Identity, Beliefs, and Political Conflict*

Nicola Gennaioli[†] and Guido Tabellini[‡]

First version: October 2018 This version: December 2018

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 toward the group's distinctive opinion. The theory yields two main implications: i) voters' beliefs are polarized and distorted along group boundaries; ii) economic shocks that induce new cleavages to emerge also bring about large changes in beliefs and preferences across many policy 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 consistent with these implications.

*We are grateful to participants in seminars at CIFAR and at the Bundesbank and to Raffella Piccarreta for helpful comments, and to Giampaolo Bonomi, Viola Corradini, Daniele d'Arienzo, Carlo Medici, Francesca Miserocchi 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.

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

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

1 Introduction

In the last few decades, the political systems of advanced economies have witnessed momentous changes. Nationalism and populism have gained support almost everywhere, often at the expenses of traditional parties. New dimensions of conflict have emerged, over immigration, globalization and civil rights, in place of the classical divide on redistribution. There is a perception of greater polarization of political views. Some of these phenomena appear to be correlated with economic and social change. U.S. areas more exposed to import competition have become politically more polarized and conservative (Autor et al. 2017). Similar shocks account for the rise of populism and anti-immigration sentiment in Europe (Colantone and Stanig 2018a,b Anelli et al. 2018). Economic insecurity is strongly correlated with cultural conservatism and support for populist parties (Guiso et al. 2017, Gidron and Hall 2017). However, the mechanisms behind these correlations are not clear. Why do losers from free trade become nationalist, hate immigrants and turn socially conservative? Why do they vote for platforms that include policies that seem to run counter some of their interests, such as tax cuts or unsustainable budget deficits?

Recent empirical work tries to shed light on these questions by studying voters' beliefs. This work confirms that beliefs are often inaccurate and systematically distorted, consistent with earlier findings in political sicence (Carpini and Keeter 1979, Luppia and McCubbins 1998). More importantly, even after controlling for individual traits, belief distortions are highly correlated across issues and reflect party affiliation. Alesina et al. (2018a) show that US voters exaggerate social mobility, more so if they identify themselves as right wing. Gimpelson and Treisman (2018) find similar patterns in a cross national survey on perceptions of inequality. Alesina et al. (2018b) show that US and EU voters overestimate the number of immigrants, but right wing voters much more so than those on the left. Kahan (2014) finds large partisan differences in beliefs over global warming in the US.

This paper offers a theory explaining both why voters' beliefs reflect political identities, and why specific economic shocks can induce large changes in policy preferences across many domains, in ways that seem puzzling on the basis of voters' narrow self-interest. The key idea is that voters' beliefs are shaped by endogenous social identities. Individuals routinely identify with a group of people or a party with similar values and interests. Once they do so, they "depersonalize": they attenuate individual specificities, slanting their beliefs toward the "prototypical" or "distinctive" group member. This general idea follows from an established tradition in social psychology, the so called "Social Identity Approach" (cf. Tajfel and Turner 1979, Turner et al. 1987). But social identities are not immutable. Important changes in the economy or society can trigger switches in identities. As this happens, beliefs change and become distorted in new correlated dimensions, amplifying economic shocks.

This mechanism offers a way to think about the phenomena we now observe. Globalization, technological progress, or cultural change create new and large conflicts of interest that reshape social identities. The more disadvantaged members of society see themselves as more similar to a nationalist or socially conservative group, whose distinctive features are the sector of employment or the region of residence, rather than income. The distorted beliefs associated with these new identities dampen demands for traditional redistributive policies, and give rise to a reassessment of political demands over many issues such as trade protection, control of immigration, less progressive civil rights.

To formalize these ideas, we study a model based on two assumptions. First, there are different groups in society defined along income, culture, etc., and each voter perceives greater similarity with a group if his interests are closer to those of the group and if the conflict between such group and the outgroup is stronger. Second, once identified with a group, the individual slants his beliefs toward the stereotypical group condition, which is modelled – following Bordalo et al. (2016, 2018) – as the position that is most distinctive of the group in question.

The model yields two main results. First, identification causes beliefs over particular issues to become more extreme, leading to greater policy conflict. A voter identifying with a group of poor workers exaggerates the risk of poverty because - in thinking about the world - he focuses on the distinguishing feature of his group, poverty itself. The reverse happens for a voter identifying with a group of rich individuals. As a result, beliefs about own income prospects become polarized and redistributive conflict is enhanced relative to a world in which identity does not matter.

Second, economic change causes identification to change, creating polarization in new dimensions, while reducing it in other dimensions. For instance, a trend in globalization clusters individual interests along exposure to foreign competition or immigration relative to, say, the traditional rich-poor divide. As a result, perceived similarities change and social identities switch from a poor-rich conflict to a conflict between globalists and nationalists. This causes individual beliefs about redistribution to become less polarized, while beliefs about globalization become more polarized. Political conflict on redistribution dampens, conflict over globalization intensifies.

More generally, our model implies that – by changing identities – economic events shape individual beliefs and preferences across many dimensions, reflecting the distinctive interests of the new groups voters identify with. For instance, suppose that there is a correlation between demand for trade protection, disliking immigrants, and opposing extensions of civil rights. This may be due to an underlying personality or cultural trait that varies across voters, for instance a belief in "communal" versus "universal" values (Graham et al. 2009, Enke 2018). When voters identify with their income class, conflict is predominantly about redistribution. In this case, a voter's beliefs reflect also his culture but the latter does not polarize politics because income classes - the politically relevant identities - encompass cultural diversity. If however new social demands, a rise in immigration or increased exposure to imports create a divide along culture or nationalism, identification switches. At this point latent cultural factors become catalyst for conflict: they lead to correlated beliefs across many issues, boosting polarization.

In the final part of the paper we make a first pass at taking two predictions of the model to the data. First, using US survey data, we construct a measure of party stereotypes based on group contrast, and show that attitudes of politically identified voters are slanted towards their party stereotype. Second, using survey data on France and the US, we show that switches in political identity or exposure to trade shocks dampen redistributive conflict and exacerbate conflicts over globalization or immigration, as predicted by the theory.

A large literature in political science studies how political identities shape voters' behavior and beliefs, through perceived affinities with political leaders whom the voters trust and identify with (e.g., Achen and Bartels 2016, Johnston et al. 2017). This literature is mostly empirical, here we offer a unified theoretical framework.

Akerlof and Kranton (2000) develop the first economic model of identity. In their approach, identification with a group changes the payoff associated with certain actions. A few recent papers have introduced identity into political economy models (e.g., Shayo 2009, Helpman and Grossman 2018). In these papers, voters obtain positive utility from the status of the group they identify with. This affects policy preferences by causing the voter to internalize the welfare of their group. In our model identification is not related to status (a voter is not penalized from identifying with underdogs), and affects policy demands through beliefs. One advantage of our model is that it sheds light on voters' belief distortions. In general, these papers highlight different mechanisms, which we view as complementary.[1]

Fryer and Jackson (2008) study the role of categorization in decision making, and draw implications for discrimination and social identities. We build on a different mechanism for social perceptions, emphasizing unlikely but representative traits (Bordalo et al. 2016), and focus on different applications. An alternative approach to beliefs is the idea that they are self-serving, namely they promote the interest or self-image of the individual (e.g., Benabou and Tirole 2011, 2016). This approach has been used to shed light on many phenomena, including cross country differences in redistribution (Benabou and Tirole 2006). For the question at hand, though, self-enhancing beliefs have difficulties in explaining why beliefs are correlated with party identity, even after controlling for many individual traits, and why the beliefs and preferences of certain voters seem sometimes to run counter their individual interests.

The paper is organized as follows. In Section 2 we review the key tenets of the "Social

[1] Diermeyer and Li (2018) study a theoretical model of electoral competition with retrospective voting, where partisan identity induces voters to pay more attention to the policies of the candidate they identify with. This mechanism can induce polarization even if candidates are entirely opportunistic.

Identity Approach" and discuss how we apply it to political groups/beliefs. In Section 3 we introduce a simple one-dimensional model of redistribution. We show that identification creates inaccurate beliefs, polarization, and can cause policy distortions. In Section 4 we consider a two dimensional problem, in which redistributive conflict coexist with conflict over culture (or trade). Here identification is non-trivial because voters must decide whether to identify with their class or cultural (trade) group. We show that the dimension of identification changes with economic shocks, and this influences beliefs and policy choice along the two domains. Section 5 studies a three dimensional model. Section 6 takes a first look at the data, and section 7 concludes.

2 The Social Psychology of Identity and Beliefs

Our model is inspired by the "Social Identity Perspective", the leading theory of group identity and intergroup relations. It combines Social Identity Theory (SIT, Tajfel and Tuner, 1979), and Self Categorization Theory (SCT, Turner et al. 1987, Hogg and Abrams 1998). In early work, Tajfel and Wilkes (1963), showed that subjects exaggerate the differences between two physical stimuli (e.g. the length of two segments) when these are assigned to groups differing in stimulus strength (i.e., a short versus a long segments group). Results of this kind, later found also in social contexts, motivated a theory in which social clusters affect the perception of individual members. SIT studies how identification with a group influences individual perceptions and behavior. SCT studies primarily which of the many possible groups (e.g., occupational, religious, etc.) an individual identifies with.

SIT holds that when certain social groups are salient ("us vs. them"), perceptions of self and others are affected. In particular, individuals "depersonalize", so that their perception of self is tainted by group features. First, individual beliefs move toward those of the group (Sherif 1936, Festinger 1950). Second, individuals derive positive self-esteem from the status of their group. Early experiments using the "minimal group paradigm" found that even arbitrary groups influence behavior in this way (Tajfel, Billig, et al. 1971). They also found that individuals tend to favor ingroup versus outgroup members, consistent with the idea that ingroup success enhances self-esteem, but the generality of this finding has been challenged (e.g., Hinkle and Brown 1990). For our analysis, depersonalization supports the key idea that political identification induces voters to see the world through the lenses of the group.

But which groups drive identification and individual perceptions at a given point in time? According to SCT, identification with a group is an individual decision, a form of self-categorization, and it is shaped by the group's accessibility and fit (Oakes, 1987). Accessibility captures the extent to which a group is salient in a given context (e.g., at a soccer match the team's fan group is accessible). Fit captures the extent to which a group reflects social reality, in the sense that it is diagnostic of relevant social differences. This gives rise to the so called "meta contrast ratio": an individual tends to identify with a group that is not only close to him, but that also displays strong differences with respect to the outgroup.

These ideas have two implications for identity politics. First, identification of a voter with a certain group does not only depend on the similarity of his interests with those of the ingroups, but also on the extent of ingroup versus outgroup conflict. For instance, workers at an assembly line fulfill different tasks and have somewhat different interests, but they see themselves as a group because their interests are sharply different from other occupations, such as white collars. Second, when social or economic reality changes, different conflicts of interest become welfare relevant, potentially changing the group with which voters identify.[2]

SCT also sheds light on depersonalization: it views it as a process whereby individual beliefs move toward the group prototype, a stereotyped representation capturing the group's distinctive features. Under high subjective uncertainty, such prototype offers a shared anchor for beliefs, influencing individual opinions (Turner et al. 1987). 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 the positions expressed individually by the members of the same group (Mackie 1986). The distinctive group trait, in fact, is often stereotyped and hence more extreme than the mean position in the group. In political conflict, this effect naturally enhances polarization in beliefs and policy preferences among voters, along the particular issues that characterize prevalent identities.

SCT also remedies some difficulties of SIT. Because identification is driven by real differences between groups, an individual can perceive similarity and hence identify also with underdogs. This may cause the individual to hold beliefs that neither enhance his self-image, nor promote his interest. This often observed pattern is puzzling if, as originally assumed in SIT, individuals only identify with high status groups to enhance self esteem.

In the paper we offer a model of identification based on the Social Identity Approach, and study its implications for individual beliefs, their polarization, and their consequences for electoral politics. Individuals are heterogeneous bundles of characteristics such as income, skills, or cultural traits, which give rise to differences in preferred economic policies. Similarities and dissimilarities in economic interests then affect identification. Depending on economic or social conditions, the individual perceives his interests to be more similar to those of a certain income, skill, or cultural group, causing him to self-categorize in that group. As the voter identifies with a group, he depersonalizes, and reassesses all of his views, slanting them toward the group stereotype, and creating policy distortions.

[2]The importance of cross-group differences in attenuating within group differences, reminiscent of the early Tajfel and Wilkes (1986) experiments, plays a key role in theories of categorization, starting from Rosch (1978). See also Tversky (1977) for a close notion of contextual similarity.

3 One dimensional model

There is a measure one of individuals. Each of them earns stochastic income $\widetilde{y} = 1 + \widetilde{\varepsilon}$, where $\widetilde{\varepsilon}$ is randomly distributed according to pdf $f(\widetilde{\varepsilon}\,|\varepsilon)$ around an individual-specific mean ε. Type ε is distributed in society according to cdf $H(\varepsilon)$ and it has zero mean, so ε captures the person's relative income. Average income in society is 1.

The individual derives utility from private consumption and from 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 utility of private and public consumption are both linear, with marginal utilities equal to 1 and $\nu > 1$ respectively. Under rational beliefs about income risk $\widetilde{\varepsilon}$, the individual's expected utility is equal to:

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

where superscript ε refers to the expected income type and the last term on the right side is the utility from the public good. The rational bliss point of type ε is

$$\tau^{\varepsilon} = \frac{1}{\varphi\nu}(\nu - 1 - \varepsilon) \equiv T(\varepsilon). \tag{2}$$

We assume that the marginal utility of the public good, ν, is large enough that even individuals with very high expected income ε are at an interior optimum, so that $T_{\varepsilon} < 0$ (throughout, subscripts of a function denote partial derivatives).

Given linear private consumption, the socially optimal tax rate τ° maximizes the welfare of the individual with average expected income, $\varepsilon = 0$. Thus, by (2),

$$\tau^{\circ} = \frac{\nu - 1}{\varphi\nu} \tag{3}$$

3.1 Group Identity and Beliefs

Beliefs about uncertain income may be distorted because the agent can focus on the upside of income gain, or on the downside of income loss, depending on his social identity. In line with SCT, group identity induces an exaggerated focus on the income states that are stereotypical of the group the individual identifies with.

Society is partitioned in two groups that differ in expected relative income: the poor $P \equiv \{\varepsilon\,|\varepsilon < \widehat{\varepsilon}\}$, and the rich $R \equiv \{\varepsilon\,|\varepsilon \geq \widehat{\varepsilon}\}$, where $\widehat{\varepsilon}$ is an historically given class boundary. Each group is then summarized by its average member $\varepsilon_P \equiv E[\varepsilon|P]$ and $\varepsilon_R \equiv E[\varepsilon|R]$. The composition of groups is exogenous. To focus on the effects of identification, we abstract from the possibility that some voters may be unengaged and assume that everybody is identified with his income group. In Section 4 we allow voters to change their identity from income to other dimensions.

To capture the effect of depersonalization on beliefs, we model a group's prototypical condition using the CBGS (2016) model of stereotypes, which is based on Kahneman and Tversky's representativeness heuristic. In this model, income realization $\widetilde{\varepsilon}$ is more stereotypical of group $G = R, P$ if it scores higher in the likelihood ratio:

$$\frac{f\left(\widetilde{\varepsilon}\,|\varepsilon_G\right)}{f\left(\widetilde{\varepsilon}\,|\varepsilon_{\overline{G}}\right)}, \tag{4}$$

where $\overline{G} = R, P$ is the alternative group $\overline{G} \neq G$. An income realization is stereotypical for a group if it is more likely to occur in that group relative to the other social group.

BCGS (2016) show that equation (4) sheds light on social stereotypes. Using data from the Moral Foundations Questionnaire (Graham et al. 2012), they show that it accounts for the perceptions about the average liberal or conservative US voter, exaggerating differences between the average liberal and conservative positions. This suggests that stereotypes enhance perceived political polarization.But Equation (4) can also shed light on the Tajfel and Wilkes (1963) experiments, where perceptions of physical stimuli exaggerate differences across groups of stimuli.[3]

Here we apply Equation (4) to beliefs about self. In line with this idea, BCGS (2018) show that stereotypes do not only affect beliefs about others, but also those about self. In their experiment, group stereotypes influence assessements about own ability in different domains of knowledge. Women appear on average too little confident in domains where women do worse than men as a group, while they appear more confident in domains where women do better than men as a group. These effects are stronger when gender is made more salient. We view Tajfel and Turner's polarization effect as resulting from a similar mechanism: when individuals identify with a group, they tilt their beliefs towards the distinctive group traits that differentiate their group relative to the other.

We formalize the effect of depersonalization on beliefs about self by assuming that individuals overweight stereotypical group conditions. In the spirit of BCGS (2016), an individual identified with group G perceives a probability density for his own future income equal to:

$$f^{\theta}\left(\widetilde{\varepsilon}\,|\varepsilon, G\right) = f\left(\widetilde{\varepsilon}\,|\varepsilon\right)\left[\frac{f\left(\widetilde{\varepsilon}\,|\varepsilon_G\right)}{f\left(\widetilde{\varepsilon}\,|\varepsilon_{\overline{G}}\right)}\right]^{\theta} Z, \tag{5}$$

where $Z \geq 0$ is a constant ensuring that the distorted density $f^{\theta}\left(\widetilde{\varepsilon}\,|\varepsilon, G\right)$ adds up to one, and $\theta \geq 0$ is a parameter capturing the strength of stereotypes. When $\theta > 0$, an individual identified with a high income group tends to exaggerate his upward income prospects, while an

[3]The same approach has been shown to account for many doumbfounding laboratory puzzles in probability judgments and for belief distortions in macro and finance (see Gennaioli and Shleifer (2018) for a comprehensive discussion).

individual identified with the poor entertains the opposite belief distortion. We set parameter θ to be the same across individuals, but relaxing this assumption is easy and may yield interesting implications.

Equation (5) characterizes beliefs as follows.

Lemma 1 *A type ε identified with group G expects his future income to be:*

$$\mathbb{E}^{\theta}\left(\widetilde{\varepsilon}\,|\varepsilon, G\right) = \varepsilon + Z \cdot cov\left(\widetilde{\varepsilon}, \left[\frac{f\left(\widetilde{\varepsilon}\,|\varepsilon_G\right)}{f\left(\widetilde{\varepsilon}\,|\varepsilon_{\overline{G}}\right)}\right]^{\theta}\right). \tag{6}$$

If $f\left(\widetilde{\varepsilon}\,|\varepsilon\right)$ is normal with variance $\sigma^2_{\widetilde{\varepsilon}}$, then expected income is:

$$\mathbb{E}^{\theta}\left(\widetilde{\varepsilon}\,|\varepsilon, G\right) = \varepsilon + \theta\left(\varepsilon_G - \varepsilon_{\overline{G}}\right). \tag{7}$$

Equation (6) implies that when the distribution $f\left(\widetilde{\varepsilon}\,|\varepsilon\right)$ exhibits the Monotone Likelihood Ratio Property (MLRP) for $\widetilde{\varepsilon}$ conditional on ε, high income is representative for the rich, low income is representative for the poor. As a result, the poor exaggerate the prospects of income loss, the rich exaggerate the prospects of upward mobility. If $f\left(\widetilde{\varepsilon}\,|\varepsilon\right)$ is normal, then perceptions of expected income simplify to (7). Stereotypes cause individual beliefs to be comparative: the income expected by a rich individual ($\varepsilon \in R$) is distorted towards the average income difference between rich and poor, and viceversa if $\varepsilon \in P$. In the rest of the paper we stick to this convenient formula by assuming that income (and other traits) are normally distributed.[4] Section 6 documents that policy beliefs of US democrats and republicans are consistent with equation (7).

In this model, beliefs concern individual conditions. It is easy, however, to see how the same mechanism also influences beliefs over aggregate conditions such as overall social mobility, immigration, etc. In this case, the density $f\left(.\,|.\right)$ reflects a voter's uncertainty about the target of assessment, and this distribution is distorted towards differences in opinions across groups by equation (7). In this way, individual beliefs are contaminated by group-based errors as found by Alesina et al. (2018a), Alesina et al. (2018b), and others.

Not only beliefs, also policy evaluations are distorted by social identities. By (7), the expected utility of type ε is:

$$W^{\varepsilon\theta}\left(\tau\,|G\right) = \left[1 + \varepsilon + \theta\left(\varepsilon_G - \varepsilon_{\overline{G}}\right)\right]\left(1 - \tau\right) + \nu(\tau - \frac{\varphi}{2}\tau^2) \tag{8}$$

and his bliss point is distorted by stereotypes: $\tau^{\varepsilon\theta} = T[\varepsilon + \theta(\varepsilon_G - \varepsilon_{\bar{G}})]$, where superscript θ denotes stereotyped evaluations.

[4] We can altogether avoid the nuisance of negative income levels by assuming that income is lognormally distributed, which yields the formula of Equation (7) for the logarithm of mean income.

Polarization of policy preferences is a a direct consequence of social identification. Indeed, by (2), the difference in the bliss points of an average rich and an average poor individual is:

$$\tau^{P\theta} - \tau^{R\theta} = (\tau^{P} - \tau^{R})(1 + 2\theta) \tag{9}$$

where τ^{G} denotes the bliss point of a rational average member of group G. The poor are too concerned with income loss and demand too much redistribution, the rich are too optimistic about future income and demand too little redistribution. Overall, stereotypes increase political polarization. The effect is stronger the larger is θ. If exposure to digital media such as Twitter or Facebook strengthens stereotypical thinking θ, then it also leads to greater polarization.

By the same logic, perceived polarization increases even more. Denote by $\widehat{\tau}^{G\theta}$ the perceived position of group G. This perception is formed by stereotyping the true position $\tau^{G\theta}$ of that group, leading to belief $\widehat{\tau}^{G\theta} = \tau^{G\theta} + \theta(\tau^{G\theta} - \tau^{\overline{G}\theta})$. By (9), then, the perceived polarization of policy preferences between the two groups is:

$$\widehat{\tau}^{P\theta} - \widehat{\tau}^{R\theta} = (\tau^{P} - \tau^{R})(1 + 2\theta)^{2}.$$

A rise in θ increases perceived polarization even more than actual polarization. This may explain why perceived polarization in the US and distrust of political opponents seems to have increased more than actual divergence in policy views (Gentzkow 2017).

SIT, and its formalizations in economics (e.g., Shayo 2009), holds that identification causes individuals to favour ingroup members relative to outgroup members. Interestingly, our belief-based mechanism can create similar effects. Specifically, by (8) individual policy preferences can also be written as:

$$W^{\varepsilon\theta}(\tau\,|G) = W^{\varepsilon}(\tau) + \theta\left[W^{G}(\tau) - W^{\bar{G}}(\tau)\right] \tag{10}$$

where $W^{G}(\tau)$ refers to an individual with expected income equal to the group average, ε_{G}. Identification distorts true individual welfare $W^{\varepsilon}(\tau)$ towards welfare difference between ingroup and outgroups. Hence, an individual may support a policy benefitting ingroups at the expenses of outgroups, thereby increasing the gap $W^{G}(\tau) - W^{\bar{G}}(\tau)$, even if it reduces own welfare $W^{\varepsilon}(\tau)$. The reason is that such policies induce the individual to focus on states of nature where the policy is beneficial for him as a typical member of his group.[5] Ingroup bias does not however always occur in our model. Whether it does or not depends on group norms or beliefs about outgroups.

[5]This can be seen generally. 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 $f\left(u|q, \overline{G}\right)$ is the distribution of outgroup $\overline{G}$ utility. The

3.2 Political Equilibrium

Consider the implications of identification for policymaking. We assume that tax policy is set as in standard models of probabilistic voting, namely 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:

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

By replacing Equation (8) in the above objective, we obtain that the equilibrium policy satisfies:

$$\tau^* = \tau^{\circ} - \theta \frac{\left(\varepsilon_P - \varepsilon_R\right)\left(\pi_P - \pi_R\right)}{\nu\varphi} \tag{11}$$

where τ° is the socially optimally policy defined in (3) and $\pi_G = \Pr\left(\varepsilon \in G\right)$ is the population share of group $G = P, R$. Since $\left(\varepsilon_P - \varepsilon_R\right) < 0$, we obtain the following immediate result:

Proposition 1 *If the poor (rich) are more numerous than the rich (poor), then the equilibrium tax is larger (smaller) than in the utilitarian benchmark. Else, the equilibrium tax is socially optimal.*

Identification creates polarization in beliefs between the poor, who demand too much taxation, and the rich, who demand too little taxation. These effects tend to offset each other, because political competition leads parties to converge to the average voter. On net, policy is distorted when excess demand for taxation by the poor is stronger than excess demand of laissez faire by the rich or viceversa. This is the case when one group is more numerous than the other.

Here we realistically assume that the poor are more numerous than the rich, namely the threshold $\widehat{\varepsilon}$ separating the two groups is large enough. The direction of the policy distortion, however, is not critical. What matters is that beliefs and policy preferences are shaped by identification in such a way that, as we will see, support for redistributive policies may

---

perceived expected utility under q is:

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

where $\mathbb{E}\left(u\left|q\right.\right)$ 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 $\mathbb{E}\left(u\left|q\right.\right)$. When the distributions are normal with identical variance, we obtain the exact formula in (10).

suddenly wane as a consequence of shocks making other social cleavages more important.[6] In fact, the mapping beween group identification and the demand for redistribution becomes more complex when, as in reality, there are multiple possible groups of identification and dimensions of conflict. We consider this case next.

4 Two-Dimensional Model

We add a second dimension of heterogeneity in individual interests, which is reflected in an additional policy instrument q. Preferences over q are determined by the quadratic loss:

$$-\frac{\kappa}{2}\left(q-\widetilde{\psi}\right)^2,$$

where $\kappa > 0$ captures the weight of policy q relative to taxation τ. In subsection 4.4 we interpret q as trade policy, and write down a trade model with microfoundations. Here we interpret q as a civil rights policy, with higher q denoting more progressiveness. Regardless the interpretation, this section illustrates how social or economic shocks (e.g., increased cultural divergence, or greater exposure to foreign competition in subsection 4.4) make new cleavages more important, changing identity, beliefs and policies.

Individual preferences over q are captured by the random variable $\widetilde{\psi}$, distributed according to pdf $z\left(\widetilde{\psi}\,|\psi\right)$, where ψ is the individual-specific mean. That is, individuals are uncertain about the consequences of expanding civil rights, and their beliefs can also be susceptible to stereotypes. To use the convenient closed form of Equation (7), unless otherwise noted we also assume $z\left(\widetilde{\psi}\,|\psi\right)$ to be normal.

An individual is summarized by his expected social progressiveness-income pair (ψ, ε), distributed according to the joint cdf $H(\psi, \varepsilon)$. As in the case of ε, the unconditional average of ψ is zero. Thus, ψ is an individual's social progressiveness relative to the mean.

The expected utility of type (ψ, ε) from the policy vector (τ, q) is therefore equal to:[7]

$$W^{\varepsilon\psi}(\tau, q) = (1+\varepsilon)(1-\tau) - \nu(\tau - \frac{\varphi}{2}\tau^2) - \frac{k}{2}(q-\psi)^2. \tag{12}$$

The utility function is conveniently separable, so that the bliss point for taxes only depends on expected income, as in (2), while the bliss point for civil rights only depends on social progressiveness, $q^{\psi} = \psi \equiv Q(\psi)$. The socially optimal tax rate τ° is still given by (3) in the previous section, while the socially optimal civil rights policy is $q^{\circ} = 0$.

[6] A large literature (eg. Alesina and Glaeser 2005) studies why in certain countries demand for redistribution is puzzlingly low. Here we do not address whether group identity may help shed light on this phenomenon.

[7] We omit for simplicity the additive constant $kVar(\psi)$ that obtains when computing $\mathbb{E}\left(q-\widetilde{\psi}\right)^2$.

With two policy dimensions, there are four possible groups, illustrated in Figure 1. Two groups are class based, and partition society into poor $P \equiv \{\varepsilon | \varepsilon < \widehat{\varepsilon}\}$ and rich $R \equiv \{\varepsilon | \varepsilon \geq \widehat{\varepsilon}\}$ voters. The other two groups are culturally based, and partition society into socially conservative, $SC \equiv \left\{\psi | \psi < \widehat{\psi}\right\}$, and socially progressive, $SP \equiv \left\{\psi | \psi \geq \widehat{\psi}\right\}$, voters. The boundary $\widehat{\psi}$ is historically determined, and such that the socially conservatives are a majority, formally $\pi_{SC} > \pi_{SP}$ (it is still true that $\pi_P > \pi_R$).

Figure 1

To simplify the analysis, we assume that groups are divided at the mean $\widehat{\psi} = \widehat{\varepsilon} = 0$ and that these two attributes are uncorrelated. In the appendix we show that in this model similar effects arise with correlation. Introducing correlation leads instead to new insights with three (or more) dimensions, as we show in Section 5.2.

Each group G is summarized by its average member along income and culture (ε_G, ψ_G). The average member can be viewed as a "model" for G, used as a reference when identifying with a group. Consider income based groups. The income of the average poor is $\varepsilon_P \equiv E[\varepsilon|P] < 0$, that of the average rich is $\varepsilon_R \equiv E[\varepsilon|R] > 0$. Due to lack of correlation between attributes, however, th∩1. If instead ε and ψ are correlated, then groups differ in both dimensions. As we shall see, this leads to belief distortions correlated across issues.

These four groups (poor, rich, socially conservatives, socially progressives) and their average types are primitives of our model: they are the unique, historically determined, social categories available for political action. The assumption that other groups cannot be formed is strong, but reflects the difficulty for political entrepreneurs to mold new identities. We could allow for mixed groups, defined along both class and culture (e.g. the rich and socially progressive group). This possibility is related to allowing some people to be unengaed or individualistic, because the latter condition is the limit case of group granularity. To focus on beliefs, we continue to disregard this possibility, but footnote 11 discusses this case. Understanding the deeper forces shaping groups is an important problem for future work.

Because the four groups form alternative partitions of the (ψ, ε) space, an individual belongs to a social class and a cultural group. For instance, individuals featuring $\varepsilon < 0$ and $\psi < 0$ belong both to the poor and to the socially conservative group. Which of these two groups will they identify with? To answer this question we study identification. We denote the dimension of identification by subscript $d = \widetilde{\varepsilon}, \widetilde{\psi}$, so that $d = \widetilde{\varepsilon}$ $(d = \widetilde{\psi})$ means that the individual identifies with his income (cultural) group.

4.1 Similarity and Identification

Consider voter (ε, ψ). If he identifies along class, $d = \widetilde{\varepsilon}$, his ingroup is either rich or poor, and so is his outgroup. If he identifies along culture, $d = \widetilde{\psi}$, his ingroup is either socially conservative or progressive, and so his outgroup. That is, when a voter identifies, he chooses a dimension of conflict that partitions society into ingroups and outgroups. A voter identifying with the poor, views all rich as outgroups, regardless of how the latter identify themselves. Likewise, a socially conservative views all socially progressive people with $\psi > 0$ as his outgroups.

We denote a voter's ingroup and outgroup along dimension d by G_d and $\overline{G}_d$, respectively. Following SCT, we assume that an individual identifies along a certain dimension, say income, when the conflict of interest with his income class is small relative to interclass conflict. Likewise, he identifies along culture when disagreement with his cultural group is small relative to intercultural conflict. Evidently, this definition of similarity and identification relies on two different notions of conflict: that between the individual (ε, ψ) and his ingroup $(\varepsilon_{G_d}, \psi_{G_d})$, and that between such ingroup and the outgroup $(\varepsilon_{\overline{G}_d}, \psi_{\overline{G}_d})$.

We capture conflict between a voter and his ingroup by the the welfare loss:

$$W^{\varepsilon\psi}\left(\tau^{\varepsilon}, q^{\psi}\right) - W^{\varepsilon\psi}\left(\tau^{G_d}, q^{G_d}\right),$$

that the voter experiences when moving from his bliss point policy $(\tau^{\varepsilon}, q^{\psi})$ to the policy $\left(\tau^{G_d}, q^{G_d}\right)$ preferred by his ingroup G_d. Similarly, intergroup conflict is captured by the loss:

$$W^{G_d}\left(\tau^{G_d}, q^{G_d}\right) - W^{G_d}\left(\tau^{\overline{G}_d}, q^{\overline{G}_d}\right),$$

borne by the ingroup $(\varepsilon_{G_d}, \psi_{G_d})$ when moving to the bliss point of the outgroup $(\varepsilon_{\overline{G}_d}, \psi_{\overline{G}_d})$.

Putting these conflicts together, we define the relative distance between voter (ε, ψ) and ingroup G_d as:

$$\Delta_d^{\varepsilon\psi} = \left[W^{\varepsilon\psi}\left(\tau^{\varepsilon}, q^{\psi}\right) - W^{\varepsilon\psi}\left(\tau^{G_d}, q^{G_d}\right)\right] - \lambda\left[W^{G_d}\left(\tau^{G_d}, q^{G_d}\right) - W^{G_d}\left(\tau^{\overline{G}_d}, q^{\overline{G}_d}\right)\right]. \qquad (13)$$

The voter feels similar to group G_d, and hence is more likely to identify with it, when $\Delta_d^{\varepsilon\psi}$ is low. That is, he feels similar to his ingroups along dimension d when conflict with them is small relative to the conflict between ingroups and outgroups in this dimension. The 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:[8]

$$\Delta_d^{\varepsilon\psi} \simeq |\varepsilon - \varepsilon_{G_d}| + \alpha \left|\psi - \psi_{G_d}\right| - \lambda \left(\left|\varepsilon_{G_d} - \varepsilon_{\overline{G}_d}\right| + \alpha \left|\psi_{G_d} - \psi_{\overline{G}_d}\right|\right). \tag{14}$$

The appendix shows that (14) is a shorthand for a second order approximation of the expression on the RHS of (13), and that $\alpha = \kappa\nu\varphi$, where κ is the importance of civil rights, ν is the marginal utility of public consumption and φ captures the severity of the tax distortions. Thus, α captures the relative importance of cultural vs income conflict. If civil rights become a more important source of disagreement (κ rises) and/or taxes become a less important source of disagreement (ν and /or φ rise), then α rises and voters feel stronger similarity along culture than along income.[9]

Individual (ε, ψ) identifies by selecting the dimension $d = \widetilde{\varepsilon}, \widetilde{\psi}$ along which relative distance $\Delta_d^{\varepsilon\psi}$ is minimal. That is, voters identify along income when they perceive stronger similarity with their income class than with their cultural group. This depends on the individual position relative to the group average, as well as on contrast between groups. Thus, strong income differences between income groups, i.e. high $|\varepsilon_P - \varepsilon_R|$, reduce relative distance along the income dimension, $\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi}$, facilitating income based identification. By this mechanism, a low middle class individual perceives himself as more similar to the average poor, so he identifies with him, because they are both very different from the rich. Likewise, strong intercultural clash, i.e. high $|\psi_{SC} - \psi_{SP}|$, reduces relative distance from one's own cultural group, $\Delta_{\widetilde{\psi}}^{\varepsilon\psi}$, facilitating culture based identification. By this mechanism, a moderate and a strongly conservative voter feel more similar, so they identify in the same group, because they are both very different from socially progressives. Due to this mechanism, the presence of a strong cleavage in society induces even moderate voters to take sides in it.

Identification is then characterized as follows.

[8] The appendix shows that (14) is a shorthand for a second order approximation of the expression on the RHS of (13). In performing the second order approximation, the conflict between the individual and his group is approximated around $(\tau^{\varepsilon}, q^{\psi})$, while the conflict among the ingroup and the outgroup is approximated around $\left(\tau^{G_d}, q^{G_d}\right)$. In these approximations we replace the quadratic distances with their absolute value, which greatly simplifies the algebra. Note that here identification is defined with respect to the true utilities of the individual and the group. Little would change if we define conflict of interest using the stereotyped bliss points. In this case, equation (14) becomes:

$$\Delta_d^{\varepsilon\psi} \simeq |\varepsilon - \varepsilon_{G_d}| + \alpha \left|\psi - \psi_{G_d}\right| - \lambda (1 + 2\theta) \left(\left|\varepsilon_{G_d} - \varepsilon_{\overline{G}_d}\right| + \alpha \left|\psi_{G_d} - \psi_{\overline{G}_d}\right|\right),$$

so that the meta contrast term becomes more important when θ is higher.

[9] Note that a higher value of ν corresponds to an increase in the importance of the public good relative to private consumption, which in turn makes income conflict less pronounced (see (bliss)). Likewise, if φ rises and taxes are more distorting, income differences become a less important source of disagreement.

Proposition 2 *There exist two thresholds* $\underline{\alpha}$ *and* $\overline{\alpha}$*, with*

$$\underline{\alpha} \equiv \frac{\lambda\left|\varepsilon_R-\varepsilon_P\right|-\varepsilon_P}{\lambda\left|\psi_{SP}-\psi_{SC}\right|+\psi_{SP}} < \overline{\alpha} \equiv \frac{\lambda\left|\varepsilon_R-\varepsilon_P\right|+\varepsilon_R}{\lambda\left|\psi_{SP}-\psi_{SC}\right|-\psi_{SC}}, \tag{15}$$

and a threshold $0<\widehat{\lambda}<\infty$ *such that:*

i) If $\lambda>\widehat{\lambda}$*, all types* (ϵ,ψ) *identify along income* ($d=\widetilde{\varepsilon}$) *for* $\alpha<\underline{\alpha}$*, while they identify along culture* ($d=\widetilde{\psi}$) *for* $\alpha>\overline{\alpha}$

ii) If $\lambda<\min\left(\pi_R,\pi_{SP}\right)$*, identification is mixed. If* $\alpha<\underline{\alpha}$*, types* (ε,ψ) *with high or low income identify with their class, while some types with* ε *close to* $\widehat{\varepsilon}=0$ *identify with their cultural group. If* $\alpha>\overline{\alpha}$*, types with high or low values of* ψ *identify with their cultural group, while some types with* ψ *close to* $\widehat{\psi}=0$ *identify with their class.*

To begin, focus on the simpler case i). The pattern of identification depends on the relative weight α placed on cultural vs income conflict. If conflict about civil rights is more welfare relevant than conflict about taxes, α is high, voters identify along the cultural dimension. Viceversa, if income conflict is more relevant, they identify based on income. Critically, only intergroup conflict determines identification. An increase in income inequality $|\varepsilon_R-\varepsilon_P|$ between rich and poor causes both $\underline{\alpha}$ and $\overline{\alpha}$ to increase, fostering class identification. But an increase in cultural contrast $|\psi_{SP}-\psi_{SC}|$, stemming for instance from a large inflow of immigrants, goes in the opposite direction, favoring culture based identification. In case ii), instead, similarity between the individual and the group also plays a role, and the dimension of identification differs across individuals. Here too, however, the relative importance of cultural vs income conflict, as captured by α, determines the prevalent dimension of identification.

Thus, as economic or social change renders new cleavages important, the dimension of identification changes. The extent of these changes depends on λ. When intergroup conflict is a key driver of identification, $\lambda>\widehat{\lambda}$, everyobody changes identity in the same way. For instance, as socially progressive positions become more extreme (ψ_{SP} rises), even fairly moderate voters identify as socially conservative. If instead λ is small ($\lambda<\min\left(\pi_R,\pi_{SP}\right)$), then voters with moderate views on the most conflictual issue do not take sides in it and maintain their original identification.[10]

Figure 2 illustrates equilibrium identification when $\lambda<\min\left(\pi_R,\pi_{SP}\right)$.

Figure 2

[10]Group contrast also favors the formation of broad rather than narrow group. One can show that if λ is large then, depending on α, a poor and socially conservative voter prefers to join either the group of poor voters or the group of socially conservative ones, even if a more granular group of "poor and socially conservative" voters is available. The reason is that group contrast is maximized along the single most important dimension, as determined by α.

Panel A reports the case $\alpha \leq \underline{\alpha}$ in which the area of class identification is larger. Panel B reports the case $\alpha \geq \overline{\alpha}$ in which the area of cultural identification is larger.

The figure shows that the voters more likely to change identity are those with extreme income and culture. These voters feel intensely about both issues, so they readily switch to reduce dissonance along the most relevant dimension. Due to an inflow of immigrants, a poor and socially conservative voter feels so strongly about cultural clash that he changes his identification from income to culture. Social groups, then, reshuffle. Poor and rich social conservatives were previously in income conflict but now become allies. The same is true for poor and rich socially progressive voters. Piketty (2018) documents that in the last sixty years in the U.S., U.K., and France, leftwing parties have lost very poor voters while gaining highly educated ones. Our analysis may shed light on this pattern as the result of changes in major political identities over time.

To compute the population shares identified with different groups we need to know the distribution of voter types (ε, ψ) in Figure 2. For tractability, we assume that each trait $x = \varepsilon, \psi$ is distributed in society according to the same "piecewise uniform" distribution displayed in Figure 3 below for expected income ε:

Figure 3

The distribution has mean zero, and $\overline{x} > \underline{x} > 0$ ensures that it is left skewed, so that poor and socially conservative voters are in majority. We then obtain the following result:

Proposition 3 *Under piecewise uniform distributions such as in Figure 3, the majority of voters are identified along income for $\alpha \leq \underline{\alpha}$ and along culture for $\alpha \geq \overline{\alpha}$. Furthermore, for any $\alpha \leq \underline{\alpha}$ and $\alpha \geq \overline{\alpha}$ the poor and socially conservative are a weak majority, namely $\pi_P(\alpha) \geq \pi_R(\alpha)$ and $\pi_{SC}(\alpha) \geq \pi_{SP}(\alpha)$.*

Even if some people stick to their original identities, a sufficiently large increase in the importance of cultural conflict α turns the identity of the majority of voters from income to culture. Within each subgroup, the socially conservative always dominate the social progressives while the poor dominate the rich. The latter feature is important to characterize equilibrium policy distortions.

Propositions 2 and 3 show that economic change, as measured for instance by α, changes cleavages and thus the dimension of identification, yielding the following chain reaction:

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

We now analyze the last two steps of of this chain. In the next subsection, we show how changes in identification cause a change in beliefs. In subsection 4.3 we show how this leads to changes in equilibrium policy.

4.2 Beliefs

For a given group of identification, beliefs overweigh the group's stereotypical state.The only difference with the analysis of Section 3 is that now identification influences beliefs in two dimensions: income risk as well as civil rights. As a result, when economic change induces identity changes, beliefs in both dimensions are affected.[11]

To see the how this works, suppose that the importance of cultural policy α is low, so that income based identification is prevalent. Then, by equation (7) the beliefs of a poor and socially conservative voter, $\varepsilon < 0$ $\psi < 0$, identifying with the poor are:

$$\mathbb{E}^{\theta}\left(\widetilde{\varepsilon}\,|\varepsilon,\psi,P\right) = \varepsilon+\theta\left(\varepsilon_P-\varepsilon_R\right)<\varepsilon, \tag{16}$$

$$\mathbb{E}^{\theta}\left(\widetilde{\psi}\,|\varepsilon,\psi,P\right) = \psi \tag{17}$$

As in the one dimensional model, the voter exaggerates downward income risk, the larger is the income contrast $|\varepsilon_P-\varepsilon_R|$ between rich and poor. The voter's view on civil rights, on the other hand, is undistorted, because cultural attitudes do not vary between rich and poor. In the presence of correlation, income identification would also distort beliefs about ψ. Here we stick to zero correlation because it helps us illustrate our results in the simplest way.

Suppose then that new social demands (e.g. the right to abortion) or an inflow of immigrants create strong cultural conflict, so that cultural policy becomes more important, namely $\alpha > \overline{\alpha}$. Then, an equally poor and socially conservative voter behaves differently. He perceives stronger similarity with social conservatives, even if rich, and thus identifies with his cultural group. His beliefs now switch to:

$$\mathbb{E}^{\theta}\left(\widetilde{\varepsilon}\,|\varepsilon,\psi,SC\right) = \varepsilon, \tag{18}$$

$$\mathbb{E}^{\theta}\left(\widetilde{\psi}\,|\varepsilon,\psi,SC\right) = \psi+\theta\left(\psi_{SC}-\psi_{SP}\right)<\psi. \tag{19}$$

As identification changes from income to culture, beliefs change along both dimensions even though the underlying type (ε,ψ) is the same. He slants his views toward the socially conservative stereotype, the more so the stronger is the contrast $|\psi_{SC}-\psi_{SP}|$. At the same time, the voter moderates his demand for redistribution, because he now identifies with a group that also includes rich people.

The result below generalizes this discussion to all types.

[11]Formally, the stereotypical income $\widetilde{\varepsilon}$ and culture $\widetilde{\psi}$ of group G_d are those scoring high in the likelihood ratios $\frac{f\left(\widetilde{\varepsilon}|\varepsilon_{G_d}\right)}{f\left(\widetilde{\varepsilon}|\varepsilon_{\overline{G}_d}\right)}$ and $\frac{z\left(\widetilde{\psi}|\psi_{G_d}\right)}{z\left(\widetilde{\psi}|\psi_{\overline{G}_d}\right)}$, where $\overline{G}_d$ is again the opposing social group. A type (ε,ψ) identifying with G_d overweights income levels or cultural views that score high in these likelihood ratios.

Proposition 4 *i) when ε and ψ are uncorrelated, beliefs are only distorted in the dimension of identification and not in the other dimension. ii) In the dimension of identification, beliefs are more distorted the larger is the contrast between groups.*

When identification changes from income to culture, the entire beliefs system changes. Voters' views on civil rights become more extreme, those on redistribution more moderate. If many voters switch, beliefs over civil rights become more polarized across identified voters, while polarization in their attitudes towards redistribution dampens.

This mechanism can account for the evidence in Johnston et al. (2017), whereby among "politically engaged" voters income is a poor predictor of preferences for redistribution. The authors argue that this pattern arises because engaged voters join parties based on their views on civil rights, and then inherit also the party's redistributive policy, regardless of economic self interest. Our model yields precisely this pattern as a result of identification with a party's cultural platform. In Section 5 we take some of these predictions to the data.

4.3 Equilibrium Policy

We now study the last step of the causal link, namely how changes in identification and beliefs translate into changes in equilibrium policy. To do so, denote by $W^{\varepsilon\psi\theta}(\tau, q\,|G_d)$ the expected utility of type (ε, ψ) when he identifies with group G_d so that his beliefs about income and culture are those distorted by stereotypes, $E^\theta(\widetilde{\varepsilon}\,|\varepsilon, \psi, G_d)$ and $E^\theta\left(\widetilde{\psi}\,|\varepsilon, \psi, G_d\right)$.[12]

Let ϕ_d denote the fraction of population that identifies along dimension d. Then the equilibrium policy satisfies:

$$(\tau^*, q^*) = \arg\max_{\tau,q} \sum_{d=\tilde{\varepsilon},\tilde{\psi}} \phi_d \int W^{\varepsilon\psi\theta}(\tau, q\,|G_d)\, dH(\psi, \varepsilon)\,. \tag{20}$$

The first order conditions of the problem imply:

$$\tau^*_\alpha = \tau^\circ - \frac{\theta(\varepsilon_P - \varepsilon_R)[\pi_P(\alpha) - \pi_R(\alpha)]}{\nu\varphi}, \tag{21}$$

$$q^*_\alpha = q^\circ + \theta(\psi_{SC} - \psi_{SP})[\pi_{SC}(\alpha) - \pi_{SP}(\alpha)]\,. \tag{22}$$

where τ° and q° denote the socially optimal policies, and $\pi_P(\alpha)$, $\pi_R(\alpha)$, $\pi_{SC}(\alpha)$, $\pi_{SP}(\alpha)$ are the shares of voters identified with the different groups as a function of α.

The presence of stereotypes, $\theta > 0$, distorts equilibrium policy. Because by Proposition 3 the poor outnumber the rich, and the socially conservative outnumber the progressives, taxation is excessive and civil rights are too restricted.

[12] To obtain $W^{\varepsilon\psi\theta}(\tau, q\,|G_d)$, just replace in (12) the true ε and ψ with their stereotyped versions.

Importantly, policy distortions depends on the parameter α, because the latter determines the most relevant political conflict and the prevalent identities. The appendix then proves the following result.

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

i) If $\lambda > \widehat{\lambda}$, all voters identify along income or culture. Thus, only taxation is distorted if $\alpha < \underline{\alpha}$, while only civil rights are distorted if $\alpha > \overline{\alpha}$.

ii) If $\lambda < \min(\pi_R, \pi_{SP})$, 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 $\alpha > \overline{\alpha}$ reduces both τ^*_{α} and q^*_{α}, so that taxes become less distorted and civil rights more distorted.

Policy distortions follow the pattern of distorted beliefs. Poor-identified voters exaggerate the benefit of redistribution, socially conservative-identified voters exaggerate the risks of liberal policies. The rich and socially progressive entertain the opposite belief distortions, but they are a minority. Hence equilibrium redistribution is excessive, civil rights too limited, to an extent that depends on whether identification is homogeneous of heterogeneous.

Crucially, by influencing beliefs, changes in identity amplify the effects of economic shocks on equilibrium policies. A change from prevalent class based to prevalent culture based identification causes redistribution and its distortions to fall, but distortions from too limited civil rights increase. In the Appendix we show that similar results obtain when ε and ψ are correlated.

5 Three Dimensional Conflict: Adding Globalization

Recent political trends in many advanced countries have highlighted the emergence of a new political cleavage, nationalism vs globalization. We now show that a setting isomorphic to the one studied above can shed light on this important phenomenon. We write down a simple model where import tariffs are an additional policy instrument. In this setting, increased exposure to foreign competition gives rise to a new trade-related cleavage. This, in turn, can trigger a change in political identities. Class base identities wane, and globalist-nationalistic conflict materializes, with rich and poor losers from trade openness joining the same nationalistic group.

To illustrate these results, we proceed in two steps. The next subsection presents a model that features a policy conflict between trade policy and income redistribution. In subsection 5.2 we add culture and study identification, beliefs, and policies in this tri-dimensional model. As we will see, this setting is most interesting when different dimensions are correlated.

5.1 Globalization and Redistribution

There is 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 price of the exported good is 1, the price of the imported good is p. Government policy consists of two policy instruments. The first is an ad valorem tariff t, so that domestic prices of imports are $p = (1+t)p^*$, where p^* is the world price of imports. The second instrument is the distorting income tax τ of the previous subsections. Revenue from both instruments is used to finance a public good g.

Individuals differ in their sources of income. First, each individual earns a stochastic and taxable income $(1+\tilde{\varepsilon})$ from employing an input in the export sector. As before, the average of $\widetilde{\varepsilon}$ captures the individual's income type ε. Second, each individual earns 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 from this input is 1 if employed in the exported good sector, and p if employed in the import competing sector.

The probability of employment in the import competing sector is $\sigma(1-\widetilde{\eta})$, where $\sigma \in (0,1)$ captures generalized exposure to import competition, while $\widetilde{\eta}$ is a random variable with individual specific mean η.[13] An individual with higher η is more employable in the export sector, perhaps because he is more mobile or skilled. Thus, the expected income of type η from the specific input is $1+\sigma(1-\eta)(p-1)$. The average value of η is zero, so that aggregate output in the import competing sector is σ.

A voter's type is summarized by the income-trade openness vector (ε, η). With this notation, the expected income of individual type (ε, η) is equal to:

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

and his budget constraint (based on expected income) is:

$$Y^{\varepsilon\eta} = x^{\varepsilon\eta} + pm^{\varepsilon\eta}$$

where $x^{\varepsilon\eta}$ and $m^{\varepsilon\eta}$ denote consumption of the exported and imported good respectively.

The preferences of a generic voter are:

$$w^{\varepsilon\eta} = x^{\varepsilon\eta} + U(m^{\varepsilon\eta}) + \nu g,$$

where νg denotes the utility of public consumption g, as in the previous section, and $\nu > 1$ and large. To match this model with the previous setting we further assume that the utility

[13] To rely on normality, which yields the exact distortion in beliefs of Equation (7), we assume that the variance is sufficiently low that $\sigma(1-\widetilde{\eta})$ is between zero and one with very large probability. We could also specify a bounded distribution for $\widetilde{\eta}$, such as a beta, and obtain very similar results.

function from the imported good is quadratic, and takes the form:

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

with $\omega, \delta > 0$ and ω large.

Since income effects are absorbed by consumption of the export good, every individual consumes the same amount of the imported good, namely:

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

so that tariff revenue is $tp^*(\widehat{m} - \sigma)$.[14]

With this notation, 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

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

denote the consumer surplus from the imported good. Then we can write the expected indirect utility function of type (ε, η) as

$$W^{\varepsilon\eta}(\tau, t) = Y^{\varepsilon\eta}(\tau, t) + S(t) + \nu G(\tau, t). \tag{23}$$

Only expected income varies across individuals, which simplifies the algebra considerably. In particular, the indirect utility function is separable in ε and η, so the tax rate preferred by type ε, τ^ε, is stil given by (2) in the previous section. Exploiting the envelope theorem and simplifying, the tariff preferred by individual η is:

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

An individual earning less from the import competing sector (higher η) demands a lower tariff, $Q_\eta < 0$; and if he has a higher taxable income (ε is higher) he demands a lower income tax, $T_\varepsilon < 0$.

In this model with two policy instruments, the distance function $\Delta_d^{\varepsilon\eta}(\tau, t)$ that determines group identification can be approximated by equation (14) above in which η replaces ψ and in

[14]We implicitly restrict parameters so that in equilibrium there are indeed imports, namely $\widehat{m} > \sigma$.

which the relative weight α is replaced by the importance of trade policy γ relative to income taxes, which in the Appendix we show to be equal to:

$$\gamma = \sigma^2 \varphi \delta / (2 - 1/\nu). \tag{24}$$

A higher value of γ makes identification on the nationalism/globalism dimension, $d = \tilde{\eta}$, more likely. Among other things, γ increases if: 1) σ increases (i.e. if individuals are more exposed as producers to the import competing sector); 2) φ increases (i.e., taxes are more distorting, which reduces conflict over income taxation); 3) δ increases (the elasticity of import demand drops, which makes tariffs less distorting, increasing conflict over trade policy).

All the results of the previous section remain unchanged, except that the interpretation of the shocks, beliefs and policies are now different. In particular, if individuals become more exposed to import competition (σ rises), then identification switches to nationalism vs globalism. Conflicting groups change, so that rich and poor opponents of globalization join, and rich and poor supporters of globalization unite. If this happens, the demand for redistributive income taxes drops and, if nationalists outnumber globalists, the overall demand for protectionist policies rises. As a result, equilibrium policy becomes less redistributive along the income lines, and more nationalistic and protectionist.

In Section 6 we take some of these predictions to the data.

5.2 Three Dimensional Conflict, Shocks and Beliefs

We can add cultural conflict to the model, by simply adding the loss function $-\frac{\kappa}{2}\left(q - \widetilde{\psi}\right)^2$ to the welfare function in (23). Now there are three policy instruments: income tax τ, import tariff t, and civil rights q. A voter's preferences are pinned down by the vector $(\varepsilon, \eta, \psi)$ reporting his expected income, his exposure to trade, and his social progressiveness. Identification can occur along any of these dimensions. The distance function is now a direct extension of (14), in which distances along the openness trait η are weighed by parameter γ, while distances along culture ψ are weighed by $\alpha = \kappa\nu\varphi$ as before. Stronger exposure to import competition (higher σ) increases the welfare relevance of trade γ, favoring openness-based identification. On the other hand, social change making civil rights more important (higher κ), increases α, favoring culture based identification.

We now show that in this setting, shocks can exert far reaching consequences on beliefs. The main implications arise in the presence of non-zero correlation between different traits. In particular, a change in identification toward traits that are more strongly correlated causes beliefs to cluster along different dimensions, increasing overall polarization. This can help shed light on the consequences of identity changes towards culture or nationalism.

To allow for correlation, we need to specify the joint distribution of voters' traits $(\varepsilon, \psi, \eta)$, where ε is income, ψ social progressiveness, and η is openness to trade. The general analysis of a three dimensional model with correlated traits is intricate. To starkly illustrate the main implications of this setting, we make two simplifying assumptions.

First, we assume that $(\varepsilon, \psi, \eta)$ is normally distributed around the zero mean $(0, 0, 0)$ with unitary variances and the following correlation structure.

$$\begin{pmatrix} \varepsilon \\ \psi \\ \eta \end{pmatrix} \rightsquigarrow N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix} \right]. \tag{25}$$

Thus, income is uncorrelated with the other traits, while social progressiveness and openness to trade are positively associated, $\rho > 0$. This is an admittedly extreme assumption, but it captures the idea that opposition to free trade and cultural conservativeness may reflect and underlying cultural predisposition to "closedness" (Johnston et al. 2017), or to "communal values" (Graham et al. 2009, Enke 2019) that is only mildly (or less) correlated with income.

Second, we assume that λ is sufficiently large, that identification is only driven by group contrast. As a result, all voters identify either along income, or openness, or culture, $d = \varepsilon, \eta, \psi$. This greatly simplifies the analysis.

With three dimensions, a voter can identify with three out of six possible groups. Along openness to trade, a voter is either globalist, $Gl \equiv \{\eta \,|\eta \geq \widehat{\eta}\}$, or nationalist $N \equiv \{\eta \,|\eta < \widehat{\eta}\}$, where $\widehat{\eta}$ is the threshold defining the trade groups. Along income or culture a voter is either rich or poor, or socially progressive or conservative, as in the previous section. To focus on the effect of changes in beliefs, we assume that there is an identical majority of poor, nationalists and conservatives by imposing $\widehat{\varepsilon} = \widehat{\eta} = \widehat{\psi} = z > 0$.

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, α, but also on the importance of trade protection relative to taxation, γ. Repeating the same steps as in the previous section, the appendix proves

Proposition 6 *If the relative importance of trade and culture with respect to taxes is low,* $\max(\gamma\rho+\alpha, \gamma+\alpha\rho) < 1$*, identification is class-based. If trade is important,* $\gamma > \max(\alpha, 1-\alpha\rho)$*, identification is trade-based. Finally, if culture is important,* $\alpha > \max(\gamma, 1-\gamma\rho)$*, identification is culture-based.*

When redistribution is important, society is divided into rich and poor. Economic or social change can change identities. Shocks induced by globalization increase the welfare relevance of trade, and can raise γ so that each voter perceives strong similarity with others sharing a similar exposure to trade. Then, identification becomes trade based. Likewise, a social change making cultural policy more relevant changes the dimension of identification to ψ.

Interestingly, stronger correlation ρ makes it more likely that culture or trade based identification prevails over income identification. In addition, stronger correlation ρ makes identities more sensitive to shocks to α and γ. Intuitively, highly correlated dimensions are an efficient source of identification: they reduce conflict over many dimensions, enhancing perceived similarity. Historically determined clusters of values are therefore prone to become vehicles of identification and political action, and thus of polarization, replacing class based identities. To put it differently, trade and technology shocks may have increased the relevance of pre-existing fault lines within traditional political groups defined on the left vs right dimension. Socially conservative poor voters, who traditionally identified with left wing groups despite their social conservatism, are now attracted by nationalism because it appeals to both their trade preferences and their cultural views, and viceversa for voters with opposite political features. As this happens, traditional income or class based conflict wanes and is repaced by new political cleavages over correlated dimensions.

Correlation amongst individual traits also has important implications for the consequences of identification. If income is uncorrelated with the other dimensions, class-based identification distorts beliefs about income, but creates no distortions along culture or nationalism. The nationalist-globalist divide, by contrast, does not only affect beliefs over trade policy; it also affects beliefs over cultural policy, while it dampens belief distortions about income. Cultural identification exerts a similar effect, except that it mostly distorts beliefs over civil rights.

To see this in detail, consider the case of class-based identification. Due to low or little correlation with other variables, only beliefs about income and hence redistribution are distorted. In particular, a poor and socially conservative voter who stands to lose from trade $(\varepsilon, \eta, \psi < z)$ has beliefs:

$$\begin{aligned}
\varepsilon^\theta &= \varepsilon + \theta(\varepsilon_P - \varepsilon_R) < \varepsilon \\
\eta^\theta &= \eta, \\
\psi^\theta &= \psi.
\end{aligned}$$

Suppose that now the same voter identifies with the nationalists, because the importance of trade (γ) has risen. His beliefs now change to:

$$\begin{aligned}
\varepsilon^\theta &= \varepsilon, \\
\eta^\theta &= \eta + \theta(\varepsilon_P - \varepsilon_R) < \eta \\
\psi^\theta &= \psi + \theta\rho(\varepsilon_P - \varepsilon_R) < \psi
\end{aligned}$$

The voter's beliefs about income prospects are no longer distorted, so he demands less redistribution. His views become polarized towards nationalism, however, and since η and ψ are positively correlated, also towards social conservatism (note that distortions are still pinned down by $\varepsilon_P - \varepsilon_R$ since by assumption gaps in different dimensions are identical). In the overall political equilibrium, then, trade policy becomes distorted, and so does civil rights policy. The analysis of all cases yields the following result.

Proposition 7 *If the relative importance of trade and culture with respect to taxes is low,* $\max(\gamma\rho + \alpha, \gamma + \alpha\rho) < 1$*, stronger stereotypes* θ *enhance polarization only in redistribution. If economic or social change makes trade and/or culture more important, so that income identification is abandoned, then polarization over redistribution drops, but stronger stereotypes* θ *enhance polarization over both trade policy and cultural policy.*

This example shows that, when a shock changes identification to a dimension along which several preferences are correlated, voters cluster into groups that disagree over many policies. This enhances polarization and conflict among different "views of the world".

This mechanism can 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" versus "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 these predictions in more detail. We consider individuals who were recently exposed to a large trade shock, and investigate how their beliefs over redistribution, globalization, and culture are affected.

6 Some Evidence

This section explores the empirical validity of some of the theoretical predictions outlined above, using survey data. In subsections 6.1 and 6.2 we study the consequences of switching identities on individual attitudes. In subsection 6.3 we study how increased group contrast influences individual beliefs. Our aim is not to estimate precise causal effect, but to show that the main predictions of the theory are consistent with observed correlations in the data.

6.1 Switching Identities

Suppose that identification switches from the traditional left vs right dimension to a nationalist vs globalist dimension. According to the theory, we should observe a dampening of redistributive conflict, and an exasperation of conflicts over globalization or immigration. In this subsection we explore the validity of this prediction, exploiting data from the French Dynamob survey.

France is an ideal testing ground, because there was a clear shift in the dimensions of political conflicts between 2013 and 2017. This is illustrated vividly in Figure 4. The vertical axis measures attitudes towards immigration, globalization and European integration (higher values correspond to more open attitudes), the horizontal axis attitudes on the role of the government in regulating the economy and protecting workers (higher values correspond to more right wing attitudes). These measures are constructed by extracting the first principal component from questions on each of these two dimensions of political conflict, and then estimating the residuals of each principal component after conditioning on income and education. Each dot corresponds to an individual. The tones indicate how respondents were split between two clusters estimated from the original questions: 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 the role of the state in the domestic economy. In 2017, the cleavage concerned attitudes towards globalization and immigration.

Figure 4

The criterion for assigning observations to a cluster is to mininimize within group variance, in the space of seven questions from which the two principal components were extracted.[15] This clustering excercise thus tells us that between 2013 and 2017 individual views became more distant between opponents and supporters of international opennes, and more similar in the traditional left vs right dimension of political conflict. According to our theory, this should be accompanied by a switch in the dimension of identification. Identities are not observed, but they can be revealed by how people voted. Indeed, the two clusters track closely how the respondent voted in the first and second rounds of presidential elections. Figure 5 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

[15] Clusters are defined by applying Ward's minimum variance method on the seven questions used for the two principal components, and stopping the hierarchical partitioning algorithm at 2 clusters. The algorithm is run after making the input variables dichotomic, to reconcile inital scale differences, and after conditoning on income and education. The appendix provides more details.

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.[16]

Figure 5

The theory 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 in favor of globalization and less opposed to state intervention in the economy. Likewise, someone who abandoned his left wing identity and became a nationalist should become more opposed to globalization and less in favor of regulating the economy. These predictions can be tested using the panel dimension of the Dynamob survey (a subset of about 450 individuals were interviewed both in 2013 and 2017, and were asked how they voted in the previous presidential election). Since identities are not observed, we assume that they are revealed by how individuals voted in the *first* round of the presidential elections. Specifcally, we assume that those who voted left or right in the first round identified along a left-right dimension, and belonged to the left (L) and right (R) wing group respectively. Similarly, those who cast their first round vote for Le Pen or Macron identified on the nationalist vs globalist dimension, and belonged to the nationalist (N) and globalist (Gl) group respectively. Voters who in the first 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.[17]

To test these predictions, we estimate two regressions on individual panel data. In the first regression, the dependent variable is the change in attitudes over globalization and immigration between 2013 and 2017, Δy_i^{NGl}, measured by the first principal component whose residuals are plotted on the vertical axis of Figure 4 (higher values denoting a more open policy). Here we estimate:

$$\Delta y_i^{NGl} = \beta_{Gl} Gl_i + \beta_N N_i + \alpha y_{i0}^{NGl} + X_i'\delta + FEs + u_i$$

where Gl (N) is a dummy variable for whether in 2017 individual i voted globalist (nationalist), y_{i0}^{NG} denotes his initial attitudes in 2013, X is a vector of individual covariates measured in the initial period (years of schooling, income, dummy variables for gender, age, immigrant

[16]Note that the clusters depicted in Figure 4 were identified without exploiting the voting information. The vote shares in Figure 5 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.

[17]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, implictly 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.

status, working status), u is the error term and FEs are dummy variables that have to be included to estimate the effect of identity changes, specifically a dummy variable for voting N both in 2012 and 2017, and for voting N in 2012. Thus, the coefficients β_{Gl} and β_N refer to those who voted Gl and N *for the first time* in 2017 respectively (i.e., under our assumption who acquired a new globalist 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.

In the second regression, the dependent variable is the change in attitudes over the role of the state in regulating the economy and protecting workers, Δy_i^{LR}, measured again by the first principal component plotted on the horizontal axis of Figure 4, with higher values indicating a more right wing policy. Here we estimate:

$$\Delta y_i^{LR} = \beta_{LGl} LGl_i + \beta_{RGl} RGl_i + \beta_{LN} LN_i + \beta_{RN} RN_i + \alpha y_{i0}^{LR} + X_i'\delta + FEs + u_i$$

where again y_{i0}^{LR} denotes initial attitude in 2013, LGl (RGl) is a dummy variable that equals one if the individual changed his identity from left (right) to globalist, LN (RN) is a dummy variable capturing changes in identity from left (right) to nationalist, X is the same vector of individual covariates as above, and FEs are dummy variables for how the individual voted in 2012 and 2017, that have to be included so that the β coefficients refer exclusively to those who switched identity from the left / right to the nationalist / globalist dimension (specifically, we include dummy variables for voting L and R in 2012, and for voting Gl and N in 2017). Thus, the coefficients of interest have the following expected signs: $\beta_{LGl} > 0 > \beta_{RGl}$ and $\beta_{LN} > 0 > \beta_{RN}$.[18]

Table 1 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, while those who turned to Macron in the first round of 2017 became more in favor, although here the estimated coefficient is only significant in the more pasimonious specification. 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 26% 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. Columns 3 and 4 refer to changes in attitudes on the role of the state in the economy. Here the estimates of interest are reported in raws 4-8. Individuals who switched from the right to Macron dampened their opposition to state intervention, as expected ($0 > \beta_{RGl}$), although the switchers from left to Macron also have a negative but insignificant estimated coefficient, contrary to our predictions ($\beta_{LGl} > 0$). Among voters who switched to Le Pen, those coming from the left increased their acceptance

[18]Note that here the predictions only concern those who previously voted left or right and then switch to the nationalist / globalist identity, while the theory has no implications for those who were not identified on the left / right dimension in 2012. This is why the specification for Δy_i^{LR} differs from that for Δy_i^{NGl}.

of a small role for the state in the domestic economy, in line with the prediction ($\beta_{LN} > 0$), while the effect of the switch from the right is not statistically significant. The point estimates imply that a switch from left in 2012 to Le Pen in 2017 is associated with a change in their attitudes towards the domestic economy by about 70% of the initial attitudes of all those who voted left in 2012, relative to those who voted left in 2012 and did not switch. The effect of changing identification from right to Macron, evaluated in a similar manner, implies a change of 20% of the mean initial attitude amongst those who voted right in 2012 and did not switch. All in all, therefore, the evidence from vote switchers in the first round of French presidential elections supports several (but not all) of the 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. We now turn to another setting, where the switch in identities reflects an exogenous increase in the relevance of the protectionist vs globalist cleavage.

6.2 Exposure to Import Competition

In this subsection we test the following reduced form implication of the theory. Suppose that imports exposure rises. This increases the welfare relevance of cleavages along the nationalist / globalist dimension (the parameter γ in the trade model of section 5 rises - cf. (24)). Then, because more exposed individuals are more likely to identify with the nationalists, they should demand more trade protection and less redistributive taxation. Moreover, if attitudes over immigrants and civil rights are positively correlated with those on globalization, the new nationalists also turn anti-immigrants and socially conservative. In this subsection we test this prediction. Even though we do not observe identities, we ask how increased exposure to imports from China affected individual attitudes on a variety of issues. Relative to the previous subsection, we gain the possibility of estimating an effect associated with exogenous changes in the environment, although the mechanism of identity changes is left implicit, and other mechanisms may act as confounds.

We follow the same strategy as Autor et al. (2017), who found that larger imports from China increased political polarization in the US and made voters more conservative. But rather than focusing on polarization and an aggregate measures of conservative ideology, as they do, we study the effect on specific attitudes. We exploit the CCES survey, a large scale survey conducted over the internet between 2006 and 2016. On average about 36 000 individuals were interviewed every year. 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).[19]

We focus on four survey 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.[20]

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. We measure the change in import exposure as in Autor et al. (2017), based on the change in imports from China. US imports from China grew particularly fast before the start of our sample period. In the repeated cross section we thus measure 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 import exposure on identity is not instantaneous, but partly operates with a lag - a very plausible assumption.[21]

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, ΔIP_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

[19]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 surevy data, obtained by aggreating 10 questions in the Pew (2014) survey; the Pew survey lacks the panel dimension and has about 25 respondents per CZ.

[20]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.

[21]Autor et al. (2013) measure the change in import exposure in each 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:

$$\Delta IP_c = \sum_{m \in M} \frac{L_{m,c,00}}{L_{c,00}} \times \frac{I_{m,t_2} - I_{m,00}}{Y_{m,91} + I_{m,91} - X_{m,91}} \tag{26}$$

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 Author et al (2013). The instrument is obtained by replacing $(I_{m,t_2} - I_{m,00})$ with $(I^{EU}_{m,t_2} - I^{EU}_{m,00})$, 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.

individual covariates (gender, race, educational attainements, age and age squared), Z_c is a vector of covariates referring to the CZ in the year 2000.[22] The inclusion of CZ fixed effects (α_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) and as explained in footonote 23 above, ΔIP_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:

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

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 ΔIP_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 ΔIP_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 excercise.

Table 2 reports the estimates for two specifications of each equation: with and without the CZ covariates Z_c. Estimation is by 2SLS and standard errors are clustered at the CZ level. Both in the repeated cross sections and in the panel estimates, an exogenous rise in imports exposure is associated with an increased willingness to accept cuts in domestic public spending, more aversion to immigrants, a stronger importance of abortion (available only in the cross sections), and an increase likelihood to vote Republican (no effect is found on votes for Democrats).

According to the 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

[22] 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 an a dummy for Republican victory in that county, all variables meaured 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.

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.

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 found in their paper. Residents of CZ more exposed to an acceleration of imports from China became less demanding of 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 if nationalism is positively correlated with social conservativeness.

6.3 Individual Beliefs and Group Contrast

According to the theory, beliefs and policy preferences are systematically distorted by group contrast: when opposing groups are more different, beliefs are more polarized. For instance, in subsection 3.1 bliss points on taxes take the form $\tau^{\varepsilon\theta} = T[\varepsilon + \theta(\varepsilon_G - \varepsilon_{\bar{G}})]$, where $T(.)$ is a linear function of the individual's relative income ε and of the income contrast between groups, $(\varepsilon_G - \varepsilon_{\bar{G}})$.

To explore the validity of this prediction, in this subsection we thus estimate an equation of the following general form:

$$y_{it} = X_{it}'\beta + (X_{it}^G - X_{it}^{\bar{G}\prime})'\beta\theta + u_{it} \tag{27}$$

where y_{it} denotes the attitudes of individual i in period t, on three different issues: redistribution, immigration and abortion; X is a vector of determinants of his rational preferences, such as income, education, gender, age, ethnicity etc.; X_{it}^G and $X_{it}^{\bar{G}}$ are the group averages of these same features, G being the group with which i identifies and $\bar{G}$ being the opposing group; and u is an unobserved error term. We want to test whether $\theta > 0$, as predicted by the theory. In the termnology of Manski (1993), this amounts to testing whether beliefs are influenced by exogenous contextual factors associated with group contrast. Unfortunately, we don't have an exogenous source of variation in group composition, so we can only estimate correlations.

We exploit US data taken from the ANES surveys during the period 1998-2016 (every four years). The relevant groups of political identification are taken to be the supporters of the Democratic and Republican parties, denoted by $G = D, R$ respectively. Since the US is a two-party system, this assumption is not overly restrictive, although it forces us to neglect important social and ethnic differences within each party.

Given the large number of potentially relevant variables in the vector X, we estimate the parameter of interest θ by a two-step procedure. First, for each policy issue we estimate $\hat{\beta}$ by OLS in the full sample, having restricted $\theta = 0$. This allows us to compute predicted group contrast, defined as $\hat{y}_t^R - \hat{y}_t^D = [X_t^R - X_t^D]'\hat{\beta}$. The vector X_{it} includes wave and macro-region fixed effects and the respondent gender, education, age, income, race, marital status, employment status, labor union membership, religion, and relevant interactions between these characteristics (see the appendix for more detail). Thus, group contrast is measured as a weighted average of differences in observed features of members of the opposite groups, with weights given by their relevance in explaining individual opinions in the full sample.

In the second stage, we estimate:

$$y_{it} = X_{it}'\delta + \theta_1 d_{it}^D * (\hat{y}_t^D - \hat{y}_t^R) + \theta_2 d_{it}^R * (\hat{y}_t^R - \hat{y}_t^D) + \alpha_1 d_{it}^D + \alpha_2 d_{it}^R + u_{it}$$

where d_{it}^G is a dummy variable that equals 1 if i identifies with $G = D, R$. This specification allows for some individuals not to be politically identified (about 38.4% of respondents do not identify with either the Democratic or Republican party). We also let the coefficient of interest θ differ between Republicans and Democrats. Recall that wave and region fixed effects are always included in X. Hence, the variation in the variable of interest comes from within wave and region variation between politically identified and non-identified individuals, and from allowing time variation in $\hat{y}_t^D - \hat{y}_t^R$ to have different effects on the attitudes of Democrats and Republicans. Since we control for group identification through the dummy variables d_{it}^G, testing for $\theta_1, \theta_2 > 0$ amounts to asking whether politically identified individuals have more extreme positions in the years in which group contrast is larger.[23]

As already noted, the contextual effects $X_{it}^G - X_{it}^{\bar{G}}$ are endogenous through group composition. In particular, time variation in $X_t^D - X_t^R$ could be correlated with omitted variables that impact differently on Democrats and Republicans. For instance, it is possible that over time the Republican and the Democratic parties have taken more extreme policy positions, and this induced moderate voters to become independent, or attracted more extreme voters. To the extent that this endogenous shift in the composition of politically identified voters is associated with unobservable individual characteristics, it could account for more extreme preferences amongst the remaining party supporters. Since we cannot rule this out, despite the fact that we control for a large number of individual features and there interaction, a positive estimate of θ cannot be interpreted as a causal effect of group contrast on individual

[23] Note that $\hat{y}_t^D - \hat{y}_t^R$ is a generated regressor. Nevertheless, under the null that $\theta = 0$ the estimated standard errors are unbiased, so that the t-statistic on θ is still a valid Lagrange multiplier test of the hypothesis that $\theta = 0$. (Pagan 1984) Given the small number of waves, we do not cluster the standard errors by political groups, but the clustered standard errors are very similar to those reported below.

beliefs. Nevertheless, it is worth asking whether the estimated correlations are consistent with the predictions of the theory.

The results are displayed in Table 3. Each column refers to a different policy dimension: willingness to redistribute, immigration policy and attitudes towards abortion. Higher values correspond to more willingness to redistribute and more liberal views on immigration and abortion. The estimates of θ always have the expected positive sign, and they are almost always highly significant. Thus, as predicted, a larger group contrast is associated with more extreme positions of politically identified individuals. The correlations are non trivial in magnitude. The estimated coefficient of 2.105 in column 1 implies that the observed change in $\hat{y}_t^R - \hat{y}_t^D$ between 1998 and 2016 is associated with a 17% reduction in the willingness to redistribute of Republican supporters, relative to the value of y_t^R in 1998. The effect on the Democrats is not significantly different from 0. For immigration the same excercise implies a change of about 20% of attitudes towards immigrants (in opposite directions) for both Democrats and Republicans, relative to their initial average attitudes. On abortion, the effect is 6% for Democrats and 12% for Republicans (in opposite directions), again relative to initial average attitudes.

7 Concluding Remarks

Identity theory provides a rich theoretical framework, that can be used to study new and puzzling political phenomena. We have illustrated how this framework 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 important implications of this general approach, but much more remains to be done. Future work should try to characterize the structure of some of these latent grouping. For instance, we argued that income based groups - while highly polarized along redistribution - may be less dispositive across other political demands relative to, say, culture based groups. This aspect could be investigated using survey or other data. It would also be interesting to analyze whether economic or social change influence identities and political demands in the way predicted by these latent clusters of interests.

A second issue concerns the role of the media. The media is 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. Moreover, 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 are not only shaped by the media. They also or perhaps especially form in spontaneous interactions with neighbors or friends sharing similar problems. Exploring the mechanisms for the diffusion of identities is another important topic for future empirical research.

In this paper we have focused on voters' behavior, since in our model opportunistic politicians simply adapt to voters' demands. But identity politics also matters on the supply side. A first 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 individual features of politicians make them more appealing to voters? Social psychology suggests that stereotypical types exert more influence. This implies that, in times of strong polarization, the most successful politician is in the tails, not in the middle. Also, because the salient aspects of politicians are related to the salient cleavages, competition or shortcomings along less salient aspects are not really important to voters. 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 the bridge between social and political identification, and the evolution of party systems. Our theory studies how individuals identify with social groups. In what circumstances does a political party primarily represent a single group, and when instead are more social groups represented by the same party? And when this happens, which groups more naturally coexist in the same party? 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 identifcation. Why is this so, and could it imply that, under some assumptions, policies are extreme in both dimensions? And why has the composition of party supporters changed significantly in the post war period in many advanced democracies, with more educated and urban workers abadoning the political right in favor of the left, as emphasized by Piketty (2018)? Can this be explained as a changing dimension of prevalent identification, with cultural and civil rights issues becoming a more salient dimension than class based conflict?

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

8 References

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.

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, 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, 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.

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).

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne (2018), "Economic Losers and Political Winners: Sweden's Radical Right", *mimeo.*

Delli Carpini, Michael X., and Scott Keeter (1993), "Measuring Political Knowledge: Putting First Things First", *American Journal of Political Science*, Vol. 37, No. 4, 1179-1206.

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.

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

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.

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

Gentzkow, Matthew (2017), "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.

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.

Graham, Jesse, Brian Nosek, and Jonathan Haidt (2012), "The Moral Stereotypes of Liberals and Conservatives: Exaggeration of Differences across the Political Spectrum," *PLoS One*, 7, 1–13.

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.*

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.

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

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

Lupia, Arthur, and Mathew D. McCubbins (1998), *The Democratic Dilemma. Can Citizens Learn What They Need to Know?*, Cambridge University Press.

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

Manski, Charles (1993), "Identification of Endogenous Social Effects: The Reflection Problem", *The Review of Economic Studies*, Vol. 60, 3, pp. 531-542.

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.

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, and John C.Turner (1986), "The social identity theory of inter-group behavior", In S. Worchel & W. G. Austin (Eds.), *Psychology of Intergroup Relations*, Chicago:

Nelson-Hall.

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), R*ediscovering the Social Group: A Self Categorization Theory*, New York, Blackwell.

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

Figure 1: Primitive Groups Defined on Income and Culture

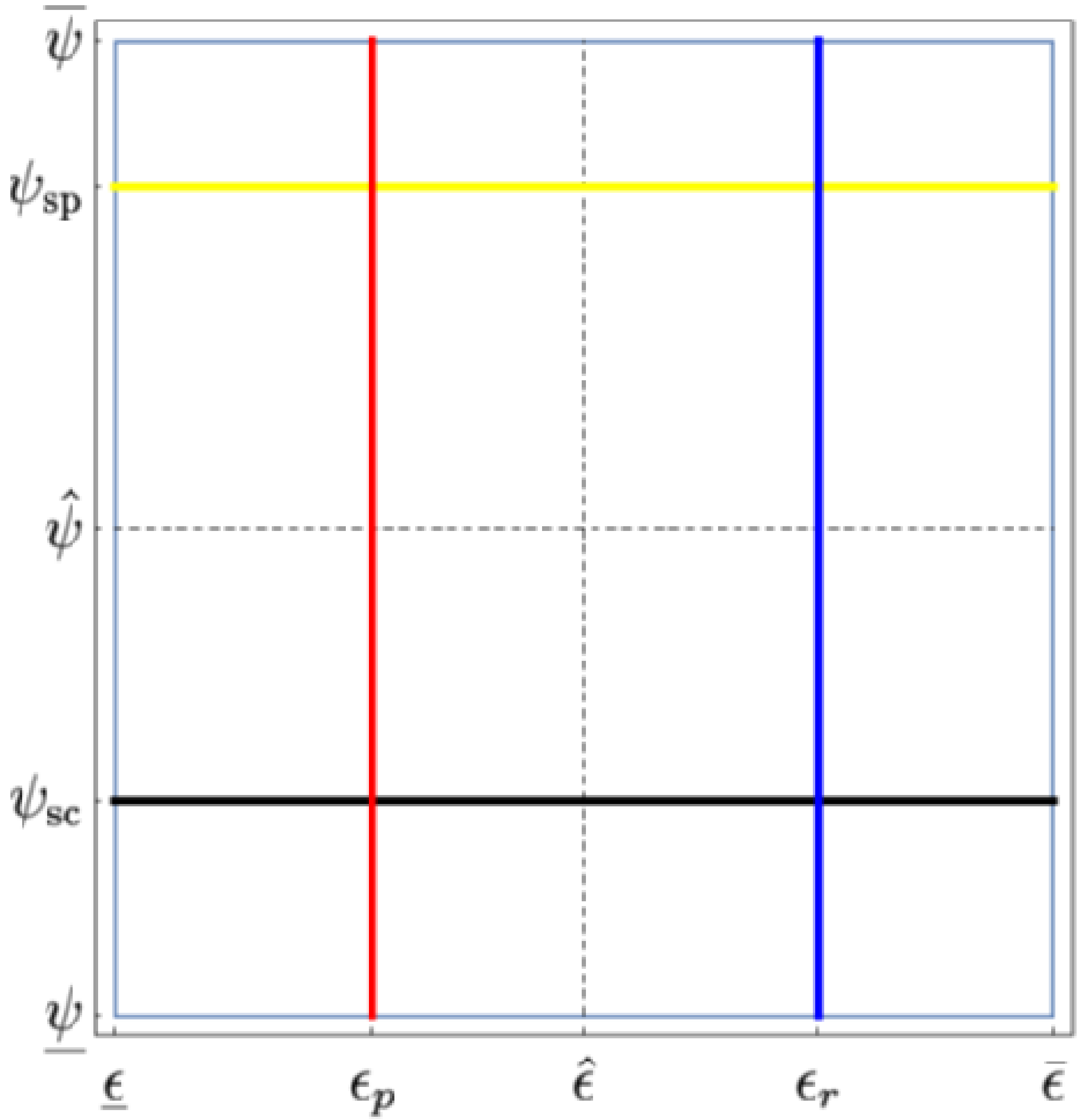
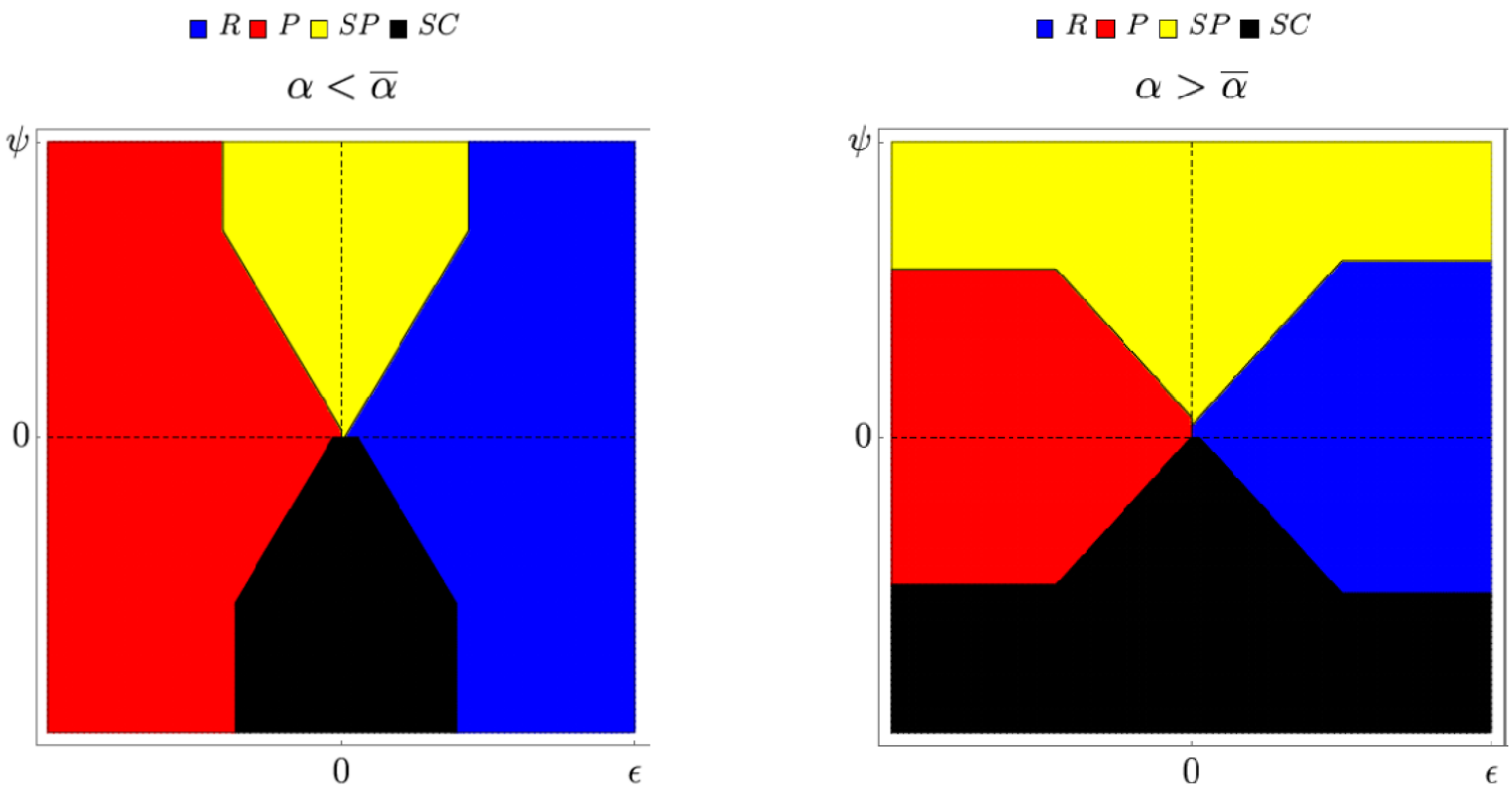


Figure 2: Mixed Identification



(a) Prevalent income-based indentification

(b) Prevalent culture-based indentification

Figure 3: Primitive Income Groups

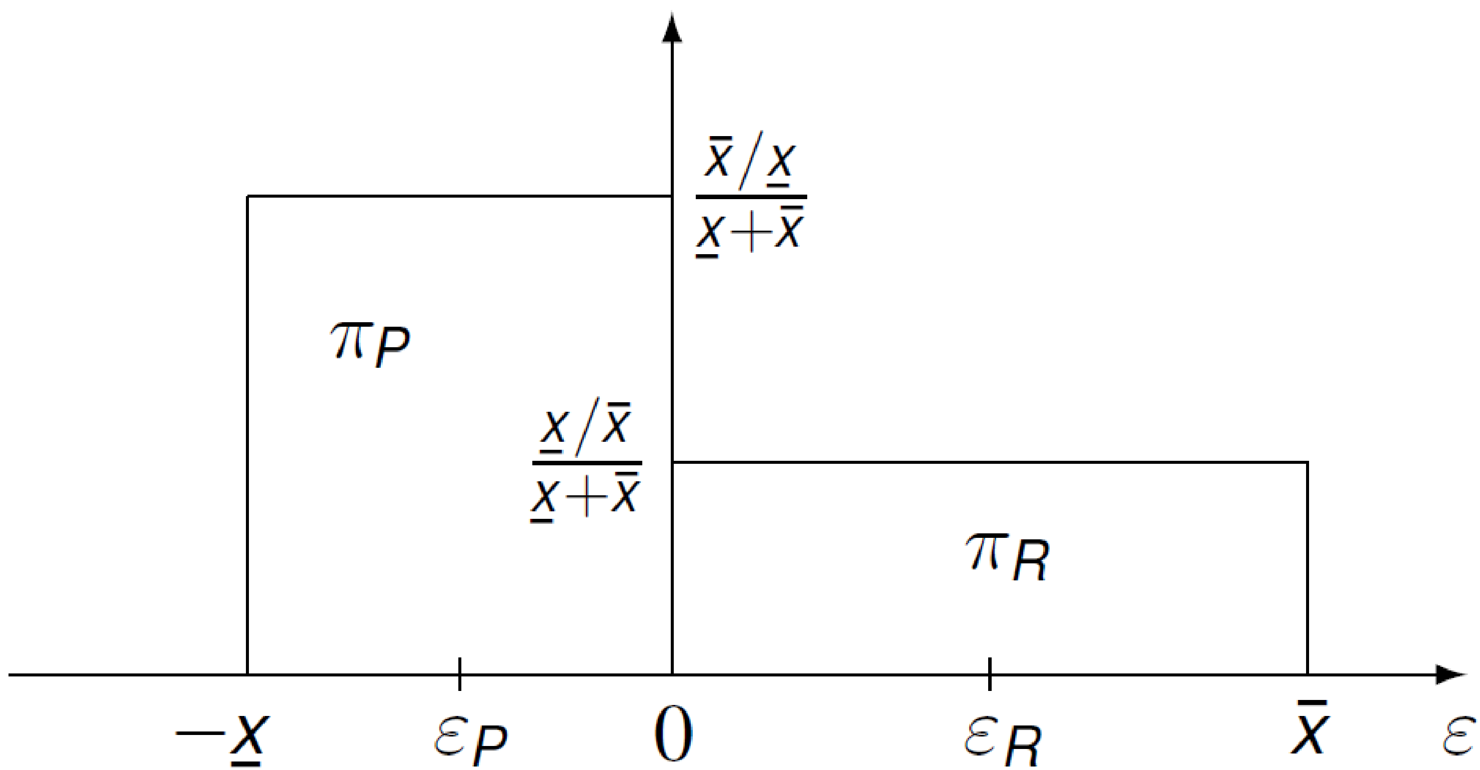
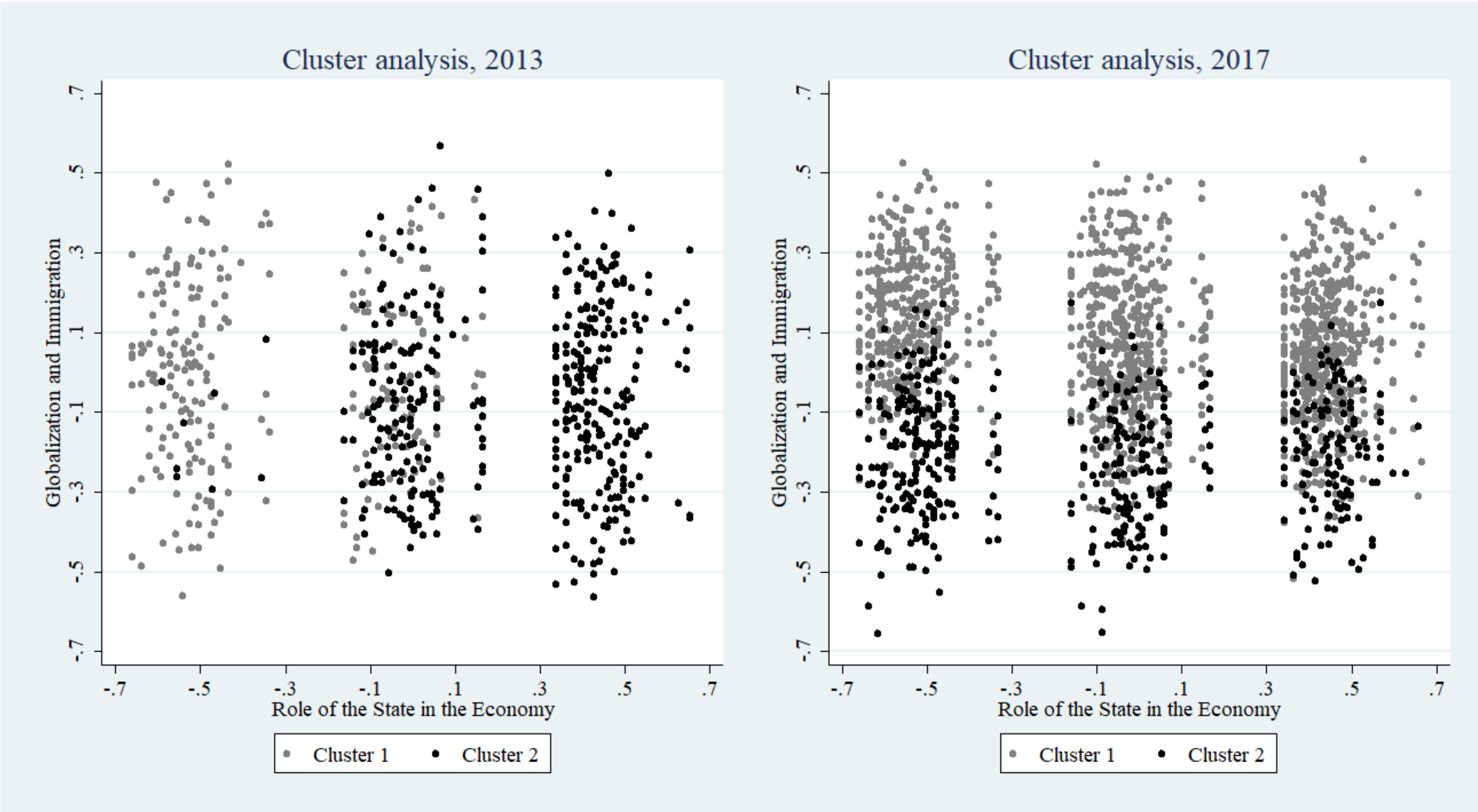
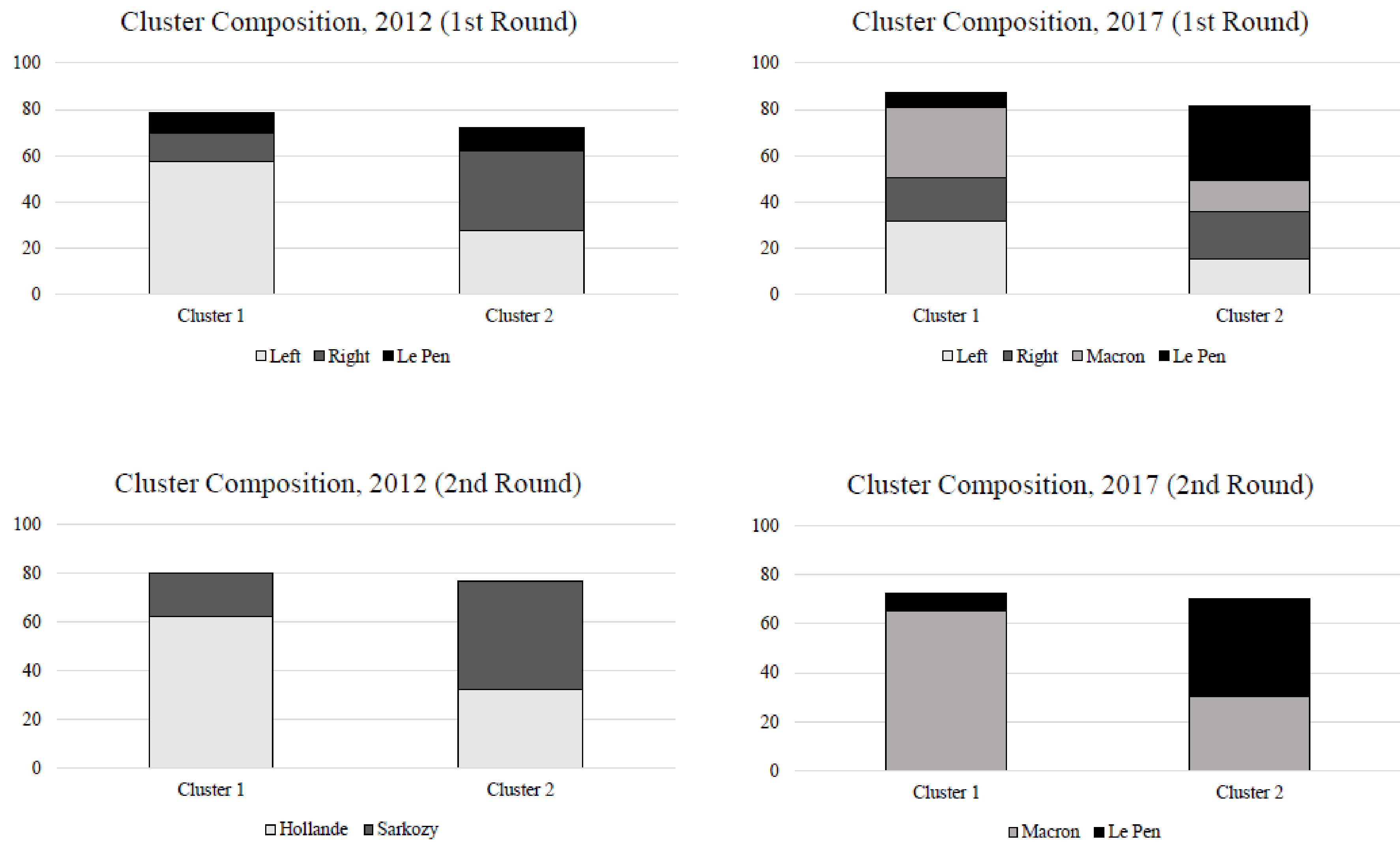


Figure 4: Changing Dimension of Political Conflict: France in 2013 and 2017



Note: 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 protecting workers (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, obtained applying Ward's method on the full set of questions on which the principal components are based. The clustering algorithm is run separately for each year. See the appendix for the list of questions, the data sources and more information on how the variables were treated before the analysis.

Figure 5: Composition of Votes in the Two Clusters



Note: 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 4. 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.

Table 1: Switching Identities and Attitudes

	Δ Globalization and Immigration		Δ Role of the State in the Economy	
Macron 17	0.0318**	0.0156	0.159***	0.141***
	(0.0147)	(0.0151)	(0.0359)	(0.0396)
Le Pen 17	-0.127***	-0.125***	0.0183	0.00776
	(0.0237)	(0.0247)	(0.0304)	(0.0345)
Macron 17 ∗ Right 12			-0.169***	-0.139**
			(0.0496)	(0.0536)
Macron 17 ∗ Left 12			-0.0779*	-0.0597
			(0.0457)	(0.0496)
Le Pen 17 ∗ Right 12			-0.0539	-0.0257
			(0.0551)	(0.0587)
Le Pen 17 ∗ Left 12			0.230***	0.227***
			(0.0714)	(0.0752)
Le Pen 17 ∗ Le Pen 12	0.0164	0.0109		
	(0.0456)	(0.0483)		
Le Pen 12	-0.0402	-0.0505		
	(0.0327)	(0.0345)		
Right 12			0.117***	0.106***
			(0.0254)	(0.0282)
Left 12			-0.0580**	-0.0716**
			(0.0263)	(0.0305)
Individual Controls	NO	YES	NO	YES
Observations	469	427	470	428
R-squared	0.294	0.333	0.214	0.231

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust 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. Individual controls are education, income and dummy variables for gender, age, immigrant status and work status. Estimation is by OLS.

Table 2: Exposure to Imports from China and Attitudes

	Cross Section		Panel	
	A.	Cut Domestic Spending (0-100)		
Δ CZ Exposure ∗ Post	0.512	1.086**		
	(0.363)	(0.525)		
Δ CZ Exposure			0.703	3.503**
			(0.938)	(1.771)
Observations	72,712	72,712	8,296	8,296
F	67.1	27.34	80.31	42.7
	B.	Migrant Aversion (PC)		
Δ CZ Exposure ∗ Post	0.010**	0.014		
	(0.004)	(0.010)		
Δ CZ Exposure			0.076***	0.120**
			(0.023)	(0.054)
Observations	73,484	73,484	9,451	9,451
F	75.17	31.01	65.11	42.1
	C.	Abortion Importance		
Δ CZ Exposure ∗ Post	0.013***	0.019**		
	(0.005)	(0.009)		
Observations	48,871	48,871		
F	72.99	27.29		
	D.	Republican Vote		
Δ CZ Exposure ∗ Post	0.0146*	0.0197*		
	(0.008)	(0.008)		
Δ CZ Exposure			0.054***	0.097***
			(0.0181)	(0.0278)
Observations	77,558	77,558	6,673	6,673
F	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 presidental 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 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 Repubblican victory in that county. Standard errors are clustered at CZ level. Estimation is by 2SLS.

Table 3: Between-Group Contrast and Individual Attitudes

	y = Redistribution	y = Immigration	y = Abortion
$d^D * (\hat{y}^d - \hat{y}^r)$	0.422	1.465***	0.834**
	(0.823)	(0.438)	(0.414)
$d^R * (\hat{y}^r - \hat{y}^d)$	2.105**	1.587***	1.336***
	(0.868)	(0.469)	(0.444)
d^D	0.0666	0.00657	0.0397**
	(0.0640)	(0.0120)	(0.0172)
d^R	0.0375	0.00348	-0.0729***
	(0.0675)	(0.0131)	(0.0181)
Observations	9,619	11,301	11,165
R-squared	0.178	0.109	0.170

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors in parentheses. The dependent variables measure attitudes towards redistribution, immigrants and abortion. Higher values indicate a preference for more redistribution and more liberal views on immigration and abortion. All first stage predictors are included as regressors in the second stage. Such regressors include year and region dummies and a set of demographic characteristics consisting of gender, a quadratic of age, education, income, race, marital status, employment status, labor union membership, religion, as well as relevant interactions between these characteristics. Estimation is by OLS.

9 Appendix

9.1 Theory

Proof of Lemma 1. Equation (5) implies:

$$\int \widetilde{\varepsilon} f^{\theta}\left(\widetilde{\varepsilon}\left|\varepsilon, G\right.\right) d\widetilde{\varepsilon} = \int \widetilde{\varepsilon}\left[\frac{f\left(\widetilde{\varepsilon}\left|\varepsilon_G\right.\right)}{f\left(\widetilde{\varepsilon}\left|\varepsilon_{\overline{G}}\right.\right)}\right]^{\theta} Z f\left(\widetilde{\varepsilon}\left|\varepsilon\right.\right) d\widetilde{\varepsilon}$$

$$= \mathbb{E}\left(\widetilde{\varepsilon}\left|\varepsilon\right.\right)\mathbb{E}\left\{\left[\frac{f\left(\widetilde{\varepsilon}\left|\varepsilon_G\right.\right)}{f\left(\widetilde{\varepsilon}\left|\varepsilon_{\overline{G}}\right.\right)}\right]^{\theta} Z\left|\varepsilon\right.\right\} + Z \cdot cov\left(\widetilde{\varepsilon}, \left[\frac{f\left(\widetilde{\varepsilon}\left|\varepsilon_G\right.\right)}{f\left(\widetilde{\varepsilon}\left|\varepsilon_{\overline{G}}\right.\right)}\right]^{\theta}\right),$$

where $\mathbb{E}\left\{\left[\frac{f\left(\widetilde{\varepsilon}\left|\varepsilon_G\right.\right)}{f\left(\widetilde{\varepsilon}\left|\varepsilon_{\overline{G}}\right.\right)}\right]^{\theta} Z\left|\varepsilon\right.\right\} = 1$ because constant Z is set so that $f^{\theta}\left(\widetilde{\varepsilon}\left|\varepsilon, G\right.\right)$ integrates to one. With normal densities, it is immediate to see that $f^{\theta}\left(\widetilde{\varepsilon}\left|\varepsilon, G\right.\right)$ is normal with variance $\sigma^2_{\widetilde{\varepsilon}}$ and mean $\varepsilon + \theta\left(\varepsilon_G - \varepsilon_{\overline{G}}\right)$. ■

Proof of Proposition 1. By inspection of Equation (11). ■

Approximate relative distance function. Here we derive the piecewise linear approximation of $\Delta_d^{\varepsilon\psi}$ in (14) from equation (13). Recall that $\tau^{G_d} = T(\varepsilon_{G_d})$ and $q^{G_d} = \psi^{G_d}$, so that $W^{\varepsilon\psi}\left(\tau^{G_d}, q^{G_d}\right) = W^{\varepsilon\psi}\left(T(\varepsilon_{G_d}), \psi_{G_d}\right)$. Taking a second order approximation of $W^{\varepsilon\psi}\left(\tau^{G_d}, q^{G_d}\right)$ with respect to ε_{G_d} and ψ_{G_d} at the point $\varepsilon_{G_d} = \varepsilon$ and $\psi_{G_d} = \psi$, and recalling that by (12) $W^{\varepsilon\psi}_{\tau q} = 0$ and that the optimality conditions imply $W^{\varepsilon\psi}_{\tau}\left(\tau^{\varepsilon}, q^{\psi}\right) = W^{\varepsilon\psi}_{q}\left(\tau^{\varepsilon}, q^{\psi}\right) = 0$, we have:

$$W^{\varepsilon\psi}\left(\tau^{G_d}, q^{G_d}\right) \simeq W^{\varepsilon\psi}\left(\tau^{\varepsilon}, q^{\psi}\right) + W^{\varepsilon\psi}_{\tau\tau}\left(\tau^{\varepsilon}, q^{\psi}\right)(T_{\varepsilon})^2(\varepsilon_{G_d} - \varepsilon)^2 + W^{\varepsilon\psi}_{qq}\left(\tau^{\varepsilon}, q^{\psi}\right)(Q_{\psi})^2(\psi_{G_d} - \psi)^2$$

where $T_{\varepsilon} = -1/\varphi v < 0$ and $Q_{\psi} = 1$. Taking a second order approximation of $W^{G_d}\left(\tau^{\bar{G}_d}, q^{\bar{G}_d}\right)$ with respect to $\varepsilon_{\bar{G}_d}$ and $\psi_{\bar{G}_d}$ at the point $\varepsilon_{\bar{G}_d} = \varepsilon_{G_d}$ and $\psi_{\bar{G}_d} = \psi_{G_d}$, we have:

$$\begin{aligned}\Delta_d^{\varepsilon\psi} \simeq\ & -[W^{\varepsilon\psi}_{\tau\tau}\left(\tau^{\varepsilon}, q^{\psi}\right)(T_{\varepsilon})^2(\varepsilon_{G_d} - \varepsilon)^2 + W^{\varepsilon\psi}_{qq}\left(\tau^{\varepsilon}, q^{\psi}\right)(\psi_{G_d} - \psi)^2] + \\ & +\lambda[W^{G_d}_{\tau\tau}\left(\tau^{G_d}, q^{G_d}\right)(T_{\varepsilon_{G_d}})^2(\varepsilon_{\bar{G}_d} - \varepsilon_{G_d})^2 + W^{G_d}_{qq}\left(\tau^{G_d}, q^{G_d}\right)(Q_{\psi_{G_d}})^2(\psi_{\bar{G}_d} - \psi_{G_d})^2]\end{aligned}$$

Given the quadratic functional form of the utility function in (12), we have: $W^{\varepsilon\psi}_{\tau\tau} = -\nu\varphi$ and $W^{\varepsilon\psi}_{qq} = -\kappa$. Thus, $\Delta_d^{\varepsilon\psi}$ can be rewritten as:

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

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 (14) in the text. ■

Finally, in the trade policy model of section 5 we have: $W^{\varepsilon\psi}_{\tau\tau} = -\varphi\nu$ and $W^{\varepsilon\psi}_{qq} = -(2\nu - 1)(p^*)^2/\delta$, and $T_\varepsilon = -1/\varphi\nu < 0$, $Q_\psi = \sigma\delta/(2\nu-1)p^*$. Inserting these expressions in the approximation above, we get: $\alpha = \sigma^2\varphi\delta/(2 - 1/\nu)$.

Proof of Proposition 2. We can summarize the two dimensional groups by the following parameters: ε_R, $\varepsilon_P = -\chi_\varepsilon\varepsilon_R$, ψ_{SP}, $\psi_{SC} = -\chi_\psi\psi_{SP}$, with $\chi_\varepsilon, \chi_\psi < 1$. Indeed, the mean zero assumption implies that $\pi_R\varepsilon_R + \pi_P\varepsilon_P = 0$ and $\pi_{SC}\psi_{SC} + \pi_{SP}\psi_{SP} = 0$, where π_G denotes the population shares of different groups. This implies that $\varepsilon_P = -\frac{\pi_R}{\pi_P}\varepsilon_R$ and $\psi_{SC} = -\frac{\pi_{SP}}{\pi_{SC}}\psi_{SP}$, where $\chi_\varepsilon = \frac{\pi_R}{\pi_P}$ 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 (ε, ψ) space. ■

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

$$\begin{aligned}
\Delta^{\varepsilon\psi}_{\widetilde{\varepsilon}} \;\le\; \Delta^{\varepsilon\psi}_{\widetilde{\psi}} &\Leftrightarrow |\varepsilon - \varepsilon_R| + \alpha\,|\psi| - \lambda\varepsilon_R(1+\chi_\varepsilon) \le |\varepsilon| + \alpha\,|\psi - \psi_{SP}| - \lambda\alpha\psi_{SP}(1+\chi_\psi), \\
&|\varepsilon - \varepsilon_R| - \alpha\,|\psi - \psi_{SP}| + \alpha\psi - \varepsilon \le \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right] \equiv \lambda MC.
\end{aligned}$$

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

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

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

$$\begin{aligned}
\Delta^{\varepsilon\psi}_{\widetilde{\varepsilon}} \;\le\; \Delta^{\varepsilon\psi}_{\widetilde{\psi}} &\Leftrightarrow -\varepsilon_R + \alpha\psi_{SP} \le \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\
&\Leftrightarrow \;\; \alpha \le \alpha_1 \equiv \frac{\varepsilon_R}{\psi_{SP}}\frac{\lambda(1+\chi_\varepsilon)+1}{\lambda(1+\chi_\psi)+1}.
\end{aligned} \tag{28}$$

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

$$\begin{aligned}
\Delta^{\varepsilon\psi}_{\widetilde{\varepsilon}} \;\le\; \Delta^{\varepsilon\psi}_{\widetilde{\psi}} &\Leftrightarrow -\varepsilon_R - \alpha\psi_{SP} + 2\alpha\psi \le \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\
&\Leftrightarrow \;\; \psi \le \psi_1 \equiv \psi_{SP}\left[1 + \frac{\lambda(1+\chi_\psi)+1}{2}\frac{\alpha_1 - \alpha}{\alpha}\right].
\end{aligned} \tag{29}$$

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

$$\begin{aligned} I^{\varepsilon\psi}_{\widetilde{\varepsilon}} &\leq I^{\varepsilon\psi}_{\widetilde{\psi}} \Leftrightarrow \varepsilon_R + \alpha\psi_{SP} - 2\varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\ &\Leftrightarrow \varepsilon \geq \varepsilon_1 \equiv \varepsilon_R\left[1 - \frac{\lambda(1+\chi_\varepsilon)+1}{2}\frac{\alpha_1 - \alpha}{\alpha_1}\right]. \end{aligned} \tag{30}$$

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

$$\begin{aligned} \Delta^{\varepsilon\psi}_{\widetilde{\varepsilon}} &\leq \Delta^{\varepsilon\psi}_{\widetilde{\psi}} \Leftrightarrow \varepsilon_R - \alpha\psi_{SP} + 2\alpha\psi - 2\varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\ &\Leftrightarrow \varepsilon \geq \alpha(\psi - \psi_{SP}) + \varepsilon_1. \end{aligned} \tag{31}$$

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

$$\begin{aligned} \Delta^{\varepsilon\psi}_{\widetilde{\varepsilon}} &\leq \Delta^{\varepsilon\psi}_{\widetilde{\psi}} \Leftrightarrow |\varepsilon - \varepsilon_R| + \alpha\,|\psi| - \lambda\varepsilon_R(1+\chi_\varepsilon) \leq |\varepsilon| + \alpha\,|\psi - \psi_{SC}| - \lambda\alpha\psi_{SP}(1+\chi_\psi), \\ &|\varepsilon - \varepsilon_R| - \alpha\,|\psi - \psi_{SC}| - \alpha\psi - \varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right] \equiv \lambda MC. \end{aligned}$$

Once again, the meta contrast ratio entails a general identification tendency for all types (ε, ψ).

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

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

$$\begin{aligned} \Delta^{\varepsilon\psi}_{\widetilde{\varepsilon}} &\leq \Delta^{\varepsilon\psi}_{\widetilde{\psi}} \Leftrightarrow -\varepsilon_R + \alpha\chi_\psi\psi_{SC} \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right] \\ &\Leftrightarrow \alpha \leq \alpha_2 \equiv \frac{\varepsilon_R}{\psi_{SP}}\frac{\lambda(1+\chi_\varepsilon)+1}{\lambda(1+\chi_\psi)+\chi_\psi}, \end{aligned} \tag{32}$$

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

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

$$\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} \leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_R - \alpha\chi_\psi\psi_{SP} - 2\alpha\psi \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right].$$
$$\Leftrightarrow \psi \geq \psi_2 \equiv -\psi_{SP}\left[\chi_\psi + \frac{\lambda(1+\chi_\psi)+\chi_\psi}{2}\frac{\alpha_2-\alpha}{\alpha}\right]. \tag{33}$$

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

$$\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} \leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_R - \alpha\psi_{SC} - 2\varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right].$$
$$\Leftrightarrow \varepsilon \geq \varepsilon_2 \equiv \varepsilon_R\left[1 - \frac{\lambda(1+\chi_\varepsilon)+1}{2}\frac{\alpha_2-\alpha}{\alpha_2}\right], \tag{34}$$

which is isomorphic to (30), but ε_2 is lower than ε_1 because α_2 is larger than α_1.

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

$$\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} \leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_R + \alpha\chi_\psi\psi_{SP} - 2\alpha(\psi-\psi_{SC}) - 2\varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right].$$
$$\Leftrightarrow \varepsilon \geq -\alpha(\psi-\psi_{SC}) + \varepsilon_2. \tag{35}$$

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

$$\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} \leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow |\varepsilon-\varepsilon_P| + \alpha|\psi| - \lambda\varepsilon_R(1+\chi_\varepsilon) \leq |\varepsilon| + \alpha|\psi-\psi_{SP}| - \lambda\alpha\psi_{SP}(1+\chi_\psi),$$
$$|\varepsilon-\varepsilon_P| - \alpha|\psi-\psi_{SP}| + \alpha\psi + \varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right] \equiv \lambda MC.$$

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

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

$$\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} \leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_P + \alpha\psi_{SP} \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right] \tag{36}$$
$$\Leftrightarrow \alpha \leq \alpha_3 \equiv \frac{\varepsilon_R}{\psi_{SP}}\frac{\lambda(1+\chi_\varepsilon)+\chi_\varepsilon}{\lambda(1+\chi_\psi)+1},$$

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

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

$$
\begin{aligned}
\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_P - \alpha\psi_{SP} + 2\alpha\psi \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\
&\Leftrightarrow \psi \leq \psi_3 \equiv \psi_{SP}\left[1 + \frac{\lambda(1+\chi_\psi)+1}{2}\frac{\alpha_3 - \alpha}{\alpha}\right]. \qquad (37)
\end{aligned}
$$

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

$$
\begin{aligned}
\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow 2\varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right] - \alpha\psi_{SP} - \chi_\varepsilon\varepsilon_R \\
&\Leftrightarrow \varepsilon \leq \varepsilon_3 \equiv -\varepsilon_R\left[\chi_\varepsilon - \frac{\lambda(1+\chi_\varepsilon)+\chi_\varepsilon}{2}\frac{\alpha_3 - \alpha}{\alpha_3}\right]. \qquad (38)
\end{aligned}
$$

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

$$
\begin{aligned}
\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_P - \alpha\psi_{SP} + 2\alpha\psi + 2\varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\
&\Leftrightarrow \varepsilon \leq -\alpha(\psi - \psi_{SP}) + \varepsilon_3. \qquad (39)
\end{aligned}
$$

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

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

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

$$
\begin{aligned}
\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_P - \alpha\psi_{SC} \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\
&\Leftrightarrow \alpha \leq \alpha_4 \equiv \frac{\varepsilon_R}{\psi_{SP}}\frac{\lambda(1+\chi_\varepsilon)+\chi_\varepsilon}{\lambda(1+\chi_\psi)+\chi_\psi}, \qquad (40)
\end{aligned}
$$

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

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

$$\begin{aligned}\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_R - \alpha\psi_{SP} + 2\alpha\psi \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\ &\Leftrightarrow \psi \geq \psi_4 \equiv -\psi_{SP}\left\{\chi_\psi + \frac{\lambda(1+\chi_\psi)+\chi_\psi}{2}\frac{\alpha_4-\alpha}{\alpha}\right\}. \end{aligned} \tag{41}$$

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

$$\begin{aligned}\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_P - \alpha\psi_{SC} + 2\varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\ &\Leftrightarrow \varepsilon \leq \varepsilon_4 \equiv -\varepsilon_R\left\{\chi_\varepsilon - (\frac{\alpha_4-\alpha}{\alpha_4})\frac{\lambda(1+\chi_\varepsilon)+\chi_\varepsilon}{2}\right\}. \end{aligned} \tag{42}$$

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

$$\begin{aligned}\Delta_{\widetilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\widetilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_P + \alpha\psi_{SC} - 2\alpha\psi + 2\varepsilon \leq \lambda\left[\varepsilon_R(1+\chi_\varepsilon) - \alpha\psi_{SP}(1+\chi_\psi)\right]. \\ &\Leftrightarrow \varepsilon \leq \alpha(\psi - \psi_{SC}) + \varepsilon_4. \end{aligned} \tag{43}$$

Patterns of Identification. The prevalent direction of identification is driven by the location of α relative to the thresholds α_1, α_2, α_3, α_4. When $\alpha < \underline{\alpha} \equiv \min(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, class identification $d = \widetilde{\varepsilon}$ is prevalent. When $\alpha > \overline{\alpha} \equiv \max(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$, cultural identification $d = \widetilde{\psi}$ 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 (rich or poor), $d = \widetilde{\varepsilon}$. Identification in regions C and D is instead ambiguous. A necessary and sufficient condition to have some types identified along the values dimension $d = \widetilde{\psi}$ in each of these regions is that:

$$\min(\varepsilon_1, \varepsilon_2, -\varepsilon_3, -\varepsilon_4) > 0. \tag{44}$$

By inspecting thresholds (30), (34), (38) and (42) it is immediate to see, given our assumption χ_ε, $\chi_\psi < 1$, that condition (44) is implied by $\min(\varepsilon_2, -\varepsilon_4) > 0$, which is in turn implied by:

$$\lambda < \frac{\chi_\varepsilon}{1+\chi_\varepsilon}. \tag{45}$$

Condition (45) is therefore sufficient to guarantee that some types feature $d = \widetilde{\psi}$ when $\alpha < \underline{\alpha}$.

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

$$\max(\varepsilon_1, \varepsilon_2, -\varepsilon_3, -\varepsilon_4) < 0, \tag{46}$$

which is implied by:

$$\frac{1 - \lambda(1 + \chi_\varepsilon)}{2} + \frac{\lambda(1 + \chi_\varepsilon) + 1}{2} \frac{\alpha}{\alpha_1} < 0. \tag{47}$$

By substituting α_1 in the above expression, the condition becomes:

$$-\lambda \left[(1 + \chi_\varepsilon) - \alpha \frac{\psi_{SP}}{\varepsilon_R} (1 + \chi_\psi) \right] + 1 + \alpha < 0, \tag{48}$$

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

Case $\alpha > \overline{\alpha}$. Inspection of the thresholds reveals that $\overline{\alpha} = \alpha_2$. 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), $d = \widetilde{\psi}$. Identification in regions B and D is instead ambiguous. A necessary and sufficient condition to have some types identified along the income dimension $d = \widetilde{\varepsilon}$ in these two regions is that:

$$\min(\psi_1, -\psi_2, \psi_3, -\psi_4) > 0. \tag{49}$$

By inspecting thresholds (29), (33), (37) and (41) one can see, given our assumption χ_ε, $\chi_\psi < 1$, that condition (49) is implied by $\min(\psi_3, -\psi_4) > 0$, which is in turn implied by:

$$\lambda < \frac{\chi_\psi}{1 + \chi_\psi}. \tag{50}$$

Condition (50) is therefore sufficient to guarantee that some types feature $d = \widetilde{\varepsilon}$ when $\alpha > \overline{\alpha}$.

On the other hand, when $\alpha > \overline{\alpha}$ identification universally occurs along the values dimension, namely $d = \widetilde{\psi}$ for all types provided:

$$\max(\psi_1, -\psi_2, \psi_3, -\psi_4) < 0, \tag{51}$$

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 $d = \widetilde{\psi}$ identification is given by:

$$\frac{1 - \lambda(1 + \chi_\psi)}{2} + \frac{\lambda(1 + \chi_\psi) + 1}{2} \frac{\alpha_1}{\alpha} < 0. \tag{52}$$

By plugging the expression for α_1 this becomes:

$$-\lambda \left[(1+\chi_\psi) - \frac{\varepsilon_R}{\psi_{SP}} \frac{(1+\chi_\varepsilon)}{\alpha} \right] + \frac{1}{2} \left(1 + \frac{\varepsilon_R}{\alpha \psi_{SP}} \right) < 0, \tag{53}$$

which holds when $\lambda > \widehat{\lambda}_1 > 0$, where $\widehat{\lambda}_1$ is a suitable threshold. Indeed, when $\alpha > \overline{\alpha}_2$ the expression in square brackets is positive for $\chi_\varepsilon, \chi_\psi < 1$.

By conditions (45), (47), (50) and (52) we therefore conclude that across the two regimes $\alpha < \underline{\alpha}$, $\alpha > \overline{\alpha}$ identification switches from being fully $d = \widetilde{\varepsilon}$ to being fully $d = \widetilde{\psi}$ provided:

$$\lambda > \widehat{\lambda} \equiv \max \left(\widehat{\lambda}_0, \widehat{\lambda}_1 \right).$$

On the other hand, a change across the two regimes $\alpha < \underline{\alpha}$, $\alpha > \overline{\alpha}$ preserves some $d = \widetilde{\varepsilon}$ types in all regions B and D and some $d = \widetilde{\psi}$ 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 $\overline{\alpha}$ it is easy to see that they can be expressed as in Equation (15). Which proves the proposition.

Proof of Proposition 3. To ease notation, 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 $h_x = \frac{\overline{x}}{\overline{x}+\underline{x}} \frac{1}{\underline{x}}$ and $l = \frac{\underline{x}}{\overline{x}+\underline{x}} \frac{1}{\overline{x}}$, so that:

$$\frac{l_x}{h_x} = \chi_x^2. \tag{54}$$

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

$$((\bar{\varepsilon} - \varepsilon_R) + (\varepsilon_P + \underline{\varepsilon}))(\bar{\psi} + \underline{\psi}) > (\varepsilon_R - \varepsilon_P)(\bar{\psi} + \underline{\psi}).$$

Given that $\varepsilon_R = \frac{\bar{\varepsilon}\chi_\epsilon}{2(1+\chi_\varepsilon)}$, $\underline{\varepsilon} = \chi_\varepsilon \bar{\varepsilon}$, $\varepsilon_P = -\chi_\varepsilon \varepsilon_R$ and plugging in such expressions, the previous inequality is simply verified. By the same argument, when $\alpha > \overline{\alpha}$ 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 $(\pi_P, \pi_R, \pi_{SC}, \pi_{SP})$ and check that $\pi_P \geq \pi_R$ and that $\pi_{SC} \geq \pi_{SP}$.

We start with the poor-rich cleavage by computing (π_P, π_R). To do so, we focus on two sub-cases. We fist consider socially conservative types, which entails computing $\pi_P(SC)$, $\pi_R(SC)$. We next consider socially progressive types, which entails computing $\pi_P(SP)$, $\pi_R(SP)$.

Class-identified Socially Conservative types, SC. Here $\pi_P(SC) = (\varepsilon_4 + \underline{\varepsilon})(0 + \underline{\psi})h_\varepsilon h_\psi + A_P(SC)$, and $\pi_R(SC) = (\overline{\varepsilon} - \varepsilon_2)(0 + \underline{\psi})l_\varepsilon h_\psi + A_R(SC)$, where $A_P(SC)$ are the poor-identified in quadrant 4.D while $A_R(SC)$ are the rich-identified in quadrant 2.D. We first show that, leaving quadrant D aside, the poor are more numerous than the rich:

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

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

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

In light of the definitions of α_2 and α_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:

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

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

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

$$A_P(SC) > A_R(SC) \Leftrightarrow \left(2\chi_\psi\psi_{SP} + \frac{\varepsilon_4}{\alpha}\right)(-\varepsilon_4) > \alpha\chi_\psi^2\psi_{SP}^2\chi_\varepsilon^2. \tag{55}$$

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 $A_P(SC)$ is a trapezoid, namely when $\widehat{\varepsilon}_4 > 0$ 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 (55) is fulfilled.

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

Class-identified Socially Progressive types, SP. Here $\pi_P(SP) = (\varepsilon_3 + \underline{\varepsilon})\overline{\psi} h_\varepsilon l_\psi + A_P(SP)$, and $\pi_R(SP) = (\overline{\varepsilon} - \varepsilon_1)\overline{\psi} l_\varepsilon l_\psi + A_R(SP)$, where $A_P(SP)$ are the poor-identified in quadrant 3.D while $A_R(SP)$ are the rich-identified in quadrant 1.D. Once again, we first show that, leaving quadrant D aside, the poor are indeed more numerous than the rich:

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

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

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

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

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

$$A_P(SP) > A_R(SP) \Leftrightarrow \left(2\psi_{SP} + \frac{\varepsilon_3}{\alpha}\right)(-\varepsilon_3) > \alpha\psi_{SP}^2\chi_\varepsilon^2. \tag{56}$$

This condition is fulfilled because $-\varepsilon_3 > \varepsilon_1 > \alpha\chi_{SP}$, and because - after some algebra - one can show that when $A_P(SP)$ is a trapezoid, namely $\widehat{\varepsilon}_3 > 0$ we have $\varepsilon_3/\alpha > -\psi_{SP}$, so that (56) is fulfilled. It is a fortiori true that $A_P(SP) > A_R(SP)$ when they are both trapezoids, because the area of $A_R(SP)$ is largest when it is a triangle. No other case needs to be considered because when $A_R(SP)$ is a trapezoid, $A_P(SP)$ is a tarpezoid also.

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

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 $\pi_{SC}(\alpha) \geq \pi_{SP}(\alpha)$. To compute whether socially conservative or socially progressive identification is more prevalent, we again focus on two subcases. The first is within the poor, which entails computing $\pi_{SC}(P)$, $\pi_{SP}(P)$. The second is within the socially progressive, which entails computing $\pi_{SC}(P)$, $\pi_{SP}(P)$.

Culture identified Poor types, P. Here $\pi_{SC}(P) = (0 - \varepsilon_4)(\psi_{SC} + \underline{\psi})h_\varepsilon h_\psi + A_{SC}(P)$, and $\pi_{SP}(P) = -\varepsilon_3(\overline{\psi} - \psi_{SP})h_\varepsilon l_\psi + A_{SP}(P)$, where $A_{SC}(P)$ are the socially conservatives in quadrant 4.D while $A_{SP}(P)$ 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:

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

which is easily verified because $\chi_\psi, \chi_\varepsilon < 1$. Compare now the regions $A_{SC}(P)$ and $A_{SP}(P)$. In light of the previous result, for the socially conservative identified P types to to outnumber the socially progressive-identified P types, it is sufficient to show that $A_{SC}(P) - A_{SP}(P) > 0$. Consider first the case in which both $A_{SC}(P)$ and $A_{SP}(P)$ are triangles. This occurs when α 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:

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

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 α is intermediate so that $\widehat{\varepsilon}_4 < 0$ and $\widehat{\varepsilon}_3 > 0$. In this case, we have that:

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

which also holds because $-\varepsilon_4 > \alpha\chi_\psi\psi_{SC}$ and because $\alpha\psi_{SC} > -\varepsilon_3$. This implies that $A_{SC}(P) - A_{SP}(P) > 0$ is a fortiori true when $A_{SP}(P)$ 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 $A_{SC}(P)$ is a triangle and $A_{SP}(P)$ a trapezoid. We thus conclude that in all cases $A_{SC}(P) > A_{SP}(P)$ and hence $\pi_{SC}(P) > \pi_{SP}(P)$.

Culture identified Rich types, R. Here $\pi_{SC}(R) = \varepsilon_2(\underline{\psi} + \psi_{SC})l_\varepsilon h_\psi + A_{SC}(R)$, and $\pi_{SP}(R) = \varepsilon_1(\overline{\psi} - \psi_{SP})l_\varepsilon l_\psi + A_{SC}(R)$, where $A_{SC}(R)$ are the socially conservatives in quadrant 2.D while

$A_{SP}(R)$ are the socially progressives in quadrant 1.D. We first show that, leaving quadrant D aside, the socially conservatives outnumber the socially progressive:

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

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

$$\begin{aligned} l_\varepsilon h_\psi \chi_\psi \psi_{SP}\left[\varepsilon_2 - \varepsilon_1\chi_\psi\right] &> 0 \\ l_\varepsilon h_\psi \chi_\psi \psi_{SP}\varepsilon_R\left[\frac{1-\lambda(1+\chi_\varepsilon)}{2}\left(1-\chi_\psi\right) + \frac{\alpha}{\alpha_2}\frac{\lambda(1+\chi_\varepsilon)+1}{2}\frac{\lambda(1+\chi_\psi)\left(1-\chi_\psi\right)}{\lambda(1+\chi_\psi)+\chi_\psi}\right] &> 0. \end{aligned}$$

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

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

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

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

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

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

which also holds because $\varepsilon_2 > \alpha\chi_\psi\psi_{SP}$ and because $\alpha\psi_{SP} > \varepsilon_1$. This implies that $A_{SC}(R) - A_{SP}(R) > 0$ is a fortiori true when $A_{SP}(R)$ 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 $A_{SC}(R)$ is a triangle and $A_{SP}(R)$ a trapezoid. We thus conclude that in all cases $A_{SC}(R) > A_{SP}(R)$ and hence $\pi_{SC}(R) > \pi_{SP}(R)$.

This proves that for the case $\alpha \leq \underline{\alpha}$, we have that $\pi_P > \pi_R$ and $\pi_{SC} > \pi_{SP}$. It still needs to be proved that this is also the case for $\alpha > \overline{\alpha}$, but it easily follows from the symmetry of the problem.

Correlation between ε and ψ.. If individual traits ε and ψ are correlated in the population, then the poor and rich groups differ not only in their income but also in their average social progressiveness, with group means $\psi_P \neq \psi_R \neq 0$. Symmetrically, cultural groups differ also in their income, with group means $\varepsilon_{SC} \neq \varepsilon_{SP} \neq 0$. We assume a plausible form of correlation whereby:

$$\begin{aligned} |\varepsilon_P - \varepsilon_R| &> |\varepsilon_{SC} - \varepsilon_{SP}| > 0, \\ |\psi_{SC} - \psi_{SP}| &> |\psi_P - \psi_R| > 0, \end{aligned} \tag{57}$$

so that income contrast is largest when groups are defined along the income dimension, while cultural contrast is largest when groups are defined along the cultural dimension. This assumption is always satisfied if ε and ψ are not too positively correlated or if they are negatively correlated in the population. ■

To simplify the analysis we assume that $\lambda \to \infty$, so that contrast among groups in the unique driver of identification. By Equation (14), then, everybody identifies along class lines if and only if:

$$\alpha \leq \hat{\alpha} \equiv \frac{|\varepsilon_P - \varepsilon_R| - |\varepsilon_{SC} - \varepsilon_{SP}|}{|\psi_{SC} - \psi_{SP}| - |\psi_P - \psi_R|} > 0.$$

As in Proposition 2, people identify with their class if redistribution is important relative to culture.

The effect of identification on belief distortions has much in common with the results we obtained in the zero correlation case. The only difference is that now beliefs about both ε and ψ are distorted, regardless of identification. If ε and ψ are positively correlated, the poor are less socially progressive than the rich, $\psi_P < \psi_R$. As a result, a poor-identified voter does not only exaggerate downward income risk, he also exaggerates his social conservatism, $E^{\theta}\left(\widetilde{\psi}\,|\varepsilon, \psi, P\right) = \psi + \theta\left(\psi_P - \psi_R\right)$. If civil rights become more important and α rises above $\overline{\alpha}$, the same voter switches and identifies with the socially conservative group. His beliefs remain distorted in the same direction, but his exaggeration of downward income risk falls while his exaggeration of social conservatism increases.

It is easy to extend this reasoning to all voters types and to the case of negative correlation between ε and ψ. This yields the following immediate qualification of Proposition 3.

Corollary 1 *i) If correlation between ε and ψ is sufficiently small in absolute value that (57) holds, 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 ε and ψ are positively correlated.*

The intuition for this result is the same as before. Due to low correlation between income and social progressiveness, a switch from class to cultural identity brings poor and rich voters in the same group. This mixing reduces income contrast across groups, dampening income stereotypes and belief distortions.

To assess equilibrium policy, note that under the maintained assumption that $\lambda \to \infty$ everyone identifies along the same dimension. Let τ_d^*, q_d^* denote the equilibrium policies when identification is along dimension $d = \widetilde{\varepsilon}, \widetilde{\psi}$. Repeating the same steps as above, we get:

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

which implies the following:

Corollary 2 *Suppose that $\pi_P, \pi_{SC} > 1/2$. If condition (57) is satisfied and $\pi_P \approx \pi_{SC}$, then $\tau_{\widetilde{\varepsilon}}^* > \tau_{\widetilde{\psi}}^*$ and $q_{\widetilde{\psi}}^* < q_{\widetilde{\varepsilon}}^*$. Moreover, $\tau_{\widetilde{\psi}}^* \gtrless \tau^\circ$ and $q_{\widetilde{\varepsilon}}^* \lesseqgtr q^\circ$ 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. What happens as the dimension of identification changes? Provided the size of the poor and socially conservative groups are similar, $\pi_P \approx \pi_{SC}$, a change in identification from income to culture reduces both taxation and civil rights. Taxes become less distorted, and civil rights become even more distorted, exactly as Corollary 2 reports.

Consider now the case in which ε and ψ 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, the socially conservative majority becomes key in shaping policies. Civil rights are now underprovided, and taxation is also underprovided, because socially conservatives are richer.

Proof of Proposition 4. Equilibrium beliefs for a type identified along dimension d are equal to:

$$\mathbb{E}^\theta\left(\widetilde{\varepsilon}\,|\varepsilon,\psi,G_d\right) = \varepsilon + \theta\left[\mathbb{E}\left(\varepsilon\,|G_d\right) - \mathbb{E}\left(\varepsilon\,|\overline{G_d}\right)\right], \tag{58}$$

$$\mathbb{E}^\theta\left(\widetilde{\psi}\,|\varepsilon,\psi,G_d\right) = \psi + \theta\left[\mathbb{E}\left(\psi\,|G_d\right) - \mathbb{E}\left(\psi\,|\overline{G_d}\right)\right]. \tag{59}$$

■

When identification occurs along class lines, $d = \widetilde{\varepsilon}$, only beliefs about income are distorted because - given zero correlation - $\mathbb{E}\left(\psi\,|P\right) = \mathbb{E}\left(\psi\,|R\right)$. Likewise, when identification is culture based, $d = \widetilde{\psi}$, only beliefs about civil rights are distorted, because $\mathbb{E}\left(\varepsilon\,|SP\right) = \mathbb{E}\left(\varepsilon\,|SC\right)$. This proves part i).

Part ii) is also straightforward to prove. The distortion in beliefs about $x = \varepsilon, \psi$, namely $\left|\mathbb{E}^{\theta}\left(\widetilde{x}\,|\varepsilon, \psi, G_d\right) - x\right|$ is equal to $\theta\left|\mathbb{E}\left(x\,|G_d\right) - \mathbb{E}\left(x\,\left|\overline{G_d}\right.\right)\right|$, which increases in group contrast $\left|\mathbb{E}\left(x\,|G_d\right) - \mathbb{E}\left(x\,\left|\overline{G_d}\right.\right)\right|$.

Proof of Proposition 5. The previous analysis established that $(\pi_P - \pi_R)$ and $(\pi_{SC} - \pi_{SP})$ 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 $(\pi_P - \pi_R)$ and $(\pi_{SC} - \pi_{SP})$ change from $\alpha \leq \overline{\alpha}$ to $\alpha \geq \overline{\alpha}$. We do so in both classes of identification equilibria. ■

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

$$\begin{aligned}
\left.(\pi_P - \pi_R)\right|_{\alpha \geq \overline{\alpha}} &= \frac{\chi_\varepsilon}{1+\chi_\varepsilon}\left(1 - \chi_\varepsilon\right) > \left.(\pi_P - \pi_R)\right|_{\alpha \leq \overline{\alpha}} = 0, \\
\left.(\pi_{SC} - \pi_{SP})\right|_{\alpha \leq \underline{\alpha}} &= \frac{\chi_\psi}{1+\chi_\psi}\left(1 - \chi_\psi\right) > \left.(\pi_{SC} - \pi_{SP})\right|_{\alpha \geq \overline{\alpha}} = 0.
\end{aligned}$$

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 $\pi_P - \pi_R$ for $\alpha \leq \underline{\alpha}$, and then compute it for $\alpha \geq \overline{\alpha}$. When $\alpha \leq \underline{\alpha}$ we have that:

$$\left.(\pi_P - \pi_R)\right|_{\alpha \leq \underline{\alpha}} \geq \widehat{\pi}_P - \widehat{\pi}_R = \left[\left(\varepsilon_4 + \underline{\varepsilon}\right)\underline{\psi} + \left(\varepsilon_3 + \underline{\varepsilon}\right)\overline{\psi}\right] h_\varepsilon - \left[\left(\overline{\varepsilon} - \varepsilon_2\right)\underline{\psi} + \left(\overline{\varepsilon} - \varepsilon_1\right)\overline{\psi}\right] l_\varepsilon, \quad (60)$$

where $\widehat{\pi}_P - \widehat{\pi}_R$ 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 (60). After some algebra, we can show that this is equal to:

$$\widehat{\pi}_P - \widehat{\pi}_R = \widehat{\pi}_0 + \alpha\psi_{SP}^2 h_\varepsilon\left(1 - \chi_\varepsilon^2\right)\left(1 - \chi_\psi^2\right)\left(1 + \lambda\right), \quad (61)$$

where

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

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

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

$$\left.(\pi_P - \pi_R)\right|_{\alpha \geq \alpha} \leq \pi_P^* - \pi_R^* = \left[-\psi_4\underline{\varepsilon} + \psi_3\underline{\varepsilon}\right] h_\varepsilon - \left[-\psi_2\overline{\varepsilon} + \psi_1\overline{\varepsilon}\right] l_\varepsilon, \quad (62)$$

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

We now show that, in fact, these trapezoids feature a majority of poor voters, so adding them to the count offers an upper bound of the poor-rich edge for $\alpha \geq \overline{\alpha}$. Consider first the set of socially consevative voters. Here the poor ($\varepsilon < 0$) culture identified voters with $\psi \in (\psi_4, 0)$ outnumber the rich ($\varepsilon > 0$) culture identified voters 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 rich group increases in the product of two terms that increase in $-\psi_2$ and in ε_2. Critically, the area of the rich group must be discounted by χ_ε^2 due to the higher density of poor voters. As a result, a sufficient condition for the poor to outnumber the rich 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 poor ($\varepsilon < 0$) culture identified voters with $\psi \in (0, \psi_3)$ outnumber the rich ($\varepsilon > 0$) culture identified voters with $\psi \in (0, \psi_1)$. This occurs because the area defining the poor group increases in the product of two terms that increase in ψ_3 and $-\varepsilon_3$, while the area defining the rich group increases in the product of two terms that increase in ψ_1 and in ε_1. Critically, the area of the rich group must be discounted by χ_ε^2 due to the higher density of poor voters. As a result, a sufficient condition for the poor to outnumber the rich 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 prelimnaries therefore establish the upper bound of Equation (62), which can be written as:

$$\pi_P^* - \pi_R^* = \pi_\infty^* + 2\varepsilon_R^2 h_\varepsilon \frac{\lambda \chi_\varepsilon \left(1 - \chi_\varepsilon^2\right)}{\alpha},$$

where

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

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

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

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

which implies that moving from any $\alpha \leq \overline{\alpha}$ to any $\alpha \geq \overline{\alpha} \equiv \alpha_2$ indeed reduces the numberical advantage of the poor among the class identified voters. Formally, $(\pi_P - \pi_R)|_{\alpha \leq \underline{\alpha}} > (\pi_P - \pi_R)|_{\alpha \geq \overline{\alpha}}$. By a symmetric argument, one can show that $(\pi_{SC} - \pi_{SP})|_{\alpha \leq \underline{\alpha}} < (\pi_{SC} - \pi_{SP})|_{\alpha \geq \overline{\alpha}}$.

As a result, moving from prevalent class identification to prevalent culture identification reduces excess taxation while exacerbates the excess resrtictiveness of civil rights, and viceversa when identification changes in the opposite direction. This proves the proposition.

Proof of Proposition 6. If $d = \widetilde{\varepsilon}$, the contrast among the rich and poor groups is:

$$K \equiv \varepsilon_R - \varepsilon_P, \tag{63}$$

where $\varepsilon_R = \mathbb{E}\left(\varepsilon \,|\varepsilon > \widehat{\varepsilon}\right)$ and $\varepsilon_P \equiv \mathbb{E}\left(\varepsilon \,|\varepsilon \leq \widehat{\varepsilon}\right)$. Income is uncorrelated with culture and trade attitudes, so (63) only involves income differences between rich and poor. ■

If $d = \widetilde{\eta}$, the contrast among globalists Gl and nationalists N is:

$$\gamma\left(\eta_{Gl} - \eta_N\right) + \alpha\rho\left(\eta_{Gl} - \eta_N\right) = \left(\gamma + \alpha\rho\right) K, \tag{64}$$

where γ is the importance, α that of culture (all relative to taxes). Contrast $\left(\eta_{Gl} - \eta_N\right)$ is equal to K in (63) due to the fact that the marginal distribution of η is identical to that of ε.

If $d = \widetilde{\psi}$, the contrast between socially conservative and progressive groups is equal to:

$$\left(\gamma\rho + \alpha\right) K. \tag{65}$$

Relative to class identification in (63), conflict along income diminishes, but conflict along the other dimensions widens. Relative to trade based identification in (64) conflict along trade dampens, but conflict along culture amplifies. By finding the minimum distances among Equations (63), (64), and (65) readily yields the proposition.

Proof of Proposition 7. Straightforward. ■

9.2 Empirics

9.2.1 Switching Identities

The analysis of the effect of switching identity on individual beliefs presented in section 6.1 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 variables of interest are defined as follows.

Main Variables 1) **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.

2) **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.

3) **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.

4) **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.

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.

6) **State Regulation** "In order to face economic difficulties it is necessary for the State to... [1. rely more on private firms, granting them more freedom; 2. or, on the contrary, control them by tightening regulation]". 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.

7) **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.

Variables 1-5 are used to build the first polychoric principal component "Globalization and

Immigration" while variables 6 and 7 are used in building the polychoric principal component "Role of the State in the Economy".

In order to reconcile scale differences, before performing the cluster analysis, all these variables were dychotomized. Values of 1 and 2 are coded as 0, values of 4 and 5 are coded as 1.

Figure 4 was constructed as follows. First we computed the two clusters using these 7 dychotomized variables. Clusters are defined by applying Ward's minimum variance method, and stopping the hierarchical partitioning algorithm at 2 clusters. Next, we extracted the first principal component on the five (non-dychotomized) variables 1-5, and on the two (non dychotomized) variables 6-7. We then estimated the residuals of a regression of each of these two first principal components on income and education (defined below). These residuals are plotted in Figure 4, with tones differentiated between the two clusters. This procedure is repeated for 2013 and 2017.

These same first principal components, rescaled between 0 and 1, are used as dependent variables in the regressions displayed in Table 1.

Regressors **Voting Decisions** The voting decision regarding the first and the second round of the presidential election held in 2012 are asked 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, if he voted, or whether he abstained or did not vote. 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.

For the 2017 presidential election, we use voting decision as reported in the survey of May 2017. We group first round candidates as follows:

Left: Arthaud, Poutou, Mélenchon, Hamon;

Right: Fillon, Assalineau, 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 eduation, based on the usual number of years required to achieve the specified educational level. The variable is treated as countinuous.

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.

9.2.2 Exposure to Import Competition

In the analysis of the effects of import shocks from China performed in section 6.2, we combine data from multiple datasets. All individual level variables we use come 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.

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 form 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 is 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 coming from the CCES. The other variables are described more in detail in the text or in the sources indicated above.

Dependent Variables **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.

Regressors **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).

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.

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

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

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

9.2.3 Individual Beliefs and Group Contrast

The source for the data in section 6.3 is *American National Election Studies* (ANES), surveys carried out in 1998, 2000, 2004, 2008, 2012 and 2016. The surveys are designed to represent the national population aged 18 or more and we use individual survey weights to further enhance its representativeness. The sample of the pooled data includes about 10 000 individuals, with minor differences in size depending on the attitudes studied. Below we report the questions used in the analysis and the main information on how the variables based on these questions are defined.

Dependent Variables **Redistribution** "Where would you place yourself on this scale, or haven't you thought much about this?" Answers are on a 7-point scale ranging from "Govt should see to jobs and standard of living" to "Govt should let each person get ahead on own". We drop those who didn't place themselves on such scale, reverse the direction of the scale (with high values indicating more willingness to redistribute) and rescale the variable between 0 and 1.

Immigration "Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be [1. increased a lot; 2. increased a little; 3 left the same as it is now4; 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 resulting variable between 0 and 1.

Abortion "There has been some discussion about abortion during recent years. Which one of the opinions on this page best agrees with your view? You can just tell me the number of the opinion you choose. [1. By law, abortion should never be permitted; 2. By law, only in case of rape, incest, or woman's life in danger; 3. By law, for reasons other than rape, incest, or woman's life in danger if need established; 4. By law, abortion as a matter of personal choice]". We treat this variable as continuous and rescale it between 0 and 1.

Controls **Party ID** "Generally speaking, do you usually think of yourself as a democrat, a republican, an independent or what? [Democrat, Republican, Independent, Other party, No preference]" Based on this question we build the two party dummy *democrat* and *republican*. The dummy democrat (republican) is equal to 1 if the individual identifies with the Democratic Party (Republican Party) and to 0 if the individual identifies with another party or answers "no party".

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 30 income brakets, which we replace by their mid-points. We take the logarithm of the resulting values. The inclusion of time fixed effects in all regression models makes general deflating unnecessary.

Education Self-reported highest level of education attained. Grades of education completed or highest degree obtained. When the latter is reported, we convert it in years of education.

Age Self reported age and its square. The latter is included 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.

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

Marital Status Self reported, indicates whether the individual is married, widowed, divorced, separated or has never been married.

Work Status Categorical variable indicating whether the respondent reports being employed, temporarily laid off, unemployed, retired, unable to work, homeworker, or a student.

Religion Based on three variables containing information on whether respondents consider themselves as religious, and if they do, on their religion. Individuals are classified in 5 categories: non religious, protestant, catholic, jewish and other.

Region Macro-area of residence. The considered areas are Northeast, Northcentral, South, West.

Labor Union Household labor union membership. Dummy equal to one if at least one member of respondent's household is member of a labor union.

We also include pairwise interactions between income, education, age and gender, as well as interactions between income and race, education and race, and interactions of religion with income, education, age and race.