \Leftrightarrow \frac{\varepsilon_2^2}{\alpha} > \frac{\varepsilon_1^2}{\alpha} \chi_\psi^2,$$

which holds because  $\varepsilon_2 > \varepsilon_1 \chi_\psi$ .

Consider next the case in which  $A_{SC}(U)$  is a trapezoid and  $A_{SP}(U)$  is a triangle, which occurs when  $\alpha$  is intermediate so that  $\hat{\varepsilon}_4 < 0$  and  $\hat{\varepsilon}_3 > 0$ . In this case, we have that:

$$A_{SC}(U) - A_{SP}(U) > 0 \Leftrightarrow (2\varepsilon_2 - \alpha \chi_\psi \psi_{SP}) \chi_\psi \psi_{SP} > \frac{\varepsilon_1^2}{\alpha} \chi_\psi^2,$$

which also holds because  $\varepsilon_2 > \alpha \chi_\psi \psi_{SP}$  and because  $\alpha \psi_{SP} > \varepsilon_1$ . This implies that  $A_{SC}(U) - A_{SP}(U) > 0$  is a fortiori true when  $A_{SP}(U)$  is a trapezoid, because the latter area is small than in the case the set is a triangle. It is easy to check that it cannot be that  $A_{SC}(U)$  is a triangle and  $A_{SP}(U)$  a trapezoid. We thus conclude that in all cases  $A_{SC}(U) > A_{SP}(U)$  and hence  $\pi_{SC}(U) > \pi_{SP}(U)$ .

This proves that for the case  $\alpha \leq \underline{\alpha}$ , we have that  $\phi_L(\alpha) > \phi_U(\alpha)$  and  $\phi_{SC}(\alpha) > \phi_{SP}(\alpha)$ . It still needs to be proved that this is also the case for  $\alpha > \bar{\alpha}$ , but it easily follows from the symmetry of the problem.

**Equilibrium Policy Outcomes.** The previous analysis established that  $(\phi_L - \phi_U)$  and  $(\phi_{SC} - \phi_{SP})$  are always weakly positive, so identity tends to distort redistribution upward and civil rights downward. To see the extent to which these distortions vary with economic shocks, we need to compute how  $(\phi_L - \phi_U)$  and  $(\phi_{SC} - \phi_{SP})$  change from  $\alpha \leq \bar{\alpha}$  to  $\alpha \geq \bar{\alpha}$ .

We do so in both classes of identification equilibria.

1)  $\lambda > \hat{\lambda}$ . In this case, everybody identifies along class lines for  $\alpha < \bar{\alpha}$  and along cultural lines for  $\alpha \geq \bar{\alpha}$ . It is then easy to find that:

$$\begin{aligned} (\phi_L - \phi_U)|_{\alpha \geq \bar{\alpha}} &= \frac{\chi_\varepsilon}{1 + \chi_\varepsilon} (1 - \chi_\varepsilon) > (\phi_L - \phi_U)|_{\alpha \leq \bar{\alpha}} = 0, \\ (\phi_{SC} - \phi_{SP})|_{\alpha \leq \underline{\alpha}} &= \frac{\chi_\psi}{1 + \chi_\psi} (1 - \chi_\psi) > (\phi_{SC} - \phi_{SP})|_{\alpha \geq \bar{\alpha}} = 0. \end{aligned}$$

Clearly, only one policy is distorted, and this policy is the one set with respect to the dimension of identification.

2)  $\lambda < \min\left(\frac{\chi_\varepsilon}{1 + \chi_\varepsilon}, \frac{\chi_\psi}{1 + \chi_\psi}\right)$ . In this case, equilibrium identification is mixed, so there is always overprovision of the public good and underprovision of civil rights. To infer the extent of policy distortions, we characterize population shares across different regimes. We start by computing  $\phi_L - \phi_U$  for  $\alpha \leq \underline{\alpha}$ , and then compute it for  $\alpha \geq \bar{\alpha}$ . When  $\alpha \leq \underline{\alpha}$  we have that:

$$(\phi_L - \phi_U)|_{\alpha \leq \underline{\alpha}} \geq \hat{\phi}_L - \hat{\phi}_U = [(\varepsilon_4 + \underline{\varepsilon}) \underline{\psi} + (\varepsilon_3 + \underline{\varepsilon}) \bar{\psi}] h_\varepsilon - [(\bar{\varepsilon} - \varepsilon_2) \underline{\psi} + (\bar{\varepsilon} - \varepsilon_1) \bar{\psi}] l_\varepsilon, \quad (62)$$

where  $\hat{\phi}_L - \hat{\phi}_U$  is computed by excluding the triangles/trapezoids in quadrants 1-4.D. We already know from the previous analysis that also in the D quadrants there is a majority of poor, which implies the inequality in Equation (62). After some algebra, we can show that this is equal to:

$$\hat{\phi}_L - \hat{\phi}_U = \hat{\pi}_0 + \alpha \psi_{SP}^2 h_\varepsilon (1 - \chi_\varepsilon^2) (1 - \chi_\psi^2) (1 + \lambda), \quad (63)$$

where

$$\hat{\pi}_0 \equiv \varepsilon_U \psi_{SP} (1 + \chi_\psi) [3\chi_\varepsilon + \lambda (1 + \chi_\varepsilon^2)] (1 - \chi_\varepsilon),$$

so that when  $\alpha \leq \underline{\alpha}$  the numerical edge of the poor increases with  $\alpha$  (because a greater importance of civil rights reduces the number of rich faster than the number of poor).

Consider now the case  $\alpha \geq \bar{\alpha}$ . In this case we have that:

$$(\phi_L - \phi_U)|_{\alpha \geq \bar{\alpha}} \leq \phi_L^* - \phi_U^* = [-\psi_4 \underline{\varepsilon} + \psi_3 \underline{\varepsilon}] h_\varepsilon - [-\psi_2 \bar{\varepsilon} + \psi_1 \bar{\varepsilon}] l_\varepsilon, \quad (64)$$

where  $\phi_L^* - \phi_U^*$  is computed by including in the count the culture-identified voters belonging to the triangles/trapezoids in quadrants 1-4.D.

We now show that, in fact, these trapezoids feature a majority of  $L$  voters, so adding them to the count offers an upper bound of the lower class size edge for  $\alpha \geq \bar{\alpha}$ . Consider first the socially conservative voters. Here the  $L$  ( $\varepsilon < 0$ ) culture identified voters with  $\psi \in (\psi_4, 0)$  outnumber the  $U$  ( $\varepsilon > 0$ ) culture identified ones with  $\psi \in (\psi_2, 0)$ . This occurs because the area defining the poor group increases in the product of two terms that increase in  $-\psi_4$  and

$-\varepsilon_4$ , while the area defining the  $U$  group increases in the product of two terms that increase in  $-\psi_2$  and in  $\varepsilon_2$ . Critically, the area of the  $U$  group must be discounted by  $\chi_\varepsilon^2$  due to the higher density of  $L$  voters. As a result, a sufficient condition for the  $L$  voters to outnumber the  $U$  is that  $-\psi_4 > -\chi_\varepsilon\psi_2$  and  $-\varepsilon_4 > \chi_\varepsilon\varepsilon_2$ . It is easy to see, after some algebra, that both conditions are fulfilled.

Consider next the set of socially progressive voters. Here the  $L$  ( $\varepsilon < 0$ ) culture identified voters with  $\psi \in (0, \psi_3)$  outnumber the  $U$  ( $\varepsilon > 0$ ) culture identified voters with  $\psi \in (0, \psi_1)$ . This occurs because the area defining the  $L$  group increases in the product of two terms that increase in  $\psi_3$  and  $-\varepsilon_3$ , while the area defining the  $U$  group increases in the product of two terms that increase in  $\psi_1$  and in  $\varepsilon_1$ . Critically, the area of the  $U$  group must be discounted by  $\chi_\varepsilon^2$  due to the higher density of  $L$  voters. As a result, a sufficient condition for the  $L$  voters to outnumber the  $U$  voters is that  $\psi_3 > \chi_\varepsilon\psi_1$  and  $-\varepsilon_3 > \chi_\varepsilon\varepsilon_1$ . It is easy to see, after some algebra, that both conditions are fulfilled.

These preliminaries therefore establish the upper bound of Equation (64), which can be written as:

$$\phi_L^* - \phi_U^* = \pi_\infty^* + 2\varepsilon_R^2 h_\varepsilon \frac{\lambda \chi_\varepsilon (1 - \chi_\varepsilon^2)}{\alpha},$$

where

$$\pi_\infty^* \equiv \varepsilon \psi_{SP} h_\varepsilon \frac{(1 - 2\lambda) (1 + \chi_\psi) (1 - \chi_\varepsilon)}{2},$$

so that for  $\alpha \geq \bar{\alpha}$  the upper bound on  $L$  voters gradually decreases as  $\alpha$  goes up (because the group of  $U$  voters expands faster than that of  $L$  voters).

It is easy to check that  $\hat{\pi}_0 > \pi_\infty^*$ , so that across the extreme cases  $\alpha = 0$  and  $\alpha \rightarrow \infty$ , the numerical superiority of  $U$  voters among the class identified narrows down. One can then see that  $\chi_\varepsilon, \chi_\psi < 1$  imply the stronger condition:

$$\hat{\pi}_0 > \pi_\infty^* + 2\varepsilon_U^2 h_\varepsilon \frac{\lambda \chi_\varepsilon (1 - \chi_\varepsilon^2)}{\alpha_2},$$

which implies that moving from any  $\alpha \leq \bar{\alpha}$  to any  $\alpha \geq \bar{\alpha} \equiv \alpha_2$  indeed reduces the numerical advantage of the  $U$  voters among the class identified ones. Formally,  $(\phi_L - \phi_U)|_{\alpha \leq \bar{\alpha}} > (\phi_L - \phi_U)|_{\alpha \geq \bar{\alpha}}$ . By a symmetric argument, one can show that  $(\phi_{SC} - \phi_{SP})|_{\alpha \leq \bar{\alpha}} < (\phi_{SC} - \phi_{SP})|_{\alpha \geq \bar{\alpha}}$ . Thus, moving from prevalent class identification to prevalent culture identification reduces excess taxation while it exacerbates the excess restrictiveness of civil rights, and viceversa when identification changes in the opposite direction. This proves the proposition. ■

**Proof of Proposition 6.** By Equation (18), by the fact that  $\bar{\psi}_L = \rho \bar{\varepsilon}_L$ ,  $\bar{\psi}_U = \rho \bar{\varepsilon}_U$ ,  $\bar{\varepsilon}_{SC} = \rho \bar{\psi}_{SC}$ ,  $\bar{\varepsilon}_{SP} = \rho \bar{\psi}_{SP}$ , and by the assumption  $\lambda \rightarrow \infty$ , an individual identifies with his



cultural group if and only if:

$$\alpha (\bar{\psi}_{SP} - \bar{\psi}_{SC}) + (\rho \bar{\psi}_{SP} - \rho \bar{\psi}_{SC}) \geq \alpha (\rho \bar{\varepsilon}_U - \rho \bar{\varepsilon}_L) + (\bar{\varepsilon}_U - \bar{\varepsilon}_L),$$

which, by the assumption in Equation (??) immediately gives rise to the threshold rule (24). In the Robustness Section of this Appendix we also prove some additional general properties of the case with two correlated attributes. Property ii) follows immediately by inspection of Equation (24).

**Proof of Corollary 4.** By inspection of the formulas for distorted beliefs. ■

**Tade Policy Model.** Assume that the utility function  $U(m)$  takes the form:

$$U(m) = -\frac{\delta}{2}(\omega - m)^2$$

with  $\omega, \delta > 0$  and  $\omega$  large. Since income effects are absorbed by consumption of the export good, every individual consumes the same amount of the imported good, namely:

$$\hat{m} = \omega - p/\delta \equiv M(p).$$

so that tariff revenue is  $tp^*(\hat{m} - \sigma)$ .<sup>31</sup> ■

Assume that the average value of  $\eta$  is zero, so that aggregate output in the import competing sector is  $\sigma$ . The government budget constraint can be written as:

$$g = \tau - \frac{\varphi}{2}\tau^2 + t[M((1+t)p^*) - \sigma]p^* \equiv G(\tau, t)$$

where the function  $G(\tau, t)$  denotes overall public revenue from the two policy instruments and where the first two terms capture tax revenue net of the tax distortions.

Let

$$S(t) = U(\hat{m}) - (1+t)p^*\hat{m}$$

denote the consumer surplus from the imported good. Then we can write the expected indirect utility function of type  $(\varepsilon, \eta)$  as

$$W^{\varepsilon\eta}(\tau, t) = Y^{\varepsilon\eta}(\tau, t) + S(t) + \nu G(\tau, t). \quad (65)$$

Only expected income varies across individuals, which simplifies the algebra considerably. In particular, the indirect utility function is separable in  $\varepsilon$  and  $\eta$ , so the tax rate preferred by type  $\varepsilon$ ,  $\tau^\varepsilon$ , is still given by (4) in the previous section. Exploiting the envelope theorem and

---

<sup>31</sup>We implicitly restrict parameters so that in equilibrium there are indeed imports, namely  $\hat{m} > \sigma$ .

simplifying, the tariff preferred by individual  $\eta$  is:

$$t^\eta = [\delta(1 - \eta)\sigma + (\nu - 1)(\delta\omega - p^*)]/(2\nu - 1)p^* \equiv Q(\eta)$$

Finally, in this model we have:  $W_{\tau\tau}^{\varepsilon\psi} = -\varphi\nu$  and  $W_{qq}^{\varepsilon\psi} = -(2\nu - 1)(p^*)^2/\delta$ , and  $T_\varepsilon = -1/\varphi\nu < 0$ ,  $Q_\eta = \sigma\delta/(2\nu - 1)p^*$ . Inserting these expressions in the approximation (30) of the proof of Proposition 2 above, replacing  $\psi$  with  $\eta$ , we get that the parameter  $\alpha$  is replaced by

$$\gamma = \sigma^2\varphi\delta/(2 - 1/\nu)$$

which has the properties discussed in the text. ■

**Proof of Proposition 7.** In what follows, to ease notation, we continue to denote average types using  $(\varepsilon_G, \psi_G)$  rather than  $(\bar{\varepsilon}_G, \bar{\psi}_G)$ . Under class identification, meta-contrast among the  $U$  and  $L$  groups is:

$$|\varepsilon_U - \varepsilon_L|.$$

Under cultural identification, meta-contrast among the  $SP$  and  $SC$  groups is:

$$\alpha |\psi_{SP} - \psi_{SC}| + \gamma\rho |\psi_{SP} - \psi_{SC}|.$$

Under trade identification, meta-contrast among the  $C$  and  $N$  groups is:

$$\alpha\rho |\eta_C - \eta_N| + \gamma |\eta_C - \eta_N|.$$

The prevailing identification occurs along the dimension in which meta-contrast is largest. After some algebra, and given that by assumption  $|\psi_{SP} - \psi_{SC}| = |\eta_C - \eta_N| = |\varepsilon_U - \varepsilon_L|$ , one obtains regions of the proposition. ■

**Proof of Corollary 5.** By inspection, using the expression for distorted beliefs. ■

## 9.1 Robustness

### 9.1.1 Two Correlated Attributes: Equilibrium Policy

Under the maintained assumption that  $\lambda \rightarrow \infty$  everyone identifies along the same dimension. Let  $\tau_d^*$ ,  $q_d^*$  denote the equilibrium policies when identification is along dimension  $d = \tilde{\varepsilon}, \tilde{\psi}$ . Define by  $G_d$  the group when the dimension of identification is  $d$ . Repeating the same steps as above, we get:

$$\begin{aligned}\tau_d^* &= \tau^\circ + 2\theta (\varepsilon_{Gd} - \varepsilon_{\bar{G}_d}) (\pi_{Gd} - 1/2) / \nu\varphi \\ q_d^* &= q^\circ + 2\theta (\psi_{Gd} - \psi_{\bar{G}_d}) (\pi_{Gd} - 1/2)\end{aligned}$$

which implies the following:

**Corollary 6** *Suppose that  $\pi_P, \pi_{SC} > 1/2$ . If condition (??) in the text is satisfied and  $\pi_P \approx \pi_{SC}$ , then  $\tau_\varepsilon^* > \tau_\psi^*$  and  $q_\psi^* < q_\varepsilon^*$ . Moreover,  $\tau_\psi^* \geq \tau^\circ$  and  $q_\varepsilon^* \leq q^\circ$  as correlation is positive (negative).*

It is straightforward to prove this result under the maintained assumption that the poor and socially conservative voters outnumber their opponents.

Consider first the case in which income and social progressiveness are positively correlated. In this case, taxation is excessive and civil rights insufficient, regardless of the dimension of identification because the stereotypes of poor and social conservatives always prevail. If identification changes to culture, provided  $\pi_L \approx \pi_{SC}$ , both taxation and civil rights are reduced. Suppose instead that  $\varepsilon$  and  $\psi$  are negatively correlated. If identification is income based, taxation is excessive, because the poor are in majority, but civil rights are also excessive, because the poor majority is also socially progressive. As identification switches to culture then, civil rights are now underprovided, and taxation is also underprovided.

### 9.1.2 Three Correlated Attributes: General Case

Denote by  $\rho_h$  the correlation between culture and trade exposure. Denote by  $\rho_l$  the correlation of income with the two other traits. Clearly,  $\rho_h \geq \rho_l$ . We also assume that contrast along culture and trade exposure is equal, but it may be higher or lower than contrast along income. We denote by  $K$  the ratio between contrast in  $\psi$  or  $\eta$  and contrast in  $\varepsilon$ . Once again, by the importance of civil rights relative to taxation,  $\alpha$ , but also by the importance of trade protection relative to taxation,  $\gamma$ . We find the following result.

**Proposition 8** *Suppose that  $K > \rho_l$ . Then, the patterns of identification are as follows:*

- i) *If  $\alpha > \max \left[ \gamma, \frac{1-K\rho_l}{K-\rho_l} - \gamma \frac{K\rho_h-\rho_l}{K-\rho_l} \right]$  everybody identifies as a socially conservative or socially progressive.*
- ii) *If  $\gamma > \max \left[ \alpha, \frac{1-K\rho_l}{K-\rho_l} - \alpha \frac{K\rho_h-\rho_l}{K-\rho_l} \right]$  everybody identifies as a nationalist or cosmopolitan.*
- iii) *In all other cases, everybody identifies as upper or lower class.*

**Proof.** In what follows, to ease notation, we continue to denote average types using  $(\varepsilon_G, \psi_G)$  rather than  $(\bar{\varepsilon}_G, \bar{\psi}_G)$ . Under class identification, meta-contrast among the  $U$  and  $L$  groups is:

$$|\varepsilon_U - \varepsilon_L| + \alpha\rho_l |\varepsilon_U - \varepsilon_L| + \gamma\rho_l |\varepsilon_U - \varepsilon_L|.$$

Under cultural identification, meta-contrast among the  $SP$  and  $SC$  groups is:

$$\rho_l |\psi_{SP} - \psi_{SC}| + \alpha |\psi_{SP} - \psi_{SC}| + \gamma \rho_h |\psi_{SP} - \psi_{SC}|.$$

Under trade identification, meta-contrast among the  $C$  and  $N$  groups is:

$$\rho_l |\eta_C - \eta_N| + \alpha \rho_h |\eta_C - \eta_N| + \gamma |\eta_C - \eta_N|.$$

The prevailing identification occurs along the dimension in which meta-contrast is largest. After some algebra, by using the definition of  $K$  in the text, one obtains regions of the proposition provided  $K > \rho_l$ .

Higher  $\rho_h$  increases contrast for culture and trade based identification, not for class identification. As a result, higher  $\rho_h$  favors the two former groupings. Higher  $\rho_l$  increases meta-contrast under all identification regimes. However, it unambiguously reduces the size of the region in which class identification prevails provided it expands both of the sets:

$$\alpha > \max \left[ \gamma, \frac{1 - K\rho_l}{K - \rho_l} - \gamma \frac{K\rho_h - \rho_l}{K - \rho_l} \right], \gamma > \max \left[ \alpha, \frac{1 - K\rho_l}{K - \rho_l} - \alpha \frac{K\rho_h - \rho_l}{K - \rho_l} \right].$$

On the other hand, higher  $\rho_l$  increases the region in which class identification prevails when it reduces the size of both of the above sets. These effects are channels by the changing intercepts and slopes of the rightmost terms in the max operators.

When  $K\rho_h > \rho_l$ , these terms decrease in  $\gamma$  and in  $\alpha$ , respectively. The intercepts of the curves on the axes  $\alpha = 0$  and  $\gamma = 0$  are equal to  $\frac{1-K\rho_l}{K-\rho_l}$  and  $\frac{1-K\rho_l}{K\rho_h-\rho_l}$  (these intercepts are on different axes for the  $\alpha$  and the  $\gamma$  regions above). Thus, higher  $\rho_l$  expands the regime of class identification when it increases both intercepts. A sufficient condition for this to happen is  $K < 1$ . On the other hand, higher  $\rho_l$  reduces the regime of class identification when it reduces both intercepts. A sufficient condition for this to happen is  $K^2 > \rho_h$ . ■

Inspection of Proposition 8 immediately yields the following comparative statics result.

**Corollary 7** *Suppose that  $K > \rho_l$ . Then, higher correlation  $\rho_h$  between  $\psi$  and  $\eta$  favors identification away from class. An increase in the correlation  $\rho_l$  between income and these traits also favors identification away from class when cultural divisions are strong,  $K > \frac{1}{\rho_h}$ , while it favors identification with class when cultural divisions are weak,  $K < 1$ .*

If cultural divisions  $K$  are sufficiently large, economic shocks that concentrate losses on socially conservative voters reduce class identity. Higher correlation  $\rho_h$  between trade exposure and cultural conservatism increases homogeneity among losers from trade and enhance their conflict with the outgroup of winners from trade, making nationalism (or social conservatism) more efficient vehicles of conflict. But even higher correlation  $\rho_l$  between income and cultural

conservatism drives identification away from class when cultural conflict is large,  $K > 1/\rho_h$ . Indeed, this shock also increases similarity among members of the same cultural (/trade) group, so its mobilization displaces the fragile identification with a non-cohesive lower class. Generally speaking, this result indicates that higher correlation among different traits tends to favor identification along the trait that displays the strongest underlying contrast.

Consider now the effect of identification on beliefs and policies.

**Proposition 9** *Suppose that  $K \in \left(\frac{\rho_l}{\rho_h}, \frac{1}{\rho_l}\right)$  and suppose that an increase in  $\alpha$  or  $\gamma$  or  $\rho_h$  or  $\rho_l$  changes identification from class to trade. Then, polarization of beliefs and policy preferences over redistribution drops,  $\tau^{L\theta} - \tau^{U\theta} > \tau^{N\theta} - \tau^{C\theta}$ , while polarization over both trade and cultural policy is enhanced,  $q^{L\theta} - q^{U\theta} > q^{N\theta} - q^{C\theta}$  and  $t^{L\theta} - t^{U\theta} < t^{N\theta} - t^{C\theta}$ .*

**Proof.** In what follows, to ease notation, we continue to denote average types using  $(\varepsilon_G, \psi_G)$  rather than  $(\bar{\varepsilon}_G, \bar{\psi}_G)$ . The policy polarization in any policy instrument increases in the polarization of beliefs of the average member of opposing groups with respect to the trait relevant for the policy. A switch from class to trade identification reduces polarization in income prospects and hence in taxes provided:

$$(1 + 2\theta)(\varepsilon_U - \varepsilon_L) > (1 + 2\theta)\rho_l(\eta_C - \eta_N). \quad (\text{income})$$

The same change enhances polarization among average group members in social progressiveness and benefits from trade and hence in civil rights policy and import tariffs provided:

$$(1 + 2\theta)\rho_l(\varepsilon_U - \varepsilon_L) < (1 + 2\theta)\rho_h(\eta_C - \eta_N), \quad (\text{social progressiveness})$$

$$(1 + 2\theta)\rho_l(\varepsilon_U - \varepsilon_L) < (1 + 2\theta)(\eta_C - \eta_N). \quad (\text{benefits from trade})$$

It is immediate to see that these inequalities are compatible for  $K \in \left(\frac{\rho_l}{\rho_h}, \frac{1}{\rho_l}\right)$ . ■

## 9.2 Evidence on the United States

### Cultural Conflict

The data used to create Figure 4 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.

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 “Immigration” and “Race Relations/Racism” in the macro category “Immigration and Race Relations”; “Abortion” and “Rights of Women Under Attack/Rolling Back” in the macro category “Abortion and Women Rights”; and “Poverty” and “Uneven Distribution of Wealth/Inequality” in “Poverty and Inequality”. To create the trends, we consider for each of the selected issues the share of respondents including such issue among their first three mentions.

The rest of subsection 6.1.1 uses data from the *American National Election Studies* (ANES), waves 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. All the surveys are representative of the national population aged 18 or more and, when computing aggregates, we use individual survey weights. Yearly sample sizes range from roughly 1200 individuals in 2004 to about 5900 individuals in 2012. Below we describe the questions used in the analysis and how the variables based on these questions are defined.

**Desired Size of Government** “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. We rescale the variable between 0 and 1.

**Desired Immigration Levels** “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 and rescale the variable between 0 and 1.

**Abortion Policy** “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]". We rescale the variable between 0 and 1.

**Trade Openness** "Some people have suggested placing new limits on foreign imports in order to protect American jobs. Others say that such limits would raise consumer prices and hurt American exports. Do you favor or oppose placing new limits on imports, or haven't you thought much about this? [1. Favor; 3. Haven't thought much about this; 5. Oppose]"

**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? Which one?" We aggregate answers "Lower Class (Volunteered)", "Average Working", "Working" and "Upper Working" in the macro category "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".

**Party** "Generally speaking, do you usually think of yourself as a democrat, a republican, an independent or what? [Democrat, Republican, Independent, Other party, No preference]"

**Church Attendance** "Lots of things come up that keep people from attending religious services even if they want to. Thinking about your life these days, do you ever attend religious services, apart from occasional weddings, baptisms or funerals? Do you go to religious services every week, almost every week, once or twice a month, a few times a year, or never?" We classify respondents that answer "Every week" or "Almost every week" as "Attending Church" and the rest of the valid cases as "Non Attending".

**Income** Self-reported family income in the year before the survey. It should include income from all sources, including salaries, wages, pensions, Social Security, dividends, interest, and all other income. The original variable consists of roughly 30 income brackets, which we replace by their mid-points. The Cumulative Data File contains a further variable, which classifies observations in 7 year-specific quantiles. Such variable is used to condition on income in Figure 5.

**Education** Grades of education completed or highest degree obtained. When the latter is reported, we convert it in years of education. In the regression analysis, education is accounted for by dummy variables for whether the respondent attained college or more, some college, high school or less.

**Race** Self-identified race. Respondents' self-identified race is used to restrict the analysis in Figure 6 and Figure 7 to the subsample of white adults.

**Age** Self reported age. The square of age is included in the regressions in order to account for non-linear relation often found when dealing with subjective dependent variables.

**Gender** Self-identified gender. Dummy equal to 1 if the respondent identifies as female.

Figure 5 has been constructed as follows. i) We estimated the residuals of the questions on Desired Size of Government and Desired Immigration by conditioning on income and education, on the pooled data from 1998 to 2016. ii) From the distribution of the estimated residuals (pooling all waves together), we have computed the threshold values corresponding to the top and bottom 5% of the residuals. iii) Figure 5 plots the frequency of respondents that, on each wave, fall in these extreme ranges. The figure looks very similar if in step (ii) the thresholds are defined as corresponding to the top and bottom 10%.

## Exposure to Import Competition

**Data** In the analysis of the effects of import shocks from China performed in subsection 6.1.2, we combine data from multiple datasets. All individual level variables are from the Cooperative Congressional Election Study (CCES), a series of surveys with questions on political attitudes, vote choices and individual demographic characteristics. The surveys are administered online on a opt-in basis, but sample matching is employed to assure representativeness of the target population, namely US individuals aged 18 or more. The cross-sectional study has been carried out yearly starting in 2006. Between 2010 and 2014 the CCES also had a longitudinal component, with questions similar to the ones administered in the cross section. We exploit both datasets.

In the repeated cross-section analysis, for each outcome variable of interest, we use the first and the last wave for which comparable questions on that outcome are asked. Following this logic, for willingness to cut spending, importance of abortion, aversion to immigrants and voting decision, we use as first years 2006, 2006, 2007 and 2008 respectively. For each of the four outcomes, the second year of measurement is 2016. In our panel analysis we rely on the data collected in 2010 and 2014.

Autor et al. (2013) measure the change in import exposure in each commuter zone (CZ) by the average change in Chinese import penetration in the CZ's industries, weighted by each industry's share in the CZ initial employment. Thus, the change in export exposure in CZ  $c$  is defined as:

$$\Delta IP_c = \sum_{m \in M} \frac{L_{m,c,00}}{L_{c,00}} \times \frac{I_{m,t_2} - I_{m,00}}{Y_{m,91} + I_{m,91} - X_{m,91}} \quad (66)$$

where the first term in summation is the share of manufacturing industry  $m$  in *total* employment of CZ  $c$ , while the second term is the increase in US imports from China of products typical of  $m$  between 2000 and year  $t_2$ , standardized by  $m$ 's market size in 1991 (i.e, prior to the boom in China's exports). Since the change in penetration is likely to be endogenous, imports are instrumented as in Acemoglu et al (2016), in a way similar to Autor et al (2013). The instrument is obtained by replacing  $(I_{m,t_2} - I_{m,00})$  with  $(I_{m,t_2}^{EU} - I_{m,00}^{EU})$ , namely the increase of Chinese imports in eight European countries over the same period, and all the other terms



in (1) with their values in 1988.

Data on bilateral imports are downloaded from the UN Comtrade database in HS-6 product classification. In particular, we obtain data on imports from China for the US as well as for eight European countries, namely Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain and Switzerland. Such data are treated following a procedure similar to Autor et al. (2013), Acemoglu et al. (2016) and Autor et al. (2017). In particular, to obtain industry-level imports, we apply the crosswalk developed by Pierce and Schott (2012), which maps each HS-6 product into a single SIC industry. In the cross-section analysis we consider changes in imports between 2000 and 2016 (the last year of measurement of our outcome variables). For consistency with the cross section, also in analyzing the panel we consider shocks starting 6 years before the first year of measurement of attitudes, and consider changes in imports between 2004 and 2014. Trade flows are made comparable across time by deflating them with the PCE index.

Import shocks are weighted using data on employment by county and industry contained in the County Business Patterns (CBS). As these employment figures are often reported in brackets, we use the fixed-point methodology developed by Autor et al. (2013) to make them continuous. We also map the counties in commuting zones (CZ), as in Acemoglu et al. (2016).

The CZ-level and county-level controls are obtained from two sources. The first is the online public database of the American University, from which we downloaded the data on the 2000 presidential elections, broken down by county. The second is the dataset built by Autor and Dorn (2013), which contains the offshorability and routine task intensity indexes that we include among the CZ controls in our models.

After combining data from all these sources we obtain a final pooled estimation sample of more than 70 000 individuals for most outcomes studied in the cross section (for abortion the number is lower since the variable is asked only to a subset of the respondents in 2016). This amounts to roughly 60 individuals per CZ and year. The sample size of the panel samples ranges from about 6700 in our models on voting, to more than 9400 in those on aversion to migrants. For willingness to cut spending, the size lies in-between, at 8300, corresponding to around 15 individuals per CZ (and year) on average.

Below, we describe the main dependent variables and the individual controls used in our analysis, all from the CCES. The other variables are described more in detail in the sources indicated above.

**Cut Domestic Spending** “If your state were to have a budget deficit this year it would have to raise taxes on income and sales or cut spending, such as on education, health care, welfare, and road construction. What would you prefer more, raising taxes or cutting spending? Choose a point along the scale from 0 to 100”.

**Aversion to Immigrants.** We extract the first polychoric principal component from two

questions: “What do you think the U.S. government should do about immigration? Grant legal status to all illegal immigrants who have held jobs and paid taxes for at least 3 years, and not been convicted of any felony crimes. [1. Yes; 2. No]” and “What do you think the U.S. government should do about immigration? Increase the number of border patrols on the US-Mexican border. [1. Yes; 2. No]”. Migrant Aversion is the resulting first principal component, rescaled between 0 and 1. Higher values indicate more anti-migrant views.

**Importance of Abortion**(Cross Section Only) “How important is this issue to you? [1. Very High Importance; 2. Somewhat high importance; 3. Somewhat low importance; 4. Very low importance; 5. No importance at all]”. The possible answers indicated above are the ones that appear in the 2016 survey. In 2006, the answer set was [1. Very important; 2. Important; 3. Somewhat important; 4. Not important]. We reconcile the difference by recoding category 5 to 4 in 2016. We also tested the models relying on an alternative harmonization, which consists in grouping together categories 1 and 2 in 2006 and aggregating categories 2 with category 3, and category 4 with category 5, in 2016. In both cases, we reverse the scale, so that higher values indicate higher perceived importance, and treat the variables as continuous after rescaling between 0 and 1. Both approaches give virtually identical results in term of the relevant coefficients’ size and significance. In the text, we present the results obtained using the first alternative.

**Republican Vote** In the cross section, we use a question referring to the Presidential election held in the year of the survey. Since in the first and last year of the longitudinal study (2010 and 2014), no Presidential election takes place, in the panel analysis we look at a question on the State Senate election. In both cases, respondents are asked to report their voting decision, namely the candidate for whom they voted, if they voted, or whether they abstained. The questionnaires also contain information on whether the respondent did not go to vote. In each of the waves considered, panel or cross section, we create a dummy equal to 1 if respondents voted for the republican candidate in the election held in that year, and 0 if they voted another candidate, abstained or did not go to vote.

**Against Migrants and Above Median (Top Third) Cuts** Dummy equal to 1 if the individual obtains the highest score in Migrant Aversion (i.e. answers "No" and "Yes" to the questions concerning granting legal status and increasing border patrols respectively), and answers above a certain threshold in the Cut Domestic Spending scale. We use two different thresholds, namely the median and the 67th percentile of the answers in 2010.

**Educational Attainment** Self-reported highest educational level achieved. Based on this question we create dummy variables for three education levels (no college, some college, college or more).

**Race** Self-identified race. Dummy equal to 1 if the respondent identifies as white.

**Age** Self reported age. We also include its square in order to account for non-linear

relations often found when dealing with subjective dependent variables.

**Gender** Self-reported gender. Dummy equal to 1 if the respondent reports being a female.

**Blog Use (Panel Only)** Dummy equal to 1 if in the first round of the panel survey (2010) the respondent reports having read blog in the 24 hours before the questionnaire was administered.

**Online News (Panel Only)** Dummy equal to 1 if in the first round of the panel survey (2010) the respondent reports having read an online newspaper in the 24 hours before the questionnaire was administered.

**CZ Mover (Panel Only)** Dummy equal to 1 if the commuting zone of residence of the respondent changed between 2010 and 2014.

**Specification and Estimation Method** Following Autor et al. (2017), the two-period repeated cross sectional regression takes the following form:

$$y_{i,c,t} = \beta_0 \Delta IP_c * d_2 + X'_{i,c,t}(\beta_1 + \beta_2 * d_2) + Z'_c \beta_3 * d_2 + \beta_4 d_2 + \alpha_c + u_{i,c,t}$$

where  $i$  denotes the individual,  $c$  the CZ and  $t = 1, 2$  the year,  $y_{i,c,t}$  measures individual attitudes in year  $t$ ,  $\Delta IP_c$  is the variable of interest, namely the increase in import exposure in the CZ,  $d_2$  is a dummy variable that equals 1 in the second period,  $X_{i,c,t}$  is a vector of individual covariates (gender, race, educational attainments, age and age squared),  $Z_c$  is a vector of covariates referring to the CZ in the year 2000.<sup>32</sup> The inclusion of CZ fixed effects ( $\alpha_c$ ) together with time varying coefficients allows us to study how attitudes have changed between 2006 and 2016 within CZ as a result of imports exposure. As in in Autor et al. (2017),  $\Delta IP_c$  is instrumented by replacing the change in US imports from China with the change in imports from China in eight European countries, to account for possible endogeneity through imports demand.

In the analysis of the panel data too we only consider the first (2010) and last (2014) years. Since the panel starts four years later than the cross section, to preserve symmetry with the cross sectional regressions we measure the change in imports between 2004 and 2014. Here we estimate the following specification:

$$\Delta y_{i,c} = \alpha + \beta_0 \Delta IP_c + X'_{i,c,1} \beta_1 + Z'_c \beta_2 + u_{i,c,t}$$

---

<sup>32</sup> As in Autor et al. (2017), the vector  $Z_c$  includes the manufacturing share of employment, the offshorability and routine task indexes of Autor and Dorn (2013) and the interaction between the county vote share for GW Bush in the presidential election and a dummy for Republican victory in that county, all variables measured in 2000. Inclusion of these variables is important for identification, given the nature of the instrument defined in the previous footnote. As in Autor et al. (2017),  $Z_c$  also includes the CZ average of the dependent variable measured at  $t = 1$ ,  $\bar{y}_{c,1}$ , to allow for mean reversion of attitudes over time.

where  $\Delta y_{i,c}$  measures the change in attitudes between 2010 and 2014;  $X_{i,c,1}$  includes the same demographic controls as above measured in the first period, plus  $i$ 's initial attitudes in 2010 to allow for mean reversion, plus a dummy variable for those who changed CZ between 2010 and 2014, as well as the interaction between this dummy and  $\Delta IP_c$ ; the vector  $Z_c$  is defined as above (except that it does not include the CZ average of the dependent variable measured at  $t = 1$ , since mean reversion is already captured by the variable included in  $X_{i,c,1}$ ); again  $\Delta IP_c$  is instrumented as described above. Note that variation in attitudes is measured over a relatively short period (five years). Since identities and opinions are likely to change slowly over time, this is a demanding exercise. Estimation is by 2SLS and standard errors are clustered at the CZ level.

### 9.2.1 Evidence on France

The analysis in subsection 6.2 is carried out using data from the *Dynamiques de mobilisation* (DYNAMOB) panel study, a longitudinal project within the framework of the ELISS internet based survey. Units of observation are individuals, and the study is designed to be representative of the population of Mainland France aged between 18 and 75. We apply survey weights to enhance representativeness of the target population. Our estimation sample comprises all individuals who were in the panel both in the first wave of the survey, in September 2013, and in the wave carried out in May 2017, after the presidential election. To achieve the largest sample size, we complement the answers on attitudes and voting decisions contained in the survey of May 2017 with information on demographic characteristics available in the previous two rounds (December 2016, March 2017). Our final estimation sample consists of around 450 individuals. The main attitudes of interest are measured as follows.

**Economy and Redistribution** is obtained by extracting the first polychoric principal component from the following variables. In the resulting measure, higher values denote more right-wing opinions.

**Competitiveness vs Labor** “In the next years, priority should be given to... [1. the competitiveness of the French economy; 2. the working conditions of employees]”. To increase the sample size, we add an intermediate category consisting of those who don’t take a side. The resulting 3-point variable is recoded so that it ranges between 1 and 5.

**Willingness to Redistribute** “In order to establish a social justice, we should take away from the rich and give to the poor” [1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]”. We reverse the scale so that higher values indicate stronger agreement with the statement. To increase the sample size, we add an intermediate category consisting of those who don’t take a side. The resulting variable ranges between 1 and 5.

**Immigration and Globalization** is obtained by extracting the first polychoric principal component from the following variables. In the resulting measure, higher values denote more liberal and open attitudes.

**Immigration Enriching** “The presence of immigrants is a source of cultural enrichment. [1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]”. We reverse the scale so that higher values indicate stronger agreement with the statement. To increase the sample size, we add an intermediate category consisting of those who don’t take a side. The resulting variable ranges between 1 and 5.

**Too Many Immigrants** “In France there are too many immigrants.[1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]”. To increase the sample size, we add an intermediate category consisting of those who don’t take a side. The resulting variable ranges between 1 and 5.

**French Muslims** “French Muslims are the same as other French people. [1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]”. We reverse the scale so that higher values indicate stronger agreement with the statement. To increase the sample size, we add an intermediate category consisting of those who don’t take a side. The resulting variable ranges between 1 and 5.

**Consequences of Globalization** “The consequences of globalization are extremely negative for France. [1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]”. To increase sample size, we add an intermediate category consisting of those who don’t take a side. The resulting variable ranges between 1 and 5.

**Views on the EU** “All in all, do you think that France has benefitted from its European Union membership? [1. France has benefitted from its European Union membership; 2. France has not benefitted from its European Union membership]”. We switch the answer values, so that 2 indicates a more positive view of the EU. To increase the sample size, we add an intermediate category consisting of those who don’t take a side. The resulting 3-point variable is recoded so that it ranges between 1 and 5.

**Social Progressiveness** is obtained by extracting the first polychoric principal component from the following variables. In the resulting measure, higher values denote more liberal attitudes.

**Role of Women** “Women are mostly made to borne and raise children. [1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]”. To increase sample size, we add an intermediate category consisting of those who don’t take a side. The resulting variable ranges between 1 and 5.

**Same-Sex Adoption** “It is normal that homosexuals can adopt children”. [1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]”. We reverse the scale

so that higher values indicate stronger agreement with the statement. To increase sample size, we add an intermediate category consisting of those who don't take a side. The resulting variable ranges between 1 and 5.

Figure 8 was constructed as follows. First we estimate the residuals after conditioning the three measures described above on income and education (defined below), and standardize such residuals on the pooled sample of 2013 and 2017, so that all three variables have the same variance. Clusters are then defined by applying Ward's minimum variance method on the standardized residuals and stopping the hierarchical partitioning algorithm at 2 clusters. We run the algorithm separately for 2013 and 2017. Finally, the residuals of "Economy and Redistribution" and "Immigration and Globalization" are plotted in Figure 8, with different tones for the two clusters.

The three first principal components "Economy and Redistribution", "Immigration and Globalization" and "Social Progressiveness", are used as dependent variables in the regressions displayed in Table 3 and Table 4. They have been rescaled between 0 and 1 to facilitate the interpretation of regression coefficients. Below, we describe the independent variables used in the regressions.

**Voting Decisions** Voting in the first and second round of the 2012 presidential election is reported in a retrospective question included in the first wave of the survey, in September 2013. These variables contain the name of the candidate voted by the respondent, or whether he abstained. We aggregate the self-reported voting decisions in the first round of the 2012 election as follows (some candidates in the center or that did not belong to the left / right or nationalist / globalist dimensions are omitted): *Left*: Arthaud, Poutou, Mélenchon, Hollande, Joly; *Right*: Sarkozy, Dupont-Aignan; *Nationalist*: Le Pen. For the second round the groups we identify correspond to the two candidates which won the first round. Voting on the 2017 presidential elections is reported in the survey of May 2017. We group first round candidates as follows: *Left*: Arthaud, Poutou, Mélenchon, Hamon; *Right*: Fillon, Asselineau, Dupont-Aignan; *Globalist*: Macron; *Nationalist* Le Pen. Again, for the second round we identify two groups corresponding to the two candidates that ran in that round. For each year, round, and main political group defined above, we define dummy variables indicating if individuals voted that political group in the year and round considered. We retain in our sample non-voters, identified as those who abstained or did not go to vote.

**Income** Self-reported monthly revenue. Reported in 10 roughly equally wide brackets, assigned values from 1 to 10. We take the logarithm of the variable and treat it as continuous.

**Years in Education** Based on a self-reported variable containing the highest level of education attained by the respondent. We convert the variable in years spent in education, based on the usual number of years required to achieve the specified educational level. The variable is treated as continuous.

**Age** The original variable classifies respondents' age based in seven age categories. We build a dummy variable for each age band.

**Gender** Dummy equal to 1 if the respondent is a woman.

**Nationality** Dummy equal to 1 if the respondent is French national.

**Work Status** Dummy equal to 1 if the respondent is employed.

**Region** NUTS 1 region dummy variables.

**Rural** Dummy equal to 1 if the respondent resides in a rural area.