

Identity, Beliefs, and Political Conflict^{*}

Nicola Gennaioli[†] and Guido Tabellini[‡]

First version: October 2018 This version: March 2019

Abstract

We present a theory of identity politics that builds on two ideas. First, voters identify with the social group whose interests are closest to theirs and that features the strongest policy conflict with outgroups. Second, identification causes voters to slant their beliefs 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 Giampaolo Bonomi, Ben Enke, Nathan Nunn, Raffella Piccarreta, Giacomo Ponzetto, Jesse Shapiro, Andrei Shleifer and Marco Tabellini 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 science (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 assume that – for the purpose of politics – the individual identifies with the group whose interests are relatively more similar to his own, relative to the outgroup. Once identified, the individual slants his beliefs toward the stereotypical group condition, which is modelled – following Bordalo et al. (2016, 2018) – as the condition relatively most likely in the group in question.

First, we study a one-dimensional setting with heterogeneous individuals, where groups are pre-defined, everyone is exogenously identified with one of the two groups, and individual beliefs are slanted towards the stereotype of their group. The main result is that identification leads to belief polarization. Once I identify with a group of poor individuals, I exaggerate my poverty, because this is the trait that distinguishes my group from the others, and the reverse happens if I identify with the rich. 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. In a standard model of electoral competition and probabilistic voting, these belief distortions influence the political platforms offered by opportunistic political candidates.

Next, we allow individuals to differ along two dimensions, income and culture. The policy space is also two dimensional, namely a proportional income tax and a civil rights or immigration policy. Here individuals have multiple potential identities. Identification occurs along the dimension that maximizes similarity between the individual and his group, and that maximizes contrast between the in-group and the out-group. The dimension of identification is endogenous and responds to the external environment. If redistribution is more important than culture for individual welfare, then he identifies with his income group, while he identifies along the cultural dimension if cultural policy is more relevant. Group contrast also determines the dimension of identification. As a result, changes in the nature of economic conflicts can change individual identities. Since identification matters for beliefs through stereotypes, changes to the economic and social environment lead to changes in political beliefs and, through voters' behavior, in political outcomes. A shock that increases the relevance of say immigration policy or civil rights induces individuals to identify along the cultural dimension, abandoning income-based identification. This dampens redistributive conflicts and exacerbates cultural conflicts, and the equilibrium policy responds accordingly. We show that this same approach can be adapted to trade policy. Globalization or trade shocks can give

rise to economic nationalism, while at the same time dampening support for more traditional redistributive policies.

If new cleavages in society are correlated across several policy dimensions, changes in underlying economic conditions can affect beliefs over seemingly unrelated domains. For instance, suppose that voters who demand trade protection also dislike immigrants and oppose extensions of civil rights, and viceversa for advocates of free trade. This may be due to education or 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 section 6 we take some of these predictions to the data. First, we explore time patterns in US survey data. The perceived importance of immigration and race has risen in recent years. This was accompanied by a strong increase of polarization over immigration and some dampening of polarization on government spending, as predicted by the theory if the increased relevance of immigration induced identity switches, from income-based to culture-based. We also find that individuals more exposed to rising imports from China became more willing to accept cuts in domestic public spending, more averse to immigrants, and consider abortion issues more important. This is what the theory predicts, if the trade shock induced relatively poor respondents to abandon class based identification and switch to nationalism. Second, we study a longitudinal survey on France. Between 2013 and 2017 the prevailing dimension of political conflict changed from left vs right to nationalist vs globalist. This change in political cleavages is reflected in how people voted in Presidential elections and in how their attitudes changed, with a dampening of redistributive conflict and an exacerbation of conflicts over globalization, immigration and civil rights, 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. A very large literature seeks to explain changing patterns of polarization in survey data, also with recent contributions by economists (see Layman et al. 2006, Gentzkow 2016, Glaeser and Ward 2006, Glaeser, Ponzetto and Shapiro 2005).

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 Identity Approach" and discuss how we apply it to political beliefs. In Section 3 we illustrate how identities distort beliefs, in a simple one-dimensional model of redistribution. Section 4 studies how endogenous identities are affected by changes in the social environment, in a two dimensional model where redistributive conflict coexist with conflict over culture. Section 5 studies cleavages associated with globalization vs nationalism in a three dimensional model. Section 6 takes 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 identity and intergroup relations. It combines Social Identity Theory (SIT, Tajfel and Turner, 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 difference between two physical stimuli (e.g. the length of two segments) when they are labelled as belonging to groups that differ in stimulus strength (a short versus a long segments group). Results of this kind, later

¹Diermeyer 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.

confirmed in social settings (Haslam and Turner 1992), motivated a program based on the idea that individual perceptions are malleable, and may be shaped by the social context. When a certain social cleavage is salient, identification with a group anchors perceptions of self, of others, and behavior. We now discuss the building blocks of this mechanism, also in reference to our model.

2.1 Identification

How does identification with a group take place? According to SCT, this is an individual decision, a form of self-categorization. Each person belongs to many groups: income class, religious group, nation, etc, but he does not identify with all of these groups, certainly not equally so. Identification depends on the social context. For instance, at a soccer match one's own team is the main group of identification, at least in the sense that it affects behavior more than other groups to which the person belongs. Experiments confirm that subjects' beliefs and behavior change as one of our membership groups is exogenously made salient.

According to SCT, the salience of a group depends on its current accessibility and fit (Oakes, 1987). Accessibility refers to the ease with which a group comes to mind. At a soccer match, the group of "team supporters" and that of "soccer lovers" are accessible, while the group of "churchgoers" is arguably much less so. Fit captures the extent to which a group is diagnostic of relevant social differences. At the soccer match, the most fitting group is that of "team supporters", for it reflects the relevant social division of the moment. The importance of group differences and conflict for identification gives rise to the so called "meta contrast ratio": given the salient cleavage, an individual tends to identify with a group of people who are: i) similar to him, but also ii) highly dissimilar from, or in conflict with, the outgroup.²

This approach remedies some difficulties of SIT. This theory viewed identification as a mean of achieving higher self esteem, which clashes with the fact that individuals often identify with low status groups. This pattern is easy to account under the meta contrast principle: it is the presence of social differentiation or conflict that prompts identification, not the status of the group.

In politics, this idea implies that the type and severity of conflict shapes individual identities. During times of redistributive conflict, one's own income class offers a fitting identity for political action, the more so the stronger conflict is. During times of regional or cultural clash, one's own membership along these fractures becomes key, and class identity subdues. The most politically relevant identities can change over time.

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

2.2 Depersonalization and Beliefs

But how does identification with a group influence individual perceptions and behavior? According to SIT and SCT, identification causes individuals to "depersonalize", namely to re-define their self in terms of group membership. As a result, the perception of self and others is tainted by group features, in line with early work by Sherif (1936) and Festinger (1950). Experiments using the "minimal group paradigm" found that even arbitrary groups influence behavior in this way (Tajfel, Billig, et al. 1971).

In particular, according to SCT, identity causes individual beliefs to move toward the group norm. Importantly, the latter is not the average opinion in the group, it is a group prototype. McGarty et al. (1992) define such prototype as follows: "...the less a person differs from ingroup members and the more he or she differs from out group members, the more representative he or she is of the in-group. Thus, the prototype is the position which best defines what the group has in common, in contrast to other relevant out-groups." Thus, the prototype is more extreme than the mean for it creates a sharper contrast with outgroups.

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). Identification with a group leads individuals to sharpen differences in opinion from outgroups, increasing extremism.

Our model captures precisely this effect, so that identification leads to perceptions of self and others that stereotype real differences in material interests among the conflicting groups. SIT highlights another effect of depersonalization, namely that it induces individuals to derive positive self-esteem from the status of their group. Some experiments found evidence of ingroup favoritism, consistent with the idea that ingroup success enhances self-esteem. We do not include this mechanism in our model, but we show that our belief-based mechanism can endogenously generate a taste for ingroup success.³

2.3 Stereotypes and Beliefs

We conclude this section by describing how we model depersonalization. To capture the group prototype toward which individual beliefs move, we adapt the CBGS (2016) model of stereotypes, which is based on Kahneman and Tversky's representativeness heuristic.

Suppose that two groups G and \bar{G} partition society along a certain trait, for instance education, with G denoting the group of highly educated, and \bar{G} its complement. Denote by X a random variable distributed in the population with density $f(X)$, for instance X could

³The generality of the ingroup bias phenomenon has been challenged in several experiment (e.g., Hinkle and Brown 1990). It is now believed that ingroup bias arises only when discrimination is a group norm. The role of ingroup norms in influencing social preferences is consistent with our emphasis on beliefs.

refer to individual income. BCGS (2016) define a certain realization X of the random variable to be more stereotypical of group G if it scores higher in the likelihood ratio:

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

where $f(X|G)$ is the distribution of X in group G . The stereotypical values of X for G are those that are relatively more likely in this group relative to the contrasting group \bar{G} . Equation (1) captures the notion of prototype in McGarty et al. (1992): the stereotypical income for G is the income level most frequent in G in contrast to \bar{G} . BCGS (2016) offer empirical support for this notion, which also accounts for many doombfounding laboratory puzzles in probability judgments and for belief distortions in macro and finance (see Gennaioli and Shleifer 2018).

Our setting is slightly different. The random variable X is an object about which individual i forms his beliefs, rather than a specific individual trait. X can be an uncertain income prospect for individual i , e.g. his future personal income, or an unobserved aggregate outcome such as income inequality or the number of immigrants. For individual i , variable X is distributed according to density $f(X|i)$. If X is an aggregate, outcome, $f(X|i)$ captures i 's information. If X is an individual outcome, $f(X|i)$ captures individual features associated with that outcome (e.g., i 's productivity). Thus each individual corresponds to a distribution (rather than to the realization of a random variable as in BCGS 2016), and the group to a family of distributions. To obtain closed form solutions, we rewrite the likelihood ratio (1) by conditioning on the average group members, i_G and $i_{\bar{G}}$ respectively, rather than on G and \bar{G} .⁴

To formalize depersonalization, we assume that, in forming his beliefs over X , individual i overweights conditions that are stereotypical of the average member of the group he identifies with, in contrast with the average member of the opposing group. In particular, an individual identified with G perceives X to be distributed according to the distorted density:

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

where $Z \geq 0$ is a constant ensuring that $f^\theta(X|i, G)$ adds up to one and $\theta \geq 0$ captures the strength of stereotypes. When $\theta > 0$, an individual identified with G slants his opinion towards those that are prototypical of his group. If X refers to personal income, i slants his assessed economic prospects toward those distinctive of his group. If X refers to an aggregate outcome, like the number of immigrants, i slants his personal opinion towards the distinctive

⁴Our qualitative results, however, are robust to mechanically applying equation (1). Formally, the complications arise because the group level distributions such as $f(X|G)$ are not in the same class of the individual distributions $f(X|i)$ out of which they are aggregated. Replacing group level distributions with the distribution of the mean individual in the group remedies this problem, allowing for convenient closed forms.

opinion of the group he identifies with.

BCGS (2016) use this approach to characterize beliefs about groups. Using data from the Moral Foundations Questionnaire (Graham et al. 2012), they show that Equation (1) accounts for perceptions about the average liberal or conservative US voter, which exaggerate differences between the true average liberal and conservative positions. Here we go one step further, allowing stereotypes to also affect the beliefs that an individual i has about himself, or about aggregate variables. This is what effectively captures depersonalization: individual perceptions are slanted toward the group stereotype. Another important difference relative to BCGS (2016) is that here we endogeneize the group G an individual identifies with. This depends, in line with SCT, on the salience of different types of social conflict as measured by similarity and the meta contrast ratio.

BCGS (2018) document that group stereotypes indeed affect beliefs about self, and crucially that they do so on a wide range of domains. In their experiment, subjects assess their own ability in different domains of knowledge. They find that women appear on average too little self-confident in domains where women do worse than men as a group, e.g. mathematics, rock bands, sports and business. By contrast, women appear more self-confident in domains where they do better than men as a group such as literature, emotion recognition, or Disney movies. Consistent with depersonalization, individuals slant their perceived ability toward the group stereotype, captured by the relative performance of their gender group, more strongly so when gender is made more salient.

In the rest of the paper we apply these ideas to analyze identity politics. When thinking about difficult questions such as "What is the optimal tax rate?", "Should same-sex adoptions be allowed?", or "Is Globalization good?", individuals do not just think about their material interests and information. Being social animals, they think of themselves as members of the salient social group, which holds more or less dispositive norms on these issues. Beliefs also reflect group norms and stereotypes. When the salience of groups changes, the voter's identification can also change, influencing his beliefs across several domains.

3 Identity and political beliefs

We now introduce group identity in the simplest political economy model of income taxation and public good provision. Economic and class conflict have been central in western political systems during the 20th century, so this is a good place to start. Section 4 then shows that our model can account for a decline in class conflict despite persistent or even increasing income inequality.

There is a measure one of individuals, each of which earns stochastic income $1 + \tilde{\varepsilon}$, where $\tilde{\varepsilon}$ has an individual-specific mean ε . Type ε is distributed in society according to cdf $H(\varepsilon)$ in

a compact support, and it has zero mean. Thus, ε is the voter's expected relative income and aggregate income is 1. Throughout the paper, unless noted otherwise, by expected income we mean rationally expected income.

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

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

where the last term on the right side is the utility from the public good. The rational bliss point of type ε is

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

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 $T_\varepsilon < 0$ (subscripts of a function denote partial derivatives).

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

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

3.1 Group Identity and Beliefs

Distorted beliefs arise because the individual's relative income is stochastic. In a political context, the voter thinks about the policy consequences for himself as a member of a social group. When assessing a redistributive income tax, the politically salient group is made up of people belonging to the same income group. For the purpose of political action, he depersonalizes and views himself close to the prototypical member of his group. As a result, he slants his assessment toward the group's distinctive income prospects. This causes the generic voter type ε to misrepresent his income risk away from his true individual reality.

Society is partitioned in two groups based on expected income: the poor $P \equiv \{\varepsilon | \varepsilon < \hat{\varepsilon}\}$, and the rich $R \equiv \{\varepsilon | \varepsilon \geq \hat{\varepsilon}\}$, where $\hat{\varepsilon}$ is an historically given class boundary. Each group is summarized by its average member $\varepsilon_P \equiv E[\varepsilon|P]$ and $\varepsilon_R \equiv E[\varepsilon|R]$, with $\varepsilon_P < \varepsilon_R$. All voters are identified with their class and groups are exogenous. This stark assumption allows us to focus on how identification affects beliefs. In Section 4 we allow voters' identification to change.

We formalize stereotypes as in equation (2) in section 2. Consider voter with mean income ε , and let $f(\tilde{\varepsilon}|\varepsilon)$ denote the true pdf of his random individual income. This is the analog of $f(X|i)$ on the RHS of equation (2). Consider now the stereotypes term in (2). If the voter is poor, $\varepsilon \in P$, the distribution of his ingroup income is $f(\tilde{\varepsilon}|\varepsilon_P)$ and that of his outgroup income is $f(\tilde{\varepsilon}|\varepsilon_R)$. The stereotype term for a rich voter is inversely defined. Inserting these terms in equation (2), we can characterize beliefs as follows. Let ε_G^θ denote the stereotyped income ranking $E^\theta(\tilde{\varepsilon}|\varepsilon, G)$ of the generic voter type ε identified with group G . We have:

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

$$\varepsilon_G^\theta = \varepsilon + Z \cdot \text{cov} \left(\tilde{\varepsilon}, \left[\frac{f(\tilde{\varepsilon}|\varepsilon_G)}{f(\tilde{\varepsilon}|\varepsilon_{\bar{G}})} \right]^\theta \right). \quad (6)$$

If $f(\tilde{\varepsilon}|\varepsilon)$ is normal with variance σ_ε^2 , then expected income is:

$$\varepsilon_G^\theta = \varepsilon + \theta(\varepsilon_G - \varepsilon_{\bar{G}}). \quad (7)$$

Equation (6) says that, if income risk $f(\tilde{\varepsilon}|\varepsilon)$ exhibits the Monotone Likelihood Ratio Property (MLRP) for $\tilde{\varepsilon}$ 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 $f(\tilde{\varepsilon}|\varepsilon)$ 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 income difference between rich and poor, and viceversa if $\varepsilon \in P$. The distortion is larger the greater is the parameter θ that weights the likelihood ratio in (2). In the rest of the paper, we set θ to be the same for all individuals and we stick to this convenient formula by assuming that income (and other traits) are normally distributed. The assumption that θ is the same for all can easily be relaxed and yield additional implications.⁵

An important implication of (7) is that policy evaluations are also distorted by social identities. By (7), the expected utility of type ε is:

$$W^{\varepsilon\theta}(\tau|G) = (1 + \varepsilon_G^\theta)(1 - \tau) + \nu(\tau - \frac{\varphi}{2}\tau^2), \quad (8)$$

and his bliss point is distorted by stereotypes: $\tau^{\varepsilon\theta} = T(\varepsilon_G^\theta)$, where θ denotes stereotyped evaluations.

SIT, and its formalizations in economics (e.g., Shayo 2009), hold that identification causes individuals to favour ingroup members relative to outgroup members. Interestingly, our belief-

⁵We can 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.

based mechanism can create similar effects. Indeed, Equation (8) can also be written as:

$$W^{\varepsilon\theta}(\tau|G) = W^\varepsilon(\tau) + \theta \left[W^G(\tau) - W^{\bar{G}}(\tau) \right] \quad (9)$$

where $W^G(\tau)$ refers to the rational welfare of an individual with expected income equal to the group average, ε_G . Identification distorts true individual welfare $W^\varepsilon(\tau)$ toward the 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.⁶ Crucially, whether ingroup bias occurs or not in our model depends on group norms or beliefs about outgroups.

Before moving to policy, we illustrate some implications of our model for policy preferences. By (4), the difference in the bliss points of an average rich and an average poor individual is:

$$\tau^{P\theta} - \tau^{R\theta} = (\tau^P - \tau^R)(1 + 2\theta) \quad (10)$$

where τ^G is 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 upward mobility and demand too little redistribution. Overall, stereotypes increase polarization across income groups. The effect is stronger the larger is θ . If exposure to digital media such as Facebook or Instagram strengthens stereotypical thinking θ , then it also leads to greater polarization.

Westfall et al. (2015) show that US voters perceive an exaggerated difference in the policy views of Democrats and Republicans, and that the gap has increased over time. Our model can create this effect. Suppose that there are two abstract groups, *Left* and *Right*, where *Left* contains a higher share of poor-identified voters than *Right*. These groups could correspond to supporters of different political parties with opposite ideological positions, or they could be defined in other ways. Denote by $\tau^{Left,\theta}$ and $\tau^{Right,\theta}$ the actual average positions (distorted

⁶This 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 $f(u|q, \bar{G})$ is the distribution of outgroup \bar{G} utility. The perceived expected utility under q is:

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

where $\mathbb{E}(u|q)$ is the rational expectation, and Z a normalizing constant. Once again, policies that increase the probability for the outgroups to achieve high utility states cause the ingroup to focus on low utility outcomes. These policies are thus disliked even if they increase true utility $\mathbb{E}(u|q)$. When the distributions are normal with identical variance, we obtain the exact formula in (9).

by stereotypes) of the *Left* and *Right* groups. Following BCGS (2016), it is easy to show that the stereotyped perception of the average positions of these two groups are even more polarized than the actual average positions. When thinking about the left, beliefs overweigh the overly represented poor voters. When thinking about the *Right*, beliefs overweigh the overly represented rich voters. Thus, the left is stereotyped as being too redistributionist, the right as being too laissez faire. This effect is not limited to redistribution. It holds with respect to any policy issue as long as the two parties differ in the distribution of opinions over such issue. When parties differ in their composition in terms of social groups, group identities transfer to political parties the "us versus them" logic of group conflict, enhancing perceived partisan polarization.

Gentzkow (2017) shows that in the US perceived polarization and distrust of political rivals (so called affective polarization) has increased more than actual divergence in policy views. Proposition 1 shows that perceptions can indeed amplify growing underlying differences. This phenomenon can be reinforced by a rise in θ , fueled for instance by stronger stereotyping in traditional or social media. Interestingly, Ahler (2018) shows that correcting misperceptions and stereotypes about the out-party also reduces affective polarization.

Bordalo et al. (2019) test a more specific prediction of this model, namely that the gap between perceived and actual polarization between Democrats and Republicans is larger on the policy issues that are more stereotypical of differences between these two groups (based on the likelihood ratio in (1)). The evidence is consistent with this prediction, particularly after 1989 (the end of the cold war).

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). The model could be extended to partisan politicians, creating a party structures similar to the *Left* and *Right* one described above. To isolate the role of beliefs, we assume that all voters have the same degree of mobility across parties. In this case, the equilibrium policy maximizes perceived utilitarian welfare:

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

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

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

where τ° is the socially optimally policy defined in (5) and $\pi_G = \Pr(\varepsilon \in G)$ is the population share of group $G = P, R$. Since $(\varepsilon_P - \varepsilon_R) < 0$, 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 outnumber the rich, namely the threshold $\widehat{\varepsilon}$ separating the two groups is large enough. The direction of the policy distortion, however, is not key. The important point is that beliefs and policy preferences are shaped by identification, which alters the effect of economic change. First, identities distort perceptions. If taxation is initially excessive, higher inequality $(\varepsilon_R - \varepsilon_P)$ boosts taxes a lot, to an extent that increases in θ . This occurs because many lower middle class voters become excessively afraid of poverty.⁷ Second, economic shocks may cause identities to change. For instance, demand for redistribution may wane if economic shocks increase the salience of other social cleavages. Sections 4 and 5 study this implication.

3.3 Discussion

In this simple model, demand for redistribution is distorted by a clash between rich and a poor identified voters. To improve realism, we could allow for the middle class to act as a third group of identification, consistent with survey data showing that many people self-classify as middle class. We conjecture that the logic of stereotypes would cause voters in this group to shrink their perceived incomes toward the middle, and in particular toward an income level that is slightly lower than their true class average. This would create a large and cohesive

⁷A large literature (eg. Alesina and Glaeser 2005) studies why in certain countries demand for redistribution is puzzlingly low. Here we do not analyze this phenomenon, but identity may play a role. If many people identify with the rich, they will exaggerate the perception of upward income mobility. These voters are thus naturally hostile to taxation, the more so the larger is inequality $(\varepsilon_R - \varepsilon_P)$.

group supporting redistribution, consistent with the intuition of Director's law. However, we expect the broad qualitative properties of our model to be robust to this modification.⁸

Another simplification is that political parties do not have partisan or ideological positions, and voters identify with social groups. In their classic study of American voters, Campbell et al. (1966) document that during the years of the New Deal many working class voters turned Democratic because they viewed this party as representing "the common man". They also document that during the same period, many American workers identified themselves as working class, consistent with an important role of class identity in shaping party allegiances. This suggests that parties may be catering to some key social groups, and perhaps gain some voter identification from doing so. Sometimes parties themselves create new identities, particularly when charismatic leaders are able to enhance the salience of a certain social cleavage. We leave this interesting issue for future work and we focus instead on grassroots social identification.

The hallmark of our model is that identity affects preferences through beliefs. This differs from Shayo (2009) and Grossman and Helpman (2018), where identification creates ingroup altruism. Recent evidence is consistent with a key role of beliefs. Alesina et al. (2018a) find that in some western countries leftwing voters are less optimistic than rightwing voters about social mobility, and that demand for redistribution increases when subjects are provided pessimistic information about redistribution. In this case, beliefs concern aggregate outcomes rather than individual conditions, but as discussed in Section 2 our model could capture this case, with competing social groups that differ in their views about social mobility or inequality.

Using a survey on Buenos Aires residents, Cruces et al. (2009) offer evidence on misperceptions of individual income. They show that these misperceptions are due to the fact that individual income is compared to a reference group. In their analysis, voters from rich neighborhoods underestimate their income rank relative to voters from poor neighborhood, as if each voter uses the local income distribution as a reference. They also find that preferences for redistribution change in the expected direction when subjects are told their true income rank. In their context belief distortions go in the opposite direction than in our model, namely they reduce perceived income polarization. However, the same "reference income" mechanism can enhance polarization when class conflict is highly salient, so that the reference group for

⁸This conjecture follows from viewing each voter in one of the three groups as slanting his personal income toward the type that is most distinctive of his group (poor, middle, rich) relative to the income distribution in the rest of society. In this comparison, the poor and the rich still feature stereotypically low and high incomes, respectively. When thinking about the middle class stereotype, then, two things must be noted. First, the average middle class income is lower than average income in society as a whole (due to right skewness of the income distribution). Second, income distribution within the middle class has much thinner tails than in the rest of society. As a result, it is likely that voters identified with the middle class would exhibit a moderate downward distortion of their income but also converge to a tighter distribution. To characterize this model analytically, it may be easier to take as a reference distribution the overall distribution of income, not the income distribution in the rest of society.

the poor becomes the rich and vice versa.

In general, existing evidence on beliefs does not address one central question in our approach, namely the extent to which belief distortions depend on group identity. This is an important avenue for future work.

4 Endogenous Identity

How do individuals form their social identity? This question is particularly relevant when groups and political conflict (and hence identities) can be defined along more than one dimension, such as income, or ethnicity, or cultural background. This section adds a second policy instrument, q , that creates another dimension of heterogeneity and conflict. Here we interpret q as a civil rights policy, or an immigration policy, with higher q denoting a more progressive policy, but other interpretations are possible. Identities can be defined based on income (rich vs poor, as in the previous section), or on attitudes towards civil rights or immigration (socially conservative vs progressive). We then study how increased conflict over q may induce identity switches, affecting beliefs and policies. The main insight is that endogenous identity amplifies the political effects of social changes. Increased conflict over civil rights or immigration creates new cleavages, and individuals react by forming new social identities. This in turn changes belief distortions: new stereotypes are formed, old stereotypes are abandoned. Hence the effects of social changes on equilibrium policies are amplified, relative to a world in which identities and stereotypes are absent.

Individual preferences over q are represented by the quadratic loss:

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

where $\kappa > 0$ captures the weight of policy q relative to taxation τ , and $\tilde{\psi}$ is an individual-specific random variable distributed according to pdf $z(\tilde{\psi} | \psi)$, with ψ denoting the individual-specific mean. Thus, 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 $z(\tilde{\psi} | \psi)$ to be normal.

An individual type 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:⁹

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

⁹We omit for simplicity the additive constant $k\text{Var}(\psi)$ that obtains when computing $\mathbb{E}(q - \tilde{\psi})^2$.

The utility function is conveniently separable, so that the bliss point for taxes only depends on expected income, as in (4), while the bliss point for civil rights only depends on social progressiveness, $q^\psi = \psi$. The socially optimal tax rate τ° is still given by (5) in the previous section, while the socially optimal civil rights policy is $q^\circ = 0$.

With two policy dimensions, there are four possible primitive groups, illustrated by the four quadrants in Figure 1. As in the previous section, individuals are either rich (R) or poor (P), depending on whether their expected income ε is above or below the threshold $\hat{\varepsilon}$. In addition, they can be either socially conservative, $SC \equiv \{\psi | \psi < \hat{\psi}\}$, or socially progressive, $SP \equiv \{\psi | \psi \geq \hat{\psi}\}$. As with income, the boundary $\hat{\psi}$ is historically determined, and such that the socially conservatives are a majority, formally $\pi_{SC} > \pi_{SP}$ (it is still true that $\pi_P > \pi_R$). To simplify the analysis, we assume that groups are divided at the mean $\hat{\psi} = \hat{\varepsilon} = 0$, and that the attributes ψ and ε are uncorrelated in the population. In the appendix we show that similar results hold with correlation. Introducing correlation leads instead to new insights with three (or more) dimensions, as we show in Section 5.2.

Figure 1

Given that there are two dimensions of heterogeneity, any individual has multiple potential identities. For instance, a type in the lower-left corner of Figure 1 is poor ($\varepsilon < 0$) and socially conservative ($\psi < 0$). He can thus identify with his social class (P), with his cultural group (SC), or both. To reduce the number of possible cases, we assume that only two alternatives are available for identification. Individuals can either identify along the income dimension (in which case they are either rich or poor), or along the cultural dimension (and then they are either socially conservative or socially progressive). This rules out identification with a more granular group, defined along both class and culture (e.g. the poor and socially conservative). This restriction is related to the assumption that individuals cannot be unengaged or individualistic, because the latter condition is the limit case of group granularity. We discuss it further below in footnote 11.

When a voter identifies, he chooses a dimension of conflict that partitions society into ingroups and outgroups. We denote a voter's ingroup and outgroup by G and \bar{G} , respectively. A poor voter identifying with his income group ($G = P$) views all rich as outgroups ($\bar{G} = R$), regardless of their cultural views and of how they identify themselves, and similarly for culture-based identification. Each group G is summarized by its average member, (ε_G, ψ_G) , which is used as a reference or "model" for the 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, the average rich and poor share the same social progressiveness $\psi_P = \psi_R = 0$. Consider next cultural groups. The average

progressive member has culture $\psi_{SP} \equiv E[\psi|SP] > 0$, the average conservative member has culture $\psi_{SC} \equiv E[\psi|SC] < 0$. Again, due to lack of correlation, the average members of cultural groups share the same income $\varepsilon_{SP} = \varepsilon_{SC} = 0$. In sum, if attributes (ε, ψ) are uncorrelated, then groups differ from the average voter (0,0) only in the dimension defining their membership, as shown in Figure 1. If instead ε and ψ are correlated, then groups differ in both dimensions. As we shall see, this would lead to belief distortions correlated across issues.¹⁰

Note that the groups available for identification and their average types are primitives of our model: they are the unique, historically determined, social categories available for political action, and do not depend on the positions taken by political parties, nor on the identity choices of other voters. This simplifying assumption captures the difficulty for political entrepreneurs to mold new identities. The strongest magnets of identification are often groups that, either due to history or to clear economic/political boundaries, are recognized as agents of political conflict. Lipset and Rokkan (1967) discuss the relevance of certain broad groupings as carriers of identity and vehicles of political conflicts throughout history. Salient examples are "center versus periphery", "tradition/religious versus modernization/secular", "rich/elite versus poor/masses" etc. In our two dimensional setting, the natural groups are defined by the rich-poor divide and by the cultural divide. Understanding how political parties can shape perceived social categories is an important problem for future work.

4.1 Similarity and Identification

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

We capture conflict between a voter (ε, ψ) and his ingroup (ε_G, ψ_G) by the welfare loss:

$$W^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) - W^{\varepsilon\psi}(\tau^G, q^G), \quad (13)$$

¹⁰Zero correlation also implies that the mean income of the poor ε_P is closer to zero than the mean income of the rich ε_R , for $\varepsilon_P = -\frac{\pi_R}{\pi_P}\varepsilon_R$. Likewise, the average socially conservative ψ_{SC} is closer to zero than the average socially progressive ψ_{SP} , for $\psi_{SC} = -\frac{\pi_{SP}}{\pi_{SC}}\psi_{SP}$. Also note that a measure of income polarization between the two groups is $\varepsilon_R - \varepsilon_P = (1 + \frac{\pi_R}{\pi_P})\varepsilon_R$, which increases in the average income of the rich, ε_R , and in the relative size of the rich vs the poor, $\frac{\pi_R}{\pi_P}$. A similar expression holds for polarization in values, $\psi_{SC} - \psi_{SP} = (1 + \frac{\pi_{SC}}{\pi_{SP}})\psi_{SC}$.

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

$$W^G(\tau^G, q^G) - W^G(\tau^{\bar{G}}, q^{\bar{G}}), \quad (14)$$

borne by the ingroup (ε_G, ψ_G) when moving to the bliss point of the outgroup $(\varepsilon_{\bar{G}}, \psi_{\bar{G}})$. Note that these conflicts are defined on the basis of rational bliss points. Footnote 12 below shows that little would change if instead they were computed from the stereotyped bliss points.

Combining these two criteria, we define the relative distance between voter (ε, ψ) and ingroup G as:

$$\Delta^{\varepsilon\psi}(G) = [W^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) - W^{\varepsilon\psi}(\tau^G, q^G)] - \lambda [W^G(\tau^G, q^G) - W^G(\tau^{\bar{G}}, q^{\bar{G}})]. \quad (15)$$

The voter feels similar to group G , and hence is more likely to identify with it, when $\Delta^{\varepsilon\psi}(G)$ is low. That is, he feels similar to his ingroups when overall conflict with them is small, relative to the overall conflict between ingroups and outgroups. The parameter $\lambda \geq 0$ captures the importance of cross group conflict relative to ingroup congruence as a driver of identification.¹¹

To make the model tractable, we approximate relative distance by the expression:¹²

$$\Delta^{\varepsilon\psi}(G) \simeq |\varepsilon - \varepsilon_G| + \alpha |\psi - \psi_G| - \lambda (|\varepsilon_G - \varepsilon_{\bar{G}}| + \alpha |\psi_G - \psi_{\bar{G}}|). \quad (16)$$

The appendix shows that (16) is a shorthand for a second order approximation of the expression on the RHS of (15), 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) then α rises and voters feel stronger similarity along culture than along income. The same happens if taxes become more distorting (φ

¹¹Note that a higher λ also favors identification with broad rather than narrow group. Intuitively, I am most congruent with a group consisting only of myself, but cross-group contrast is increased by including in the in-group more extreme members than myself. One can show that if λ is large then, depending on parameter values, a poor and socially conservative voter identifies as either poor or socially conservative, even if a more granular group of "poor and socially conservative" voters is available for identification. The reason is that group contrast is maximized along the single most important dimension. Hence, our restriction that the most granular groups are not available for identification can be interpreted as saying that cross-group contrast has a strong weight in determining the salience of alternative social categories.

¹²To obtain (16), we take a second order approximation of (13) around $(\tau^\varepsilon, q^\psi)$, and of (14) around (τ^G, q^G) , and replace quadratic distance with absolute value. Little would change if we define the conflict of interest using the stereotyped (rather than rational) bliss points. In this case, equation (16) becomes:

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

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

risers) and if the public good is more valuable (ν rises), since in both cases income differences become a less important source of disagreement - see (4).

Individual (ε, ψ) identifies by selecting the dimension along which relative distance $\Delta^{\varepsilon\psi}(G)$ 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, 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, 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. Hence, a strong cleavage in society induces even moderate voters to take sides in it.

Identification is then characterized as follows.

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

$$\underline{\alpha} \equiv \frac{\lambda |\varepsilon_R - \varepsilon_P| - \varepsilon_P}{\lambda |\psi_{SP} - \psi_{SC}| + \psi_{SP}} < \bar{\alpha} \equiv \frac{\lambda |\varepsilon_R - \varepsilon_P| + \varepsilon_R}{\lambda |\psi_{SP} - \psi_{SC}| - \psi_{SC}}, \quad (17)$$

and a threshold $\infty > \hat{\lambda} > \min(\pi_R, \pi_{SP})$ such that:

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

ii) If $\lambda < \min(\pi_R, \pi_{SP})$, 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 $\hat{\varepsilon} = 0$ identify with their cultural group. If $\alpha > \bar{\alpha}$, types with high or low values of ψ identify with their cultural group, while some types with ψ close to $\hat{\psi} = 0$ 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 $\bar{\alpha}$ 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, if economic or social change renders new cleavages sufficiently important, the dimension of identification changes. The extent of these changes depends on λ . When intergroup conflict is a key driver of identification, $\lambda > \hat{\lambda}$, everybody 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 ($\lambda < \min(\pi_R, \pi_{SP})$), then voters with moderate views on the most conflictual issue do not take sides in it and maintain their original identification.

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

Figure 2

Panel A reports the case $\alpha \leq \underline{\alpha}$ in which the area of class identification is larger. Panel B reports the case $\alpha \geq \bar{\alpha}$ 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.¹³

4.2 Beliefs

For a given social identity, beliefs overweight the stereotypical features of the group with which the individual identifies. The only difference with the analysis of Section 3 is that now identification influences beliefs in two dimensions: income prospects as well as civil rights. As a result, when economic change triggers a switch in identity, beliefs in both dimensions are affected.¹⁴

To see how this works, consider a poor and socially conservative voter, $\varepsilon < 0$ $\psi < 0$, who identifies based on income, and hence with the poor. Then, by equation (7) his beliefs of are:

$$\varepsilon_P^\theta = \varepsilon + \theta(\varepsilon_P - \varepsilon_R) < \varepsilon, \quad \psi_P^\theta = \psi \quad (18)$$

As in the one dimensional model, the voter exaggerates downward income risk, the larger

¹³To determine how many voters identify along income or culture, we need an assumption on the distribution of (ε, ψ) . The appendix shows that, if both variables have a piecewise uniform density like in Figure 3 below, then the majority of voters is identified along income for $\alpha \leq \underline{\alpha}$ and along culture for $\alpha \geq \bar{\alpha}$. Thus, even if some people stick to their original identities, a sufficiently large increase in the importance of cultural conflict α turns the identity of the majority of voters from income to culture.

¹⁴Formally, the stereotypical income $\tilde{\varepsilon}$ and culture $\tilde{\psi}$ of group G_d are those scoring high in the likelihood ratios $\frac{f(\tilde{\varepsilon}|\varepsilon_{G_d})}{f(\tilde{\varepsilon}|\varepsilon_{\bar{G}_d})}$ and $\frac{z(\tilde{\psi}|\psi_{G_d})}{z(\tilde{\psi}|\psi_{\bar{G}_d})}$, where \bar{G}_d is again the opposing social group. A type (ε, ψ) identifying with G_d overweighs income levels or cultural views that score high in these likelihood ratios.

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 between ε and ψ , 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. the appendix illustrates what happens when ψ and ε are correlated.

Suppose then that new social demands (e.g. the right to abortion) or an inflow of immigrants increase cultural conflict, so that civil rights or immigration policy becomes more important, namely $\alpha > \bar{\alpha}$. 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:

$$\varepsilon_P^\theta = \varepsilon, \quad \psi_P^\theta = \psi + \theta(\psi_{SC} - \psi_{SP}) < \psi. \quad (19)$$

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.

Proposition 3 *i) If 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 of an individual changes. His views on civil rights become more extreme, those on redistribution more moderate. If many voters switch, beliefs over civil rights become more polarized, while polarization in attitudes towards redistribution dampens.

4.3 Equilibrium Policy

We now study the last step of the causal link, namely how identity and belief distortions feed into equilibrium policy. To do so, denote by $W^{\varepsilon\psi\theta}(\tau, q|G)$ the expected utility of type (ε, ψ) when he identifies with group G , for $G = P, R, SC, SP$ so that his beliefs about income and culture are those distorted by stereotypes, ε_G^θ and ψ_G^θ .¹⁵

Let $\phi_G(\alpha)$ denote the fraction of the population that identifies with group G . Then the equilibrium policy satisfies:

¹⁵To obtain $W^{\varepsilon\psi\theta}(\tau, q|G)$, just replace in (12) the true ε and ψ with their stereotyped versions.

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

The first order conditions of the problem imply:

$$\tau_\alpha^* = \tau^\circ + \frac{\theta(\varepsilon_R - \varepsilon_P)[\phi_P(\alpha) - \phi_R(\alpha)]}{\nu \varphi}, \quad (21)$$

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

where τ° and q° denote the socially optimal policies.

The presence of stereotypes, $\theta > 0$, distorts equilibrium policy. Since $(\varepsilon_R - \varepsilon_P) > 0$ and $(\psi_{SP} - \psi_{SC}) > 0$, equilibrium taxation is excessive relative to the social optimum if those identified with the poor outnumber those identified with the rich (if $\phi_P(\alpha) - \phi_R(\alpha) > 0$), while the opposite is true if $\phi_P(\alpha) - \phi_R(\alpha) < 0$. Similarly, civil rights are too restricted if identification with the social conservatives exceeds that with the socially progressive (if $\phi_{SC}(\alpha) - \phi_{SP}(\alpha) > 0$), and viceversa if $\phi_{SC}(\alpha) - \phi_{SP}(\alpha) < 0$. Crucially, equilibrium policy distortions depend on the parameter α , because the latter determines the most relevant political conflict and the prevalent identities.

To compute the population shares identified with different groups, $\phi_G(\alpha)$, we need to specify 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

This distribution has mean zero, and $\bar{x} > \underline{x} > 0$ ensures that it is left skewed, so that poor and socially conservative voters are in majority. The appendix then proves that, for any $\alpha \leq \underline{\alpha}$ and $\alpha \geq \bar{\alpha}$, we have $\phi_P(\alpha) \geq \phi_R(\alpha)$ and $\phi_{SC}(\alpha) \geq \phi_{SP}(\alpha)$, with strict inequality if at least some individuals identify along that dimension. Thus, if some individuals identify along income, then the poor-identified always outnumber the rich-identified. And similarly, if there is some culture-based identification, then conservative identification dominates progressive identification.

These results are important to characterize equilibrium policies:

Proposition 4 *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 $\lambda > \hat{\lambda}$, all voters identify along either income or culture. Thus, only taxation is distorted if $\alpha < \underline{\alpha}$, while only civil rights are distorted if $\alpha > \bar{\alpha}$.

ii) If $\lambda < \min(\pi_R, \pi_{SP})$, 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 $\alpha > \bar{\alpha}$ 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 or heterogeneous.

Crucially, by influencing beliefs, changes in identity amplify the effects of economic shocks, according to the following chain reaction:

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

Suppose that income based identification was initially prevalent. As the importance of cultural conflict α rises, for a while nothing happens. But if α keeps rising, once it reaches the critical thresholds $\underline{\alpha}$ and $\bar{\alpha}$, social identities switch and become based on culture. At that point, beliefs, political outcomes and policies react accordingly. Given our assumptions on the distribution of ε and ψ , redistribution and its distortions fall, while distortions from too limited civil rights increase. Thus, gradual but sustained changes in the social and economic environment induce sudden and large shifts in political demands by voters and in policy outcomes. In the Appendix we show that similar results obtain if ε and ψ are correlated.

4.4 Discussion

This mechanism can shed light on important transformations that have taken place in the US and other Western democracies in recent years. The parameter α increases if cultural issues become more important in individual welfare evaluations (κ rises), and if taxes become more distorting (φ rises). Anecdotal evidence and survey data strongly suggest that both of these changes have indeed occurred, possibly on a large scale. Increased mobility of capital and the difficulties of taxing profits in the new digital sectors have likely exacerbated the distorting effects of capital taxation, making redistributive taxation a more moot policy instrument. Modern societies have also become much more liberal and more permissive than they were in the 1950s and 1960s. At the political level, these changes were induced by the increased number and influence of educated and progressive voters, but they did not have unanimous support (cf. Fukuyama 2018, Goodhart 2017). More traditional social strata did not alter their value system to the same extent, and they resent the new status quo. According to our

theory, at some point these transformations trigger identity switches, exacerbating conflict over immigration and civil rights, while income related redistributive conflict is dampened. In section 6 we described survey evidence showing that indeed cultural issues related to race, immigration and civil rights are now perceived as more important than in the past. We also show that, as predicted by the theory, social classes have converged in their views over economic policy, and now diverge on social and cultural issues and on immigration. Similarly, views on immigration have become more extreme, while this did not happen for views on economic policy.

The evidence in Alesina et al. (2019) and Alesina et al. (2018a) is also consistent with our theoretical results. Using survey data, these papers show that respondents are less in favor of redistribution if they are exposed to larger inflows of immigrants, or if they are primed to think about immigrants. Similarly, Tabellini (2018) shows that US cities that received more immigrants between 1910 and 1930 became politically more conservative and anti-immigrants, and reduced local government spending. An interpretation of these findings is that individuals are averse to redistributing towards strangers. But an equally plausible mechanism is that these treatments increase the salience of cultural identities, relative to income based identification. Socially conservative and low-middle class voters become more focused on their native identities, and this makes them discount the benefit from redistribution because they no longer perceive themselves as typical members of low-income groups.

In our model, politicians only care about winning the election and announce identical platforms. In a richer setting with partisan politicians who also care about policy outcomes (as in Persson and Tabellini 2000, chp 5), identity switches have additional implications. Suppose in particular that there are two citizen-candidate politicians: the left wing candidate is poor and socially progressive (i.e. he is located in the upper left corner of Figure 1), while the right wing candidate is rich and socially conservative (he is in the lower right corner). If prevalent identification is on income, as in panel A of Figure 2, then the left wing candidate predominantly draws the support of poor-identified voters, while the right wing candidate is predominantly voted by those identified as rich. Now suppose that candidate positions remain fixed, but the importance of cultural conflict α rises to the point where the prevalent dimension of identification switches to culture, as in panel B of Figure 2. Social groups, then, reshuffle. Cultural progressive and rich voters are attracted by the left wing candidate, while poor but socially conservative voters switch to the right wing candidate. Rational voters would react similarly to a rise in α (or to shifting candidate positions). But endogenous identity and stereotyping amplify the effect of a rise in α and make it more sudden. As we show in section 6 below, this sorting of voters across parties according to their cultural views is exactly what has taken place in the US and other countries - see also Piketty (2018).

5 Nationalism and 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. In the next subsection we write down a simple model where import tariffs are an additional policy instrument. Increased exposure to foreign competition gives rise to a new trade-related cleavage. This, in turn, can trigger a change in political identities. Class based identities wane and globalist-nationalistic conflict materializes, with rich and poor losers from trade openness joining the same nationalistic group.

The cleavage between opponents and supporters of globalization is related to the cultural divide. To explore this connection, in subsection 5.2 we study a three dimensional model where the government sets three policy instruments: redistributive taxes, tariffs and a civil rights policy. Here we allow individual traits to be correlated across dimensions, with socially conservative types more frequent amongst opponents of free trade. The main insight is that the effects of trade shocks are also felt in the correlated dimension. Those negatively affected by globalization become protectionist *and* more socially conservative, while the opposite happens to those that benefit from it.

5.1 Trade Protection and Redistribution

Consider a small open economy with a continuum of individuals of size 1 and two traded goods: x is the exported good, and m is the imported good. The preferences of a generic voter of type i (to be defined below) are:

$$w^i = x^i + U(m^i) + \nu g, \quad (23)$$

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

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

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 $\tilde{\varepsilon}$ 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 $\sigma(1 - \tilde{\eta})$, where $\sigma \in (0, 1)$ captures generalized exposure to import competition, while $\tilde{\eta}$ is a random variable with individual specific mean η . An individual with higher η is more employable in the export sector and hence less exposed to import competition, perhaps because he is more mobile or skilled. As we shall see, higher η thus corresponds to higher benefits from trade openness.¹⁶

Government policy consists of two policy instruments. The first is the distorting income tax τ of the previous subsections. The second instrument 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. Revenue from both instruments is used to finance a public good g .

A voter's type is summarized by the income-trade openness vector (ε, η) , and his expected income is:

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

The first term is taxable income net of taxes. The remaining terms denote expected income from the specific input. Thus, the income tax τ redistributes between individuals with high vs low expected taxable income ε , while the tariff redistributes between individuals with high vs low benefits from trade openness, η .

The appendix solves for the economic equilibrium, assuming that the utility function $U(m)$ from the imported good is quadratic and that tax distortions on τ take the same quadratic form as in the previous sections. Since individuals differ only in their expected income and they are risk neutral, their indirect utility function is separable in ε and η . Thus, the tax rate rationally preferred by type ε , τ^ε , is still given by (4) in the previous section, while the rationally preferred tariff only depends on type η , t^η . The appendix shows that an individual more exposed to import competition (with low η) demands a higher tariff; and if he has a higher taxable income (ε is higher) he demands a lower income tax, as in the previous sections.

In this model with two policy instruments, the distance function $\Delta^{\varepsilon\eta}(G)$ that determines group identification can still be approximated by equation (16) above, except that η replaces ψ and the relative weight α is replaced by a new parameter γ that captures the importance of trade policy relative to income taxes. Thus, a higher value of γ makes identification on the nationalism/globalism dimension more likely, and the Appendix shows that this happens if: 1) individuals are more exposed as producers to the import competing sector (σ increases); 2) income taxes are more distorting (φ increases), which reduces conflict over income taxation as above; 3) the elasticity of import demand drops, which makes tariffs less distorting, increasing conflict over trade policy.

¹⁶To rely on normality, which yields the exact distortion in beliefs of Equation (7), we assume that the variance is sufficiently low that $\sigma(1 - \tilde{\eta})$ is between zero and one with very large probability. We could also specify a bounded distribution for $\tilde{\eta}$, such as a beta, and obtain very similar results.

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 protectionist on trade.

5.2 Three Dimensional Conflict, Shocks and Beliefs

We can add cultural conflict to the model, by simply adding the loss function $-\frac{\kappa}{2}(q - \tilde{\psi})^2$ to the utility 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 (16), in which distances along the openness trait η are weighted by parameter γ defined above, 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. More distorting taxes (higher φ) reduce the relevance of redistributive conflicts, relative to both cultural and trade conflicts - i.e., both α and γ rise.

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 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. This greatly simplifies the analysis.

Second, we assume that $(\varepsilon, \psi, \eta)$ is normally distributed around the zero mean $(0, 0, 0)$ with unitary variances and the following correlation structure: income (ε) is uncorrelated with

the other traits, while social progressiveness (ψ) and openness to trade (η) are positively associated, with correlation coefficient $\rho > 0$. This is an admittedly extreme assumption, but it captures the idea that opposition to free trade and cultural conservativeness may reflect an underlying cultural or psychological 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. Moreover, as described in Goodhart (2017) and Fukuyama(2018), the losers from globalization are often less educated and less mobile individuals living in declining and less integrated areas, which also makes them more likely social conservatives.¹⁷

Individual traits that determine policy preferences are not unambiguously observable, but we observe the patterns of correlation in policy attitudes. Appendix Table A1 reports the correlation matrix for attitudes towards government spending, trade protection, immigration and abortion in the 2000 ANES survey. Consistently with our hypothesis, attitudes towards civil rights and trade protection are more strongly correlated with each other than with those on the size of government.

With three dimensions, a voter can identify with three out of six possible groups. Along openness to trade, a voter is either globalist, $Gl \equiv \{\eta | \eta \geq \hat{\eta}\}$, or nationalist $N \equiv \{\eta | \eta < \hat{\eta}\}$, where $\hat{\eta}$ 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 $\hat{\varepsilon} = \hat{\eta} = \hat{\psi} = z > 0$.

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 5 *If the importance of trade and culture relative 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 .

¹⁷The results discussed below are qualitatively similar if all traits are correlated, as long as openness and social progressiveness are more positively correlated with each other than with income.

Interestingly, stronger correlation ρ between cultural and trade preferences 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 that determine correlated attitudes towards globalization and civil rights are therefore prone to become vehicles of identification and political action, replacing class based identities. To put it differently, trade and technology shocks 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 replaced 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:

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

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

$$\begin{aligned}\varepsilon_N^\theta &= \varepsilon, \\ \eta_N^\theta &= \eta + \theta(\eta_N - \eta_{Gl}) < \eta \\ \psi_N^\theta &= \psi + \theta\rho(\eta_N - \eta_{Gl}) < \psi\end{aligned}$$

The voter's beliefs about income prospects are no longer distorted, so he demands less redis-

tribution. His views become polarized towards nationalism, however, and since η and ψ are positively correlated, also towards social conservatism. 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 6 *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".

5.3 Discussion

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.

These results also contribute to explain why adverse trade shocks induce a political shift to the right. Rational voters would not react in this way: a trade shock would make voters more protectionist (typically a feature of left wing parties), but it would not reduce their demand for redistribution nor would it change their cultural attitudes or their views towards immigrants. With endogenous identities, instead, trade shocks change the entire belief system. Individuals more exposed to import competition tend to be less educated and less mobile residents of declining regions, who are typically also more socially conservative. Rising exposure to import competition eventually triggers identification with social groups defined by nationalism but also by more traditional value systems. As the traditional left-wing and class based distinctiveness of their social identities wanes, those hit by adverse trade shocks demand trade

protection and more conservative social policies, and reduce their demand for redistributive programs. As a result, they are naturally attracted by right-wing populist parties, consistently with the evidence in Autor et al. (2017) and Colantone and Stanig (2017, 2018a). Technology shocks could have similar effects.

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

6 Some Evidence

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

6.1 Evidence on the US

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

6.1.1 Cultural conflict

As discussed in subsection 4.4, there is much anecdotal evidence that the relevance of cultural conflict related to race, immigration and civil rights has risen over time. This is reflected in survey data. Figure 4 plots the percentage of respondents in a repeated Pew survey that mention different issues as one of the three most important problems facing the US, between 2001 and 2018. Employment dominated the rankings during the years of the great recession, but after 2013 race and immigration are mentioned with a higher frequency. As emphasized by Abramowitz (2019), this is likely to reflect a continuing decline of the share of white voters over the eligible population, due to immigration from Asia and Latin America and a higher

fertility of non-white.¹⁸ Similar trends are observed in several European countries.

Figure 4

Changing patterns of polarization An implication of the model in section 4 is that, as the relevance of cultural conflict increases (as α rises), more and more individuals identify based on culture, abandoning income based identification. This leads to new stereotypes, and social groups become more polarized on cultural issues, and less polarized on redistributive policies.

We do not observe social identities, but much anecdotal evidence is consistent with this prediction. Frank (2004) beautifully describes the "cultural backlash" in Kansas, a state that was firmly Democrat in the past, where socially conservative and low income voters have become obsessed with civil rights, immigration or racial issues, and have reduced their demand for redistribution. Similarly, Goodhart (2017) and Fukuyama (2018) emphasize how political conflict in the UK and the US has shifted from traditional left vs right economic conflict to a new cleavage concerning cultural issues.

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

Figure 5

A similar phenomenon is observed if we compare the average opinions of different social classes. One can think of a social class as reflecting a bundle of correlated individual traits such as income, education and occupation. Opposite social classes have converged in their

¹⁸"According to ANES data, (in 2016) nonwhites made up 39% of eligible voters under the age of 30, compared with only 17% of eligible voters over seventy" - Abramowitz (2019, chp 1).

views over government spending, and now diverge in their views on immigration. This pattern is evident in Figure A1 in the Appendix, that plots the views of white respondents belonging to the working class and to the middle/upper class on government spending and immigration. Working class individuals tend to be in favor of higher government spending and to oppose immigrants (they are in the lower-left quadrant of Figure 1), relative to members of the middle/upper class (who are in the upper-right quadrant of Figure 1). But during the last 15 years conflict between these two social groups over the size of government has dampened, while it has widened over immigration policy. This is what our theory predicts, if individuals within each social class abandon income based identification and switch to identification based on cultural preferences.

This average pattern could reflect changes in the composition of social classes in terms of education, age or other individual features. To control for possibly confounding factors, we have estimated multiple OLS regressions, one per wave, where the dependent variables are those depicted in Figure A1. Throughout we control for age, age-squared and dummy variables for gender and education. Figure 6 plots the estimated coefficients of dummy variables for whether the respondent belongs to the working class (on the left hand side of each panel) or upper-middle and upper class (on the right hand side), together with 95% confidence intervals; thus, the omitted categories are the lower-middle and middle classes. Working class individuals tend to favor higher government spending and oppose immigration, while the opposite is true for upper-middle and upper class individuals, relative to the middle classes. But the important fact is that, since 2008, social class has become less correlated with preferences on the size of government, and more strongly correlated with attitudes towards immigration.

Figure 6

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

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

Changing composition of party supporters As discussed in subsection 4.4, in a setting where parties have policy preferences, endogenous identities also have implications for how social groups sort themselves as supporters of different parties. Suppose that party positions

remain fixed, with Democrats standing for socially progressive and redistributive platforms, and Republicans standing for a small size of government and socially conservative platforms. If the dimension of identification switches from culture to income, then social progressive and rich voters are attracted to the Democratic party, while social conservative and poor voters move towards the Democrats. This is what we observe in the US.

Figure 7 illustrates this phenomenon using ANES data. The left hand panel depicts the time path of the difference in the share of self-identified Democrats in different social groups, the right hand panel does the same for Republicans. Since 1998, richer, more educated, less religious and urban respondents have increasingly been attracted towards the Democratic party, while the opposite social groups have moved towards the Republicans. In 2016, the share of Democrats is higher in the top decile of the income distribution than in the bottom decile, and it is much higher amongst urban, non religious and top-educated respondents than amongst rural, religiously observant and low-educated respondents.

Figure 7

Piketty (2018) documents that a similar phenomenon has also taken place in the U.K. and France. He also shows that these trends are long dated, and started in the 1960s, although they recently accelerated.

Switching identities are not the only mechanism that can explain the reshuffling of voters across parties. With fixed party positions, rational voters would behave in the same way if cultural conflict becomes more relevant (if α rises). Endogenous identities amplify the magnitude of these effects, however. A similar reshuffling would be observed if voters' preferences remained constant, but party positions changed, with both parties becoming more concerned with cultural issues (though in opposite ways). Glaeser and Ward (2006) and Glaser et al. (2005) argue that this also happened in the US, through the influence of party insiders. Survey data alone cannot discriminate between these alternative explanations, and they are not mutually exclusive. Nevertheless, it is not clear why party organizations changed in the same direction in several countries, while the changes in voters' preferences that we emphasized are likely to be common across countries.

Increased partisan polarization The reshuffling of voters across parties induced by a switch from income-based to cultural identification has an additional implication: supporters of opposite parties become more polarized on cultural issues, while polarization on economic issues is dampened. Do we see this in the data? Figure A2 in the appendix plots average attitudes of ANES respondents who self-identify as Democrats and Republicans on three issues: the size of government, immigration and abortion. Polarization between Democrats and Republicans increased on all issues, but much more so on immigration and abortion

than on the size of government, where polarization was already historically very high. Thus, contrary to the predictions of the theory, and unlike between social classes, we don't observe a dampening of economic conflict between supporters of different parties. Of course, this may be due to other offsetting changes over this period, such as rising economic inequality.

A large literature seeks to explain observed changes in the patterns of polarization, and we do not pretend to explain them all in this paper. Some of the most striking facts, such as the rise in conflict over immigration and the changing composition of party supporters, are consistent with the predictions of our theory. Our theory also predicts a dampening of redistributive conflict, which is apparent in the data except for the reversal in 2016.

6.1.2 Exposure to import competition

One of the main results of section 5 is that trade shocks can induce political reactions in several correlated policy dimensions. As those hurt by the trade shock switch to nationalistic identities, they demand more trade protection and less income based redistribution, but they also turn anti-immigrants and socially conservative. These results are consistent with recent evidence that increased imports from China are associated with more aversion to immigrants in Europe (Colantone and Stanig 2018) and with more political polarization and a turn to the right in the US (Autor et al. 2017). Here we further explore the effect of increased exposure to imports from China on individual attitudes on a variety of issues. Unlike in the previous subsections, we go beyond mere correlations and estimate a causal effect. We follow the same strategy as Autor et al. (2017), but rather than focusing on political polarization and an aggregate measures of conservative ideology, as they do, we study the effect on particular attitudes on which our theory has specific predictions.¹⁹

We exploit the CCES survey, a large scale survey conducted in the US over the internet between 2006 and 2016. The advantage over ANES data is that here about 36 000 individuals were interviewed every year on average. For a subset of 8300 respondents there is also a panel dimension over the period 2010-14. Import exposure varies at the commuter zone (CZ) level, and we have about 60 respondents per CZ and year on average (15 in the panel analysis). We focus on four indicators: whether the respondent prefers to cut domestic spending or raise taxes in order to reduce the budget deficit, aversion to immigrants (measured by the first principal component of 2 variables on immigrants), how important is the issue of abortion (not available in the panel), and how the respondent voted in the closest election.²⁰

¹⁹The main outcome variables in Autor et al. (2017) are different voting variables (in the elections and in the US Congress). They also look at indicators of conservative ideology from survey data, obtained by aggregating 10 questions in the Pew (2014) survey; the Pew survey lacks the panel dimension and has about 25 respondents per CZ.

²⁰The 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

We estimate two sets of regressions. First, a two-period repeated cross section, where individual attitudes are sampled in the first (2006-8) and last (2016) year of our sample - the initial year varies by question. Second, a panel regression in first difference, between the first (2010) and last (2014) years of the panel. Our treatment variable is the change in US imports from China, that we instrument with the change in European imports from China, as in Autor et al. (2017). US imports from China grew particularly fast before the start of our sample period. In the repeated cross section we thus take the change in imports exposure between the year 2000 and the last year of our sample period (2016). This amounts to assuming that the full effect of increased imports on identity is not instantaneous, but entails some lag. Since the panel starts four years later than the cross section, to preserve symmetry with the cross sectional regressions, in the panel we take the change in imports between 2004 and 2014.

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

Table 1

Residents of CZ more exposed to an acceleration of imports from China became more willing to accept cuts in domestic public spending, as predicted by the theory if relatively poor respondents abandoned class identification. They also became more averse to immigration and more attentive to the issue of abortion, as predicted given that protectionist attitudes are positively correlated with sentiments against immigrants and abortion (cf. Table A.1 in the Appendix). Finally, they became more likely to vote Republican (no effect is found on votes for Democrats). These results are in line with those of Autor et al. (2017), but they clarify which attitudinal dimensions account for the shift towards a more conservative ideology.²¹

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.

²¹According to our panel estimates, an acceleration in exposure by one standard deviation increases the change in willingness to cut spending by 2%, and increases the change in aversion to immigrants by 8%, both relative to average attitudes in the first year. The change in the probability of voting Republican is 3.6 percentage points higher, if the change in exposure rises by one standard deviation. The magnitude of the estimated effects in the repeated cross sections is smaller: a one standard deviation increase in imports exposure changes attitudes by about 1.5%-2.5% of the sample mean in the first period, and increases the probability of voting Republican by about 2 percentage points.

6.2 Evidence on France

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

We consider 12 questions on three sets of policy issues: (i) the traditional left vs right economic cleavage, namely attitudes towards redistribution, and on the role of the state in protecting workers and regulating the economy; (ii) civil rights (role of women and rights of homosexual); (iii) the globalist vs nationalis cleavage, namely attitudes towards immigrants, globalization and European integration - the appendix provides more detail. On each set of issues we have extracted the first polychoric principal component, estimated the residuals by conditioning on the respondent's income and education, and the resulting variables have been normalized so that they all have the same variance. These variables thus describe individual attitudes on issues (i)-(iii), after removing the component correlated with income and education.

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

Figure 8

The criterion for assigning observations to a cluster is to minimize within group variance, in the space of the three variables described above.²² This clustering exercise thus tells us that, between 2013 and 2017, individual views on issues (i)-(iii) became more distant between opponents and supporters of international openness, and more similar in the traditional left vs right dimension of political conflict. According to our theory, this should be associated with a switch in the dimension of identification, from left / right to nationalism/ globalism. 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 9 reports the vote shares within each cluster. The two vertical clusters of

²²Clusters are defined by applying Ward's minimum variance method on the normalized residuals of the three first principal components described above. The appendix provides more details.

2013 largely correspond to how votes were split between left and right wing candidates in the 2012 Presidential election. The two horizontal clusters of 2017 instead correspond to how votes were split in 2017 between Le Pen and Macron, two candidates that are hard to pin down on the left vs right divide, but that differ instead in their positions over globalization and immigration.²³

Figure 9

The theory discussed in section 5 also predicts that identity switches lead to a change in policy preferences. Someone who identified with the right in 2012, and switched to globalism in 2017, should become more in favor of globalization and less opposed to redistribution. To the extent that progressive views on civil rights are more correlated with support for globalization than with attitudes towards redistribution, he should also become more socially progressive. Likewise, someone who abandoned his left wing identity and became a nationalist should become more opposed to globalization, more socially conservative and less in favor of redistribution.²⁴

These predictions can be tested using the panel dimension of the Dynamob survey (a subset of about 470 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. Specifically, 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.²⁵

To test these predictions, we estimate two sets of regressions on individual panel data. In the first set of regressions the dependent variable, Δy_i , is the change in attitudes towards (i) globalization and immigration and (ii) civil rights, between 2013 and 2017. Attitudes are measured by the first principal components described above, rescaled to lie between 0 and 1

²³Note that the clusters depicted in Figure 8 were identified without exploiting the voting information. The vote shares in Figure 9 do not sum to 100% because some survey respondents abstained or voted for minor candidates not classifiable on the left / right or nationalist / globalist dimensions.

²⁴Table A.2 in the appendix shows that indeed, as assumed in section 5 and as for the US, attitudes of French respondents towards civil rights are more correlated with attitudes towards globalization than towards redistribution

²⁵Both in 2012 and 2017, there was more than one candidate on the left and on the right at the first round elections (the main candidates on the left and right were Hollande and Sarkozy respectively, but there were other minor candidates). Le Pen was a candidate in both 2012 and 2017, whereas Macron only in 2017. Thus, implicitly we assume that nobody identified as a globalist in 2012, a likely exaggeration. See the appendix for how we classify presidential candidates on the left and right dimensions.

(higher values denoting a more open or socially progressive policy). Here we estimate:

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

where Gl (N) is a dummy variable for whether in 2017 individual i voted globalist (nationalist), y_{i0} denotes his initial attitudes in 2013, X is a vector of individual covariates measured in the initial period (education, gender, age, immigrant status, region and rural area) or differenced (income and employment status), u is the error term and FEs are dummy variables that have to be included to estimate the effect of acquiring a globalist or nationalist identity, 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 set of regressions, the dependent variable is the change in attitudes over redistribution, measured by the first principal component described above and rescaled to lie between 0 and 1, with higher values indicating a more right wing policy. Here the theory has predictions only for those who switched dimension of identification, from the left vs right axis to the globalist vs nationalist axis (views on redistribution should not be affected by the acquisition of a nationalist or globalist identity, for those who were not previously identified on the left / right dimension). Thus, here we estimate:

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

where again y_{i0} 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 $\beta_{LN} > 0 > \beta_{RN}$.

Table 2 reports the results, with and without the individual covariates X . As expected, those who voted Le Pen for the first time in 2017 became more opposed to international openness and more socially conservative, while those who turned to Macron in the first round of 2017 became more in favor of globalization, although here the estimated coefficient is only significant in the more parsimonious specification, and did not change their views on civil

rights. The point estimates imply that voting for Le Pen for the first time in 2017 is associated with a drop of attitudes in favor of openness by about 25% of average initial attitudes amongst non-Le Pen voters in 2012, relative to the non-Le Pen voters in 2012 who did not vote for either Le Pen or Macron in 2017. The negative effect of voting le Pen on support for civil rights is about -11% of average initial attitudes, computed in the same way.

The last two columns refer to changes in attitudes towards redistribution. Here the estimates of interest are reported in rows 4-8. Individuals who switched from the right to either Macron or Le Pen dampened their opposition to state intervention, as expected ($0 > \beta_{RGL}, \beta_{RN}$). Switching to either Macron or Le Pen from the left has no significant association with attitudes towards redistribution, once controlling for the individual covariates. The point estimates imply that a switch from right in 2012 to Le Pen in 2017 is associated with a change in attitudes towards redistribution by about -25% 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 - 15%. 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.

Table 2

Of course, these estimated coefficients only capture correlations in the data, and cannot be interpreted as causal effects of identity changes. Individuals may have changed how they voted because they rationally changed opinions, or both voting and attitudes could reflect relevant omitted variables.

7 Concluding Remarks

Identity theory provides a rich 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 key implications of this general approach, but much more remains to be done. An important 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 an 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 inexperienced 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 identification. Why is this so, and could it imply that, under some assumptions, policies are extreme in both dimensions?

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

8 References

- Abramowitz, Alan I. (2019), *"The Great Alignment - Race, Party Transformation and the Rise of Donald Trump"*, Yale University Press, New Haven
- Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson, and Brendan Price (2016), "Import Competition and the Great US Employment Sag of the 2000s", *Journal of Labor Economics*, 34(S1), S141-S198.
- Achen, Christopher H., and Larry M. Bartels (2016), *Democracy for Realists: Why Elections Do Not Produce Responsive Government*, Princeton University Press.
- Ahler, Douglas (2018) "The Group Theory of Parties: Identity Politics, Party Stereotypes , and Polarization in the 21st Century", *The Forum*, 16 (1): 3-22.
- Akerlof, George , and A. Rachel E. Kranton (2000), "Economics and Identity", *The Quarterly Journal of Economics*, Volume 115, Issue 3, 715–753.
- Alesina, Alberto and Edward Glaeser (2005), *Fighting Poverty in the US and Europe: A World of Difference*, Oxford UP, Oxford.
- Alesina, Alberto, Elie Murard and Hillel Rapoport (2019) "Immigration and Preferences for Redistribution in Europe" Harvard, mimeo
- Alesina, Alberto, Stefanie Stantcheva, and Armando Miano (2018), "Immigration and Redistribution", *mimeo*.
- Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso (2018), "Intergenerational Mobility and Preferences for Redistribution", *mimeo*.
- Anelli, Massimo, Italo Colantone, and Piero Stanig (2018), "We Were The Robots: Automation and Voting Behavior in Western Europe", *mimeo*.
- Autor, David, David Dorn and Gordon Hanson (2013), "The China Syndrome: Local Labor Market Effects of Import Competition in the United States", *American Economic Review*, 103(6), 2121-68.
- Autor, David, David Dorn, Gordon Hanson and Kaveh Majlesi (2017), "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure", *MIT Working Paper*.
- Bénabou, Roland, and Jean Tirole (2006), Belief in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, 121(2) (2006), 699-746
- Bénabou, Roland, and Jean Tirole (2011), "Identity, Morals, and Taboos: Beliefs as Assets", *The Quarterly Journal of Economics*, Volume 126, Issue 2, 805–855.
- Bénabou, Roland, and Jean Tirole (2016), "Mindful Economics: The Production, Consumption, and Value of Beliefs", *Journal of Economic Perspectives*, 30(3), 141-164.
- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer (2016), "Stereotypes", *The Quarterly Journal of Economics*, 1753–1794.
- Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer (2018), "Beliefs about Gender", *American Economic Review*, forthcoming.

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

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

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

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

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.

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

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

Frank, Thomas (2005), *"What's the Matter with Kansas? - How Conservatives Won the Heart of America"*, Picador, New York

Fryer, Roland, and Matthew Jackson (2008), "A Categorical Model of Cognition and Biased Decision-Making", *B.E. Journal of Theoretical Economics*, 8 (2008), 1-42

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

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

Gentzkow, Matthew (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.

Glaser, Edward and Bryce Ward (2006) "Myths and Realities of American Political Geography", *Journal of Economic Perspectives*, 20(2), 119–144

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

Goodhart, David (2017), *"The Road to Somewhere - The New Tribes Shaping British Politics"*, Penguin Books

Graham, Jesse, Brian Nosek, and Jonathan Haidt (2009), "Liberals and Conservatives Rely on Different Sets of Moral Foundations," *Journal of Personality and Social Psychology*, 96 (5), 1029.

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.

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

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

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

Tajfel, Henri, 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), *Rediscovering 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

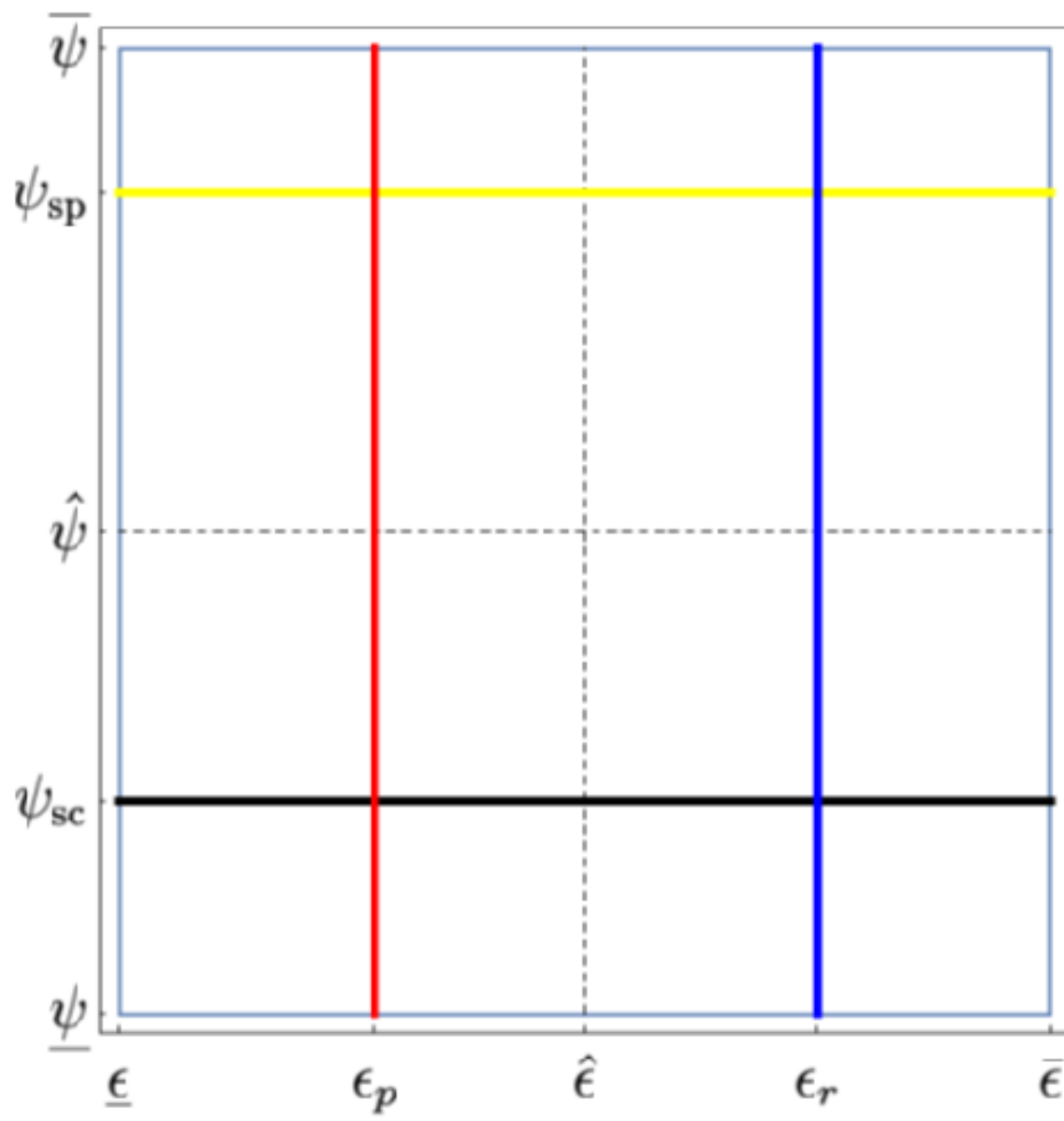
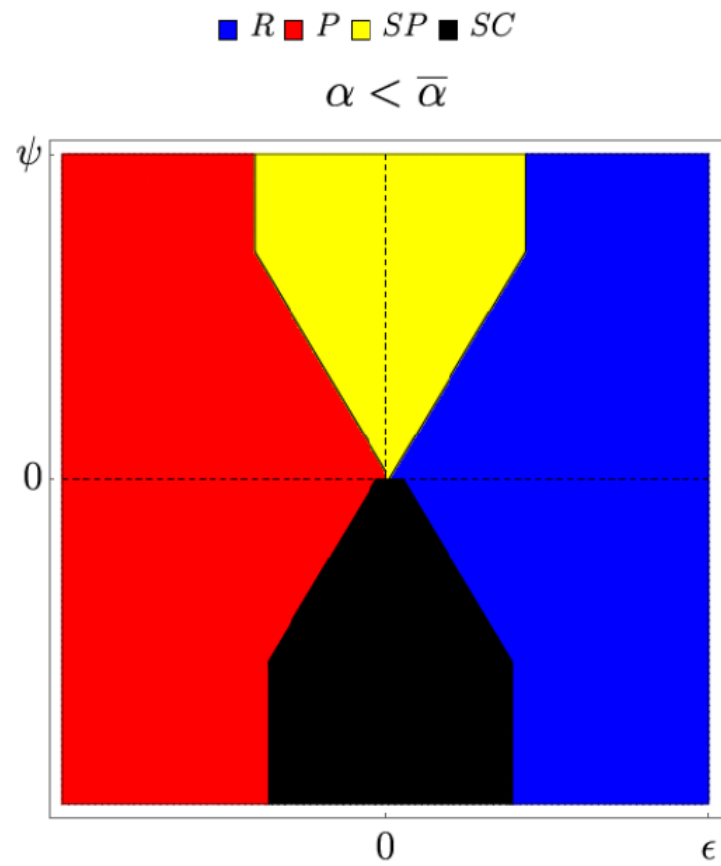
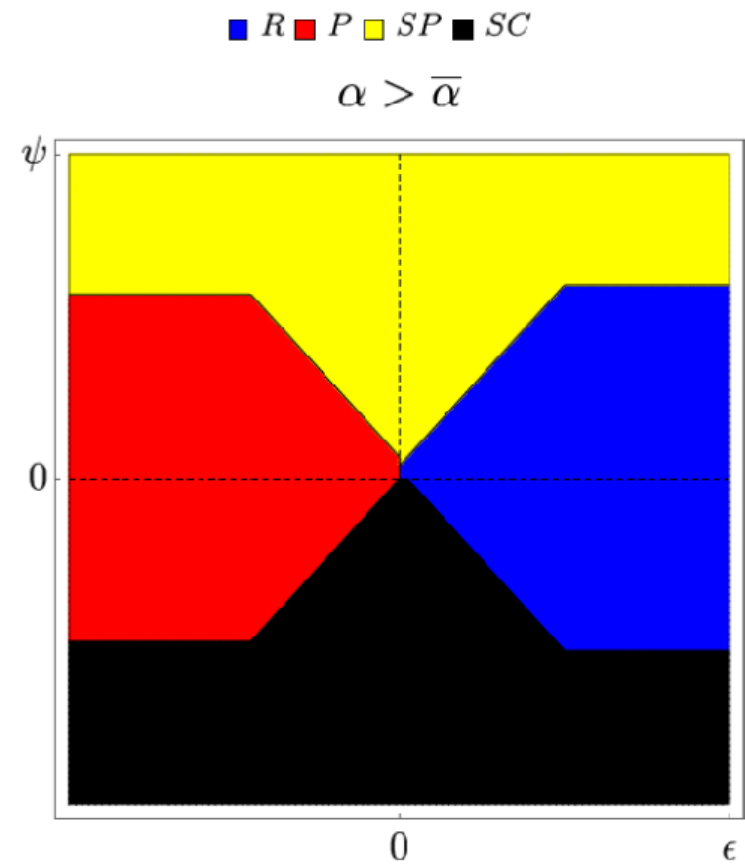


Figure 2: Mixed Identification



(a) Prevalent income-based identification



(b) Prevalent culture-based identification

Figure 3: Primitive Income Groups

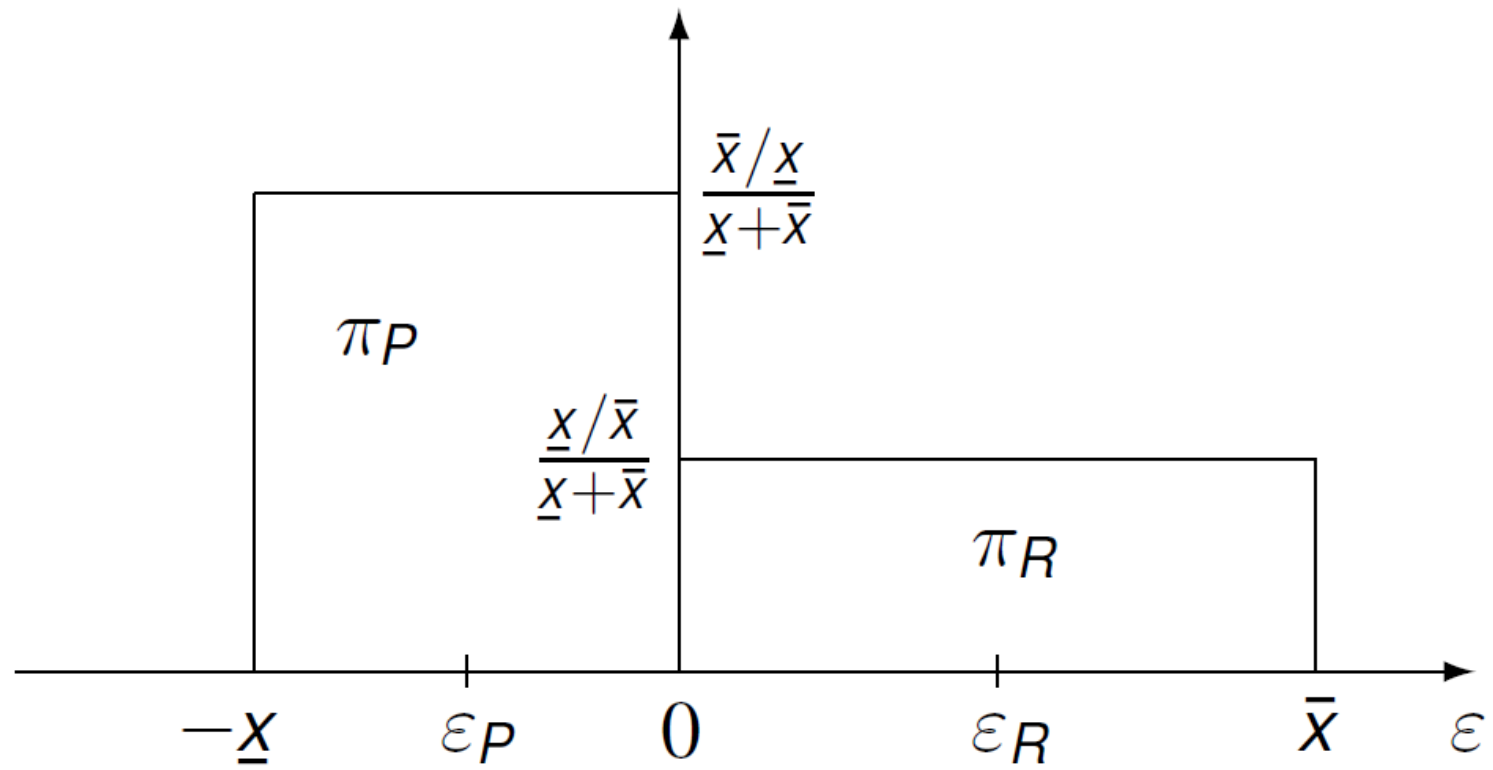
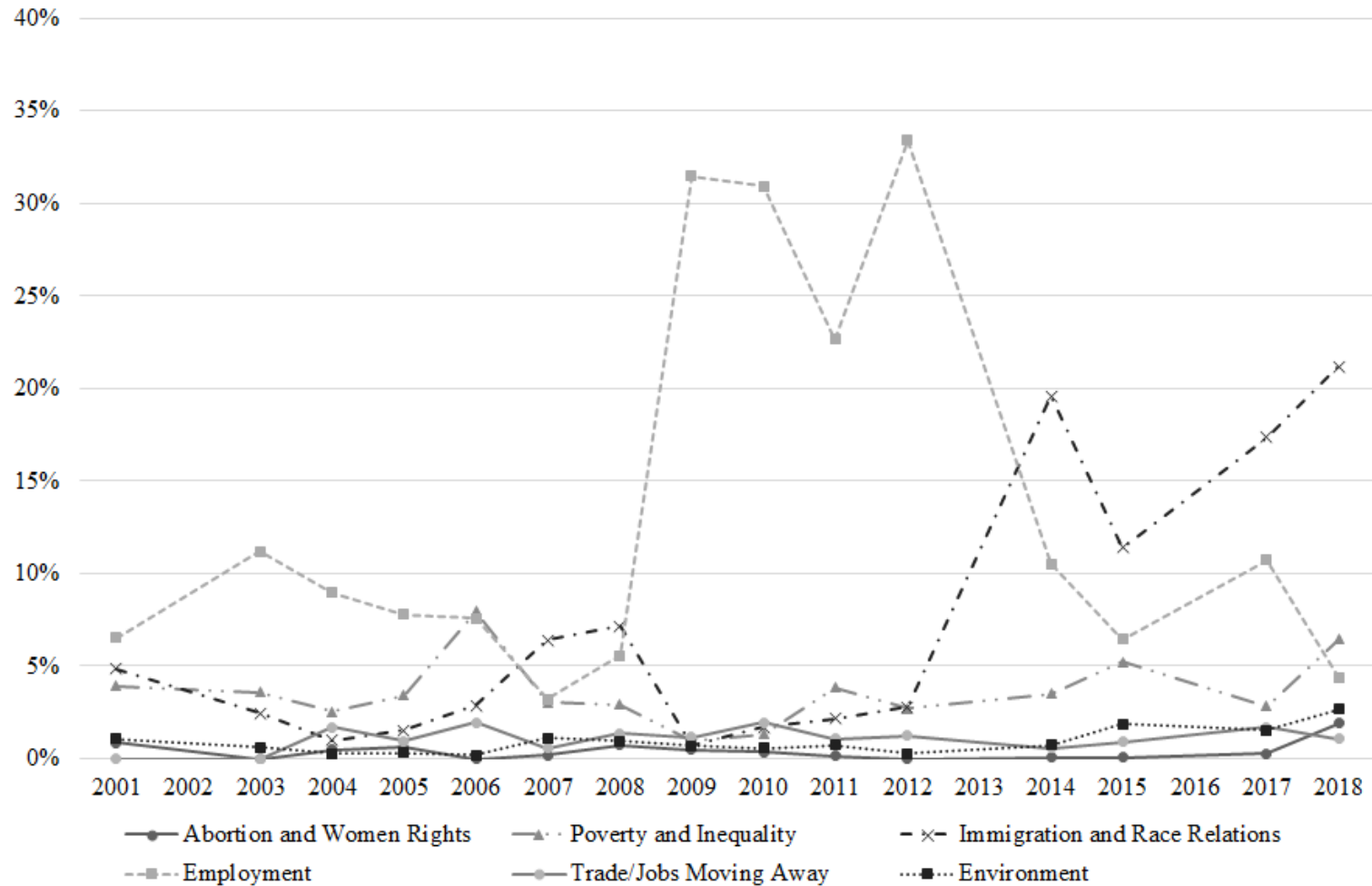
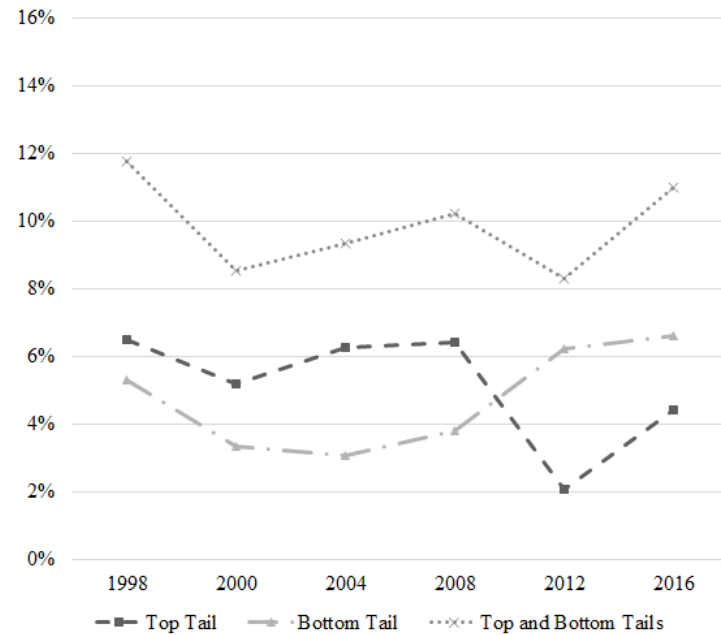


Figure 4: Most Important Problem Facing The Country

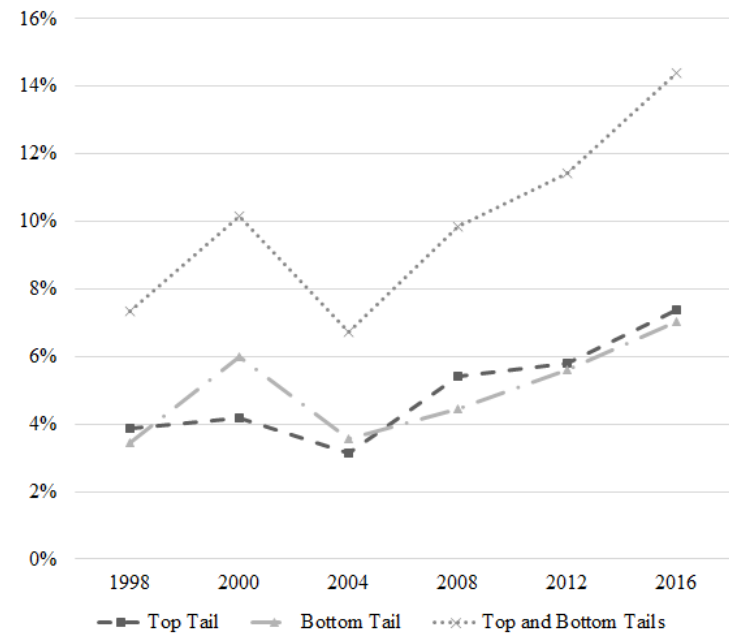


Notes: The graph shows the share of respondents mentioning the selected issues among the top three most important problems facing the US. Source: Pew Research Center public data.

Figure 5: Extremism in the Population



(a) Desired Size of Government



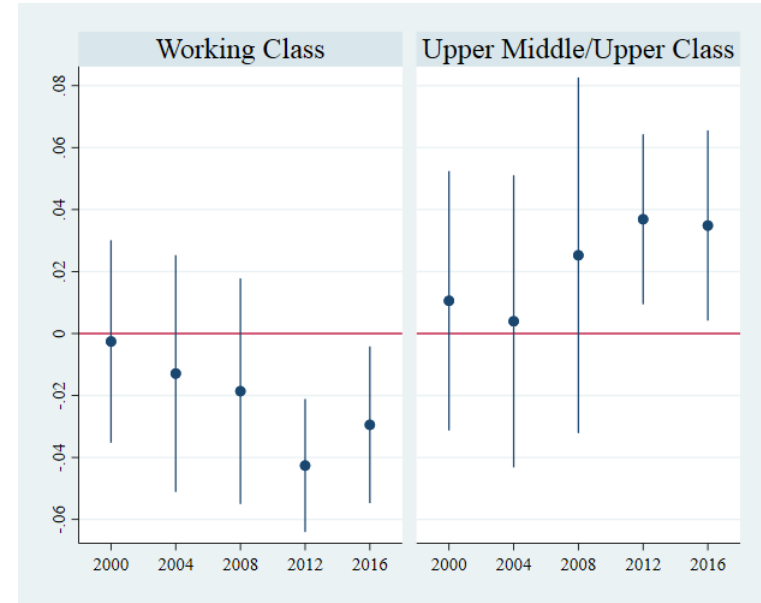
(b) Desired Immigration Levels

Notes: Desired size of government is measured as the answer to the following question: “Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel that it is important for the government to provide many more services even if it means an increase in spending. Where would you place yourself on this scale?” Answers are given on a seven-point scale, and recoded so that the variable is increasing in respondents’ desired size of government. Desired immigration levels are measured using the following question: “Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?”. The resulting five-point variable is recoded so that higher values reflect more liberal views on immigration. Both variables are rescaled between 0 and 1. The figure is constructed as follows. We have estimated the residuals of these two variables, by conditioning on income and education. For each residual we have computed the ranges corresponding to the top and bottom 5% in the pooled distribution that combines all waves between 1998 and 2016. Individuals holding extreme views are those whose estimated residuals are included in these extreme ranges. Both pictures look very similar if extremism is defined on a range corresponding to top and bottom 10% on the pooled distribution. The sample consists of all individuals aged 18 or more. Source: ANES Time Series Studies.

Figure 6: Social Class and Policy Preferences



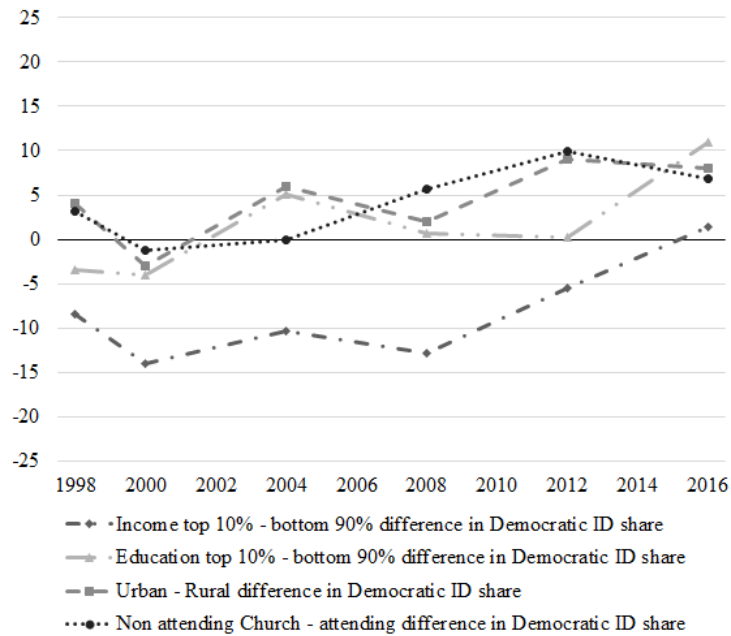
(a) Desired Size of Government



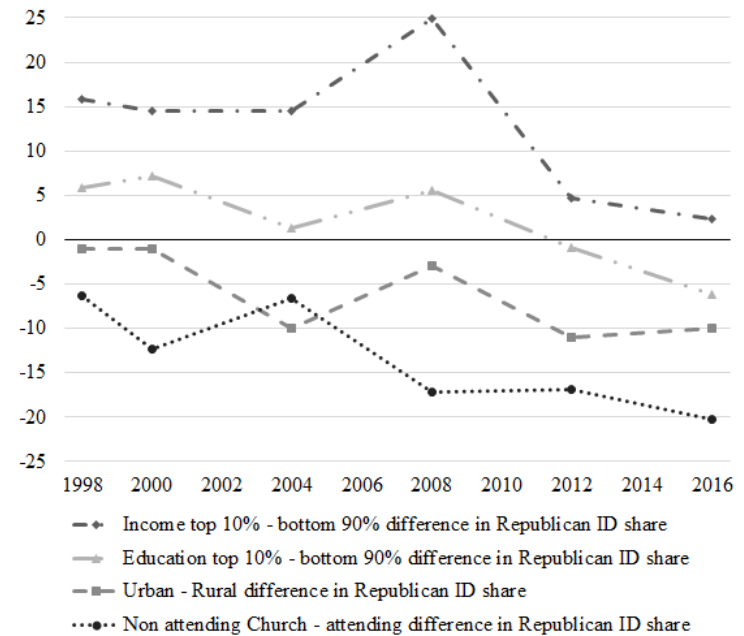
(b) Desired Immigration Levels

Notes: Desired size of government is measured as the answer to the following question: "Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel that it is important for the government to provide many more services even if it means an increase in spending. Where would you place yourself on this scale?" Answers are given on a seven-point scale, and recoded so that the variable is increasing in respondents' willingness to redistribute. Desired immigration levels are measured using the following question: "Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?". The resulting five-point variable is recoded so that higher values reflect more liberal views on immigration. Both variables are rescaled between 0 and 1. Coefficients are obtained from OLS regressions on dummy variables for whether respondents identify as working class or upper middle/upper class. The omitted category are those identifying as lower middle/middle class. Controls include age, age-squared and dummies for gender and education. 5% confidence intervals displayed. The analysis on immigration is restricted to white individuals. Source: ANES Time Series Studies.

Figure 7: Changing Composition of Party Supporters



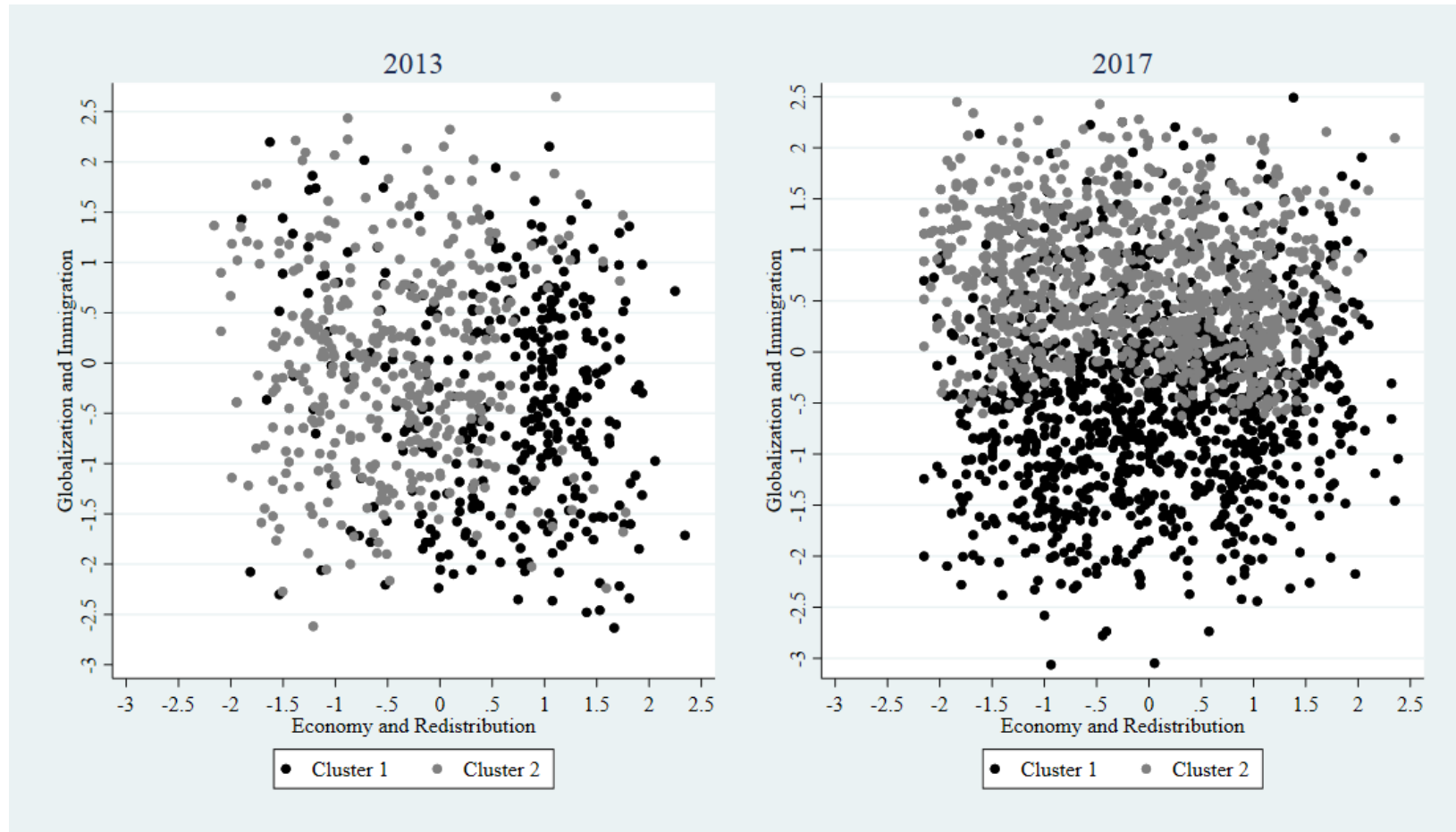
(a) Differences in Democratic share



(b) Differences in Republican share

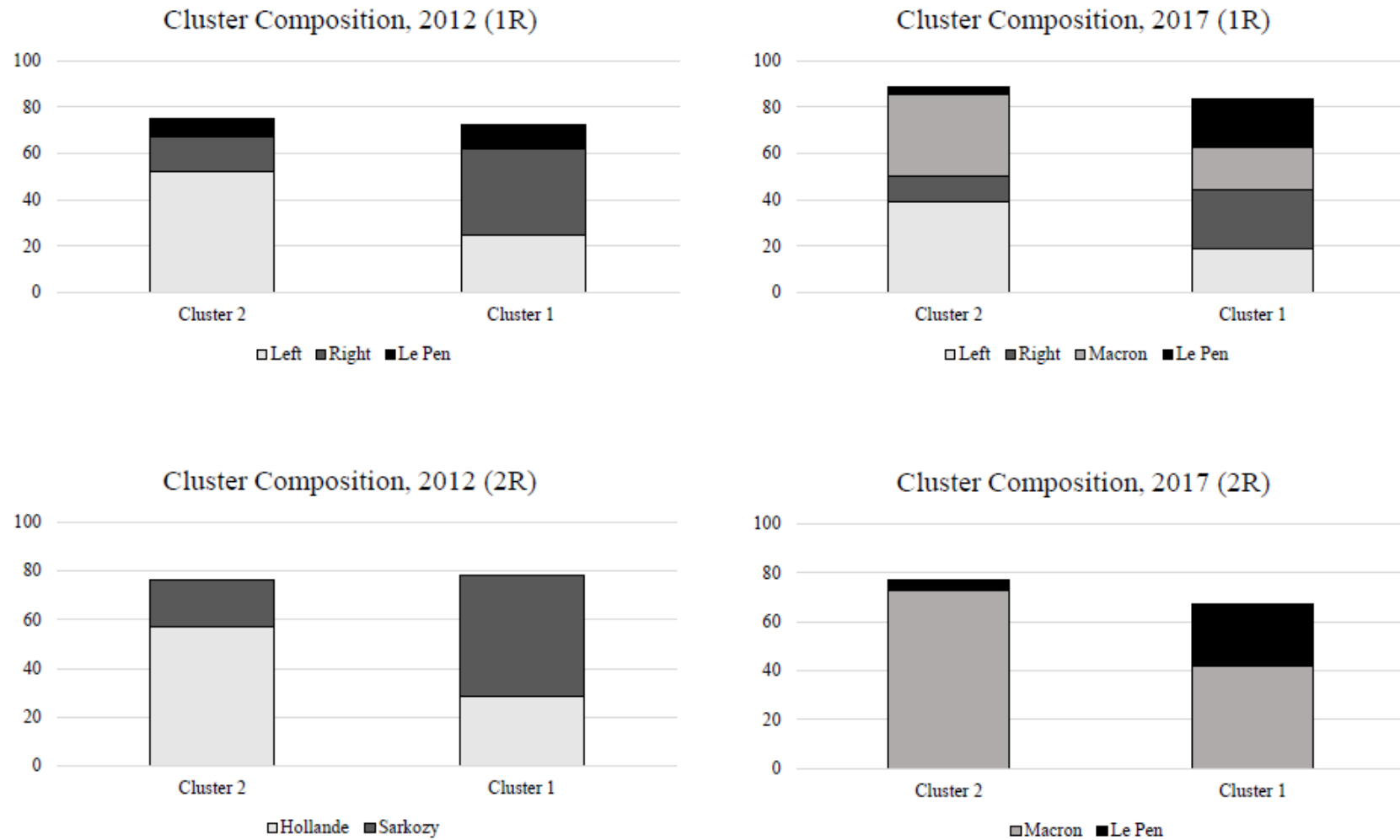
Notes: For respondents who self-identify with party P (= Democratic or Republican), the corresponding graph plots: (i) the difference between the percentage of high-income earners (top 10%) who identify with party P and the share of those in the rest of the population who identify with P; (ii) the difference in the share of party P identification between those who are among the top 10% of the population in terms of years of schooling and the rest of the population; (iii) the difference in party P identification share between those residing in rural areas and the rest of the population, and (iv) the difference in party P identification shares between those who do not and those who do attend religious ceremonies. Individuals are classified as attending religious ceremonies if they report attending church at weekly or almost weekly, while the rest of the population is considered non-attending. Source: ANES Time Series Studies.

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



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

Figure 9: Composition of Votes in the Two Clusters



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

Table 1: Exposure to Imports from China and Attitudes

	Cross Section		Panel	
	A.	Cut Domestic Spending (0-100)		
Δ CZ Exposure * Second Period	0.512 (0.363)	1.086** (0.525)		
Δ CZ Exposure			0.703 (0.938)	3.503** (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 * Second Period	0.010** (0.004)	0.014 (0.010)		
Δ CZ Exposure			0.076*** (0.023)	0.120** (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 * Second Period	0.013*** (0.005)	0.019** (0.009)		
Observations	48,871	48,871		
F	72.99	27.29		
	D.	Republican Vote		
Δ CZ Exposure * Second Period	0.0146* (0.008)	0.0197* (0.008)		
Δ CZ Exposure			0.054*** (0.0181)	0.097*** (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

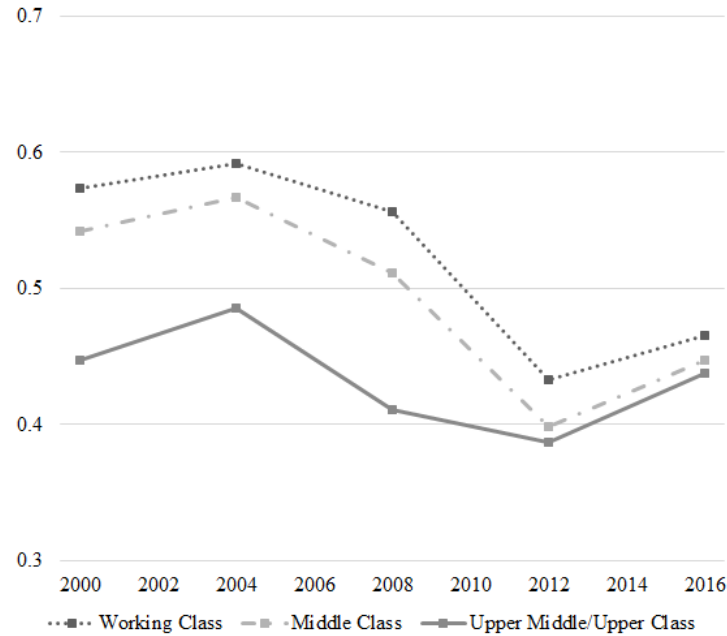
Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. For each commuter zone (CZ), the change in exposure refers to the period between 2000 and 2016 in the cross section and between 2004 and 2014 in the panel. In the cross section, the dependent variables are measured in the following pairs of years: 2006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

Table 2: Switching Identities and Attitudes

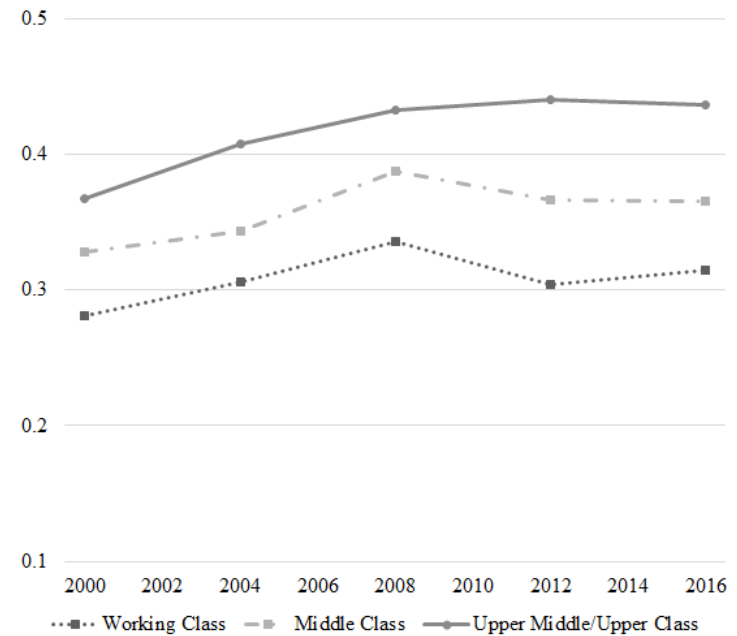
	Δ Globalization and Immigration		Δ Social Progressiveness		Δ Economy and Redistribution	
Macron 17	0.0318** (0.0147)	0.0161 (0.0153)	-0.0111 (0.0213)	-0.00919 (0.0210)	0.224*** (0.0413)	0.100** (0.0463)
Le Pen 17	-0.127*** (0.0237)	-0.121*** (0.0260)	-0.0506 (0.0322)	-0.0678** (0.0343)	-0.00376 (0.0347)	0.0232 (0.0411)
Macron 17 * Right 12					-0.217*** (0.0569)	-0.101* (0.0609)
Macron 17 * Left 12					-0.149*** (0.0521)	-0.00792 (0.0588)
Le Pen 17 * Right 12					-0.162** (0.0628)	-0.172** (0.0679)
Le Pen 17 * Left 12					0.0212 (0.0820)	-0.0111 (0.0841)
Le Pen 17 * Le Pen 12	0.0164 (0.0456)	0.0457 (0.0539)	0.0663 (0.0650)	-0.0401 (0.0733)		
Le Pen 12	-0.0402 (0.0327)	-0.0779* (0.0398)	-0.0661 (0.0456)	0.0468 (0.0534)		
Right 12					0.153*** (0.0285)	0.139*** (0.0324)
Left 12					-0.0106 (0.0289)	-0.0436 (0.0340)
Individual Controls	NO	YES	NO	YES	NO	YES
Observations	469	398	472	401	471	400
R-squared	0.294	0.333	0.309	0.371	0.292	0.272

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, employment status, NUTS 1 region and rural area. Income and employment are differenced. Estimation is by OLS. Source: Dynamob

Figure A1: Social Class and Policy Preferences



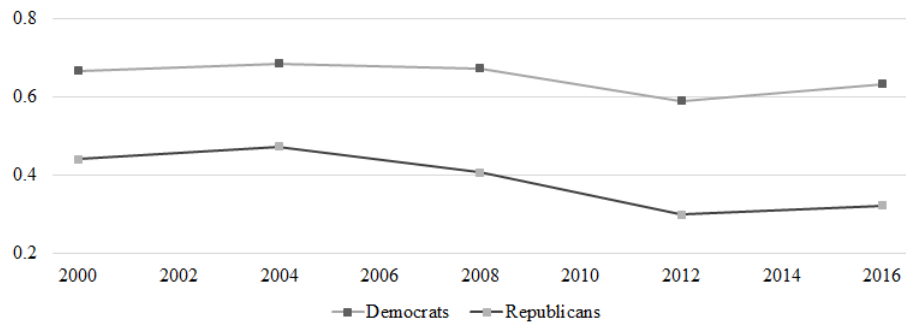
(a) Desired Size of Government



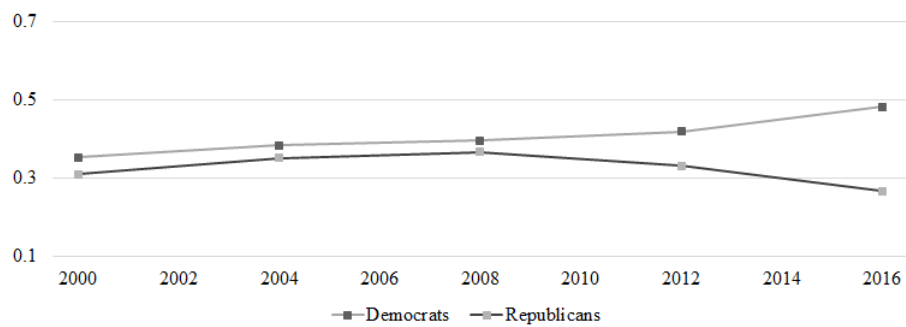
(b) Desired Immigration Levels

Notes: Desired size of government is measured as the answer to the following question: “Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel that it is important for the government to provide many more services even if it means an increase in spending. Where would you place yourself on this scale?” Answers are given on a seven-point scale, and recoded so that the variable is increasing in respondents’ desired size of government. Desired immigration levels are measured using the following question: “Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?”. The resulting five-point variable is recoded so that higher values reflect more liberal views on immigration. Both variables are rescaled between 0 and 1. The sample consists of all white individuals aged 18 or more. Source: ANES Time Series Studies.

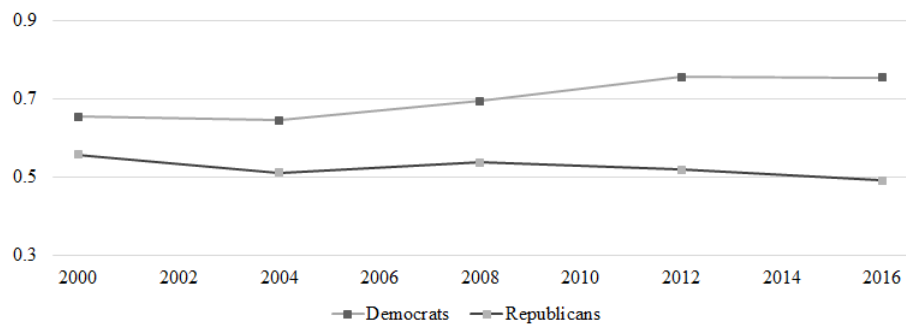
Figure A2: Partisan Polarization



(a) Desired Size of Government



(b) Desired Immigration Levels



(c) Abortion Policy Preferences

Notes: Individuals are classified as Democrats (Republicans) if they report self-identifying with the Democratic (Republican) Party. Desired size of government is proxied with the following question: “Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel that it is important for the government to provide many more services even if it means an increase in spending. Where would you place yourself on this scale?” Answers are given on a seven-point scale, and recoded so that the variable is increasing in respondents’ willingness to redistribute. Desired immigration levels are measured using the following question: “Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?”. The resulting five-point variable is recoded so that higher values reflect more liberal views on immigration. Abortion policy preferences are measured using the following question: “Which one of the opinions on this page best agrees with your view? You can just tell me the number of the opinion you choose” [1. By law, abortion should never be permitted; 2. The law should permit abortion only in case of rape, incest, or when the woman’s life is in danger; 3. The law should permit abortion for reasons other than rape, incest, or danger to the woman’s life, but only after the need for the abortion has been clearly established; 4. By law, a woman should always be able to obtain an abortion as a matter of personal choice.]. The three variables are rescaled between 0 and 1. Source: ANES Time Series Studies.

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

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

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

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

	Redistribution	Globalization	Culture
Redistribution	1		
Globalization	0.0077 0.6684	1	
Culture	0.1439 0.0000	0.3485 0.0000	1

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

9 Appendix

9.1 Theory

Proof of Lemma 1. Equation (2) implies:

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

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

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 (16) from equation (15). Define the bliss point of types ψ by $q^\psi = \psi \equiv Q(\psi)$. Recall that $\tau^G = T(\varepsilon_G)$ and $q^G = \psi^G \equiv Q(\psi^G)$, so that $W^{\varepsilon\psi}(\tau^G, q^G) = W^{\varepsilon\psi}(T(\varepsilon_G), Q(\psi_G))$. Taking a second order approximation of $W^{\varepsilon\psi}(\tau^G, q^G)$ with respect to ε_G and ψ_G at the point $\varepsilon_G = \varepsilon$ and $\psi_G = \psi$, and recalling that by (12) $W_{\tau q}^{\varepsilon\psi} = 0$ and that the optimality conditions imply $W_\tau^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) = W_q^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) = 0$, we have:

$$W^{\varepsilon\psi}(\tau^G, q^G) \simeq W^{\varepsilon\psi}(\tau^\varepsilon, q^\psi) + W_{\tau\tau}^{\varepsilon\psi}(\tau^\varepsilon, q^\psi)(T_\varepsilon)^2(\varepsilon_G - \varepsilon)^2 + W_{qq}^{\varepsilon\psi}(\tau^\varepsilon, q^\psi)(Q_\psi)^2(\psi_G - \psi)^2$$

where $T_\varepsilon = -1/\varphi v < 0$ and $Q_\psi = 1$. Taking a second order approximation of $W^G(\tau^G, q^G)$ with respect to ε_G and ψ_G at the point $\varepsilon_G = \varepsilon_G$ and $\psi_G = \psi_G$, we have:

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

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

$$\Delta_d^{\varepsilon\psi} \simeq \frac{1}{\nu\varphi} \{ [(\varepsilon_G - \varepsilon)^2 + \alpha(\psi_G - \psi)^2] - \lambda[(\varepsilon_G - \varepsilon_G)^2 + \alpha(\psi_G - \psi_G)^2] \}$$

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 (16) in the text. ■

Trade policy model

Assume that the utility function $U(m)$ takes the form:

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

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

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

so that tariff revenue is $tp^*(\hat{m} - \sigma)$.²⁶

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

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

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

Let

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

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

$$W^{\varepsilon\eta}(\tau, t) = Y^{\varepsilon\eta}(\tau, t) + S(t) + \nu G(\tau, t). \quad (24)$$

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 still given by (4) in the previous section. Exploiting the envelope theorem and simplifying, the tariff preferred by individual η is:

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

Finally, in this model we have: $W_{\tau\tau}^{\varepsilon\psi} = -\varphi\nu$ and $W_{qq}^{\varepsilon\psi} = -(2\nu-1)(p^*)^2/\delta$, and $T_\varepsilon = -1/\varphi\nu < 0$, $Q_\eta = \sigma\delta/(2\nu-1)p^*$. Inserting these expressions in the approximation above,

²⁶We implicitly restrict parameters so that in equilibrium there are indeed imports, namely $\hat{m} > \sigma$.

replacing ψ with η , we get that the parameter α is replaced by

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

which has the properties discussed in the text.

Proof of Proposition 2. We can summarize the two dimensional groups by the following parameters: $\varepsilon_R, \varepsilon_P = -\chi_\varepsilon \varepsilon_R, \psi_{SP}, \psi_{SC} = -\chi_\psi \psi_{SP}$, 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 $C_1 \equiv \{(\varepsilon, \psi) \mid \varepsilon > 0, \psi > 0\}$. A type (ε, ψ) from this set identifies along class lines if and only if:

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

The term on the right is the meta contrast ratio, which - as we saw before - becomes uniquely dispositive for $\lambda \rightarrow \infty$. When 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. $\{(\varepsilon, \psi) \in C_1 \mid \varepsilon > \varepsilon_R, \psi > \psi_{SP}\}$. For these types, class identity prevails provided:

$$\begin{aligned} \Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_R + \alpha \psi_{SP} \leq \lambda [\varepsilon_R (1 + \chi_\varepsilon) - \alpha \psi_{SP} (1 + \chi_\psi)] . \\ \Leftrightarrow \alpha &\leq \alpha_1 \equiv \frac{\varepsilon_R}{\psi_{SP}} \frac{\lambda(1 + \chi_\varepsilon) + 1}{\lambda(1 + \chi_\psi) + 1} . \end{aligned} \tag{25}$$

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

$$\begin{aligned} \Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_R - \alpha \psi_{SP} + 2\alpha \psi \leq \lambda [\varepsilon_R (1 + \chi_\varepsilon) - \alpha \psi_{SP} (1 + \chi_\psi)] . \\ \Leftrightarrow \psi &\leq \psi_1 \equiv \psi_{SP} \left[1 + \frac{\lambda(1 + \chi_\psi) + 1}{2} \frac{\alpha_1 - \alpha}{\alpha} \right] . \end{aligned} \tag{26}$$

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

$$\begin{aligned} I_{\varepsilon}^{\varepsilon\psi} &\leq I_{\psi}^{\varepsilon\psi} \Leftrightarrow \varepsilon_R + \alpha\psi_{SP} - 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \\ \Leftrightarrow \varepsilon &\geq \varepsilon_1 \equiv \varepsilon_R \left[1 - \frac{\lambda(1 + \chi_{\varepsilon}) + 1}{2} \frac{\alpha_1 - \alpha}{\alpha_1} \right]. \end{aligned} \quad (27)$$

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

$$\begin{aligned} \Delta_{\varepsilon}^{\varepsilon\psi} &\leq \Delta_{\psi}^{\varepsilon\psi} \Leftrightarrow \varepsilon_R - \alpha\psi_{SP} + 2\alpha\psi - 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \\ \Leftrightarrow \varepsilon &\geq \alpha(\psi - \psi_{SP}) + \varepsilon_1. \end{aligned} \quad (28)$$

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

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

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

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

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

$$\begin{aligned} \Delta_{\varepsilon}^{\varepsilon\psi} &\leq \Delta_{\psi}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_R + \alpha\chi_{\psi}\psi_{SC} \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \\ \Leftrightarrow \alpha &\leq \alpha_2 \equiv \frac{\varepsilon_R}{\psi_{SP}} \frac{\lambda(1 + \chi_{\varepsilon}) + 1}{\lambda(1 + \chi_{\psi}) + \chi_{\psi}}, \end{aligned} \quad (29)$$

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_R - \alpha\chi_{\psi}\psi_{SP} - 2\alpha\psi \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\
&\Leftrightarrow \psi \geq \psi_2 \equiv -\psi_{SP} \left[\chi_{\psi} + \frac{\lambda(1 + \chi_{\psi}) + \chi_{\psi} \alpha_2 - \alpha}{2} \right] .
\end{aligned} \tag{30}$$

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_R - \alpha\psi_{SC} - 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\
&\Leftrightarrow \varepsilon \geq \varepsilon_2 \equiv \varepsilon_R \left[1 - \frac{\lambda(1 + \chi_{\varepsilon}) + 1 \alpha_2 - \alpha}{2} \right] ,
\end{aligned} \tag{31}$$

which is isomorphic to (27), but ε_2 is lower than ε_1 because α_2 is larger than α_1 .

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_R + \alpha\chi_{\psi}\psi_{SP} - 2\alpha(\psi - \psi_{SC}) - 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\
&\Leftrightarrow \varepsilon \geq -\alpha(\psi - \psi_{SC}) + \varepsilon_2 .
\end{aligned} \tag{32}$$

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

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

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

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_P + \alpha\psi_{SP} \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] \\
&\Leftrightarrow \alpha \leq \alpha_3 \equiv \frac{\varepsilon_R}{\psi_{SP}} \frac{\lambda(1 + \chi_{\varepsilon}) + \chi_{\varepsilon}}{\lambda(1 + \chi_{\psi}) + 1} ,
\end{aligned} \tag{33}$$

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

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_P - \alpha\psi_{SP} + 2\alpha\psi \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\
&\Leftrightarrow \psi \leq \psi_3 \equiv \psi_{SP} \left[1 + \frac{\lambda(1 + \chi_{\psi}) + 1}{2} \frac{\alpha_3 - \alpha}{\alpha} \right] .
\end{aligned} \tag{34}$$

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] - \alpha\psi_{SP} - \chi_{\varepsilon}\varepsilon_R \\
&\Leftrightarrow \varepsilon \leq \varepsilon_3 \equiv -\varepsilon_R \left[\chi_{\varepsilon} - \frac{\lambda(1 + \chi_{\varepsilon}) + \chi_{\varepsilon}}{2} \frac{\alpha_3 - \alpha}{\alpha_3} \right] .
\end{aligned} \tag{35}$$

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_P - \alpha\psi_{SP} + 2\alpha\psi + 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\
&\Leftrightarrow \varepsilon \leq -\alpha(\psi - \psi_{SP}) + \varepsilon_3 .
\end{aligned} \tag{36}$$

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

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

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow \varepsilon_P - \alpha\psi_{SC} \leq \lambda [\varepsilon_R(1 + \chi_{\varepsilon}) - \alpha\psi_{SP}(1 + \chi_{\psi})] . \\
&\Leftrightarrow \alpha \leq \alpha_4 \equiv \frac{\varepsilon_R}{\psi_{SP}} \frac{\lambda(1 + \chi_{\varepsilon}) + \chi_{\varepsilon}}{\lambda(1 + \chi_{\psi}) + \chi_{\psi}} ,
\end{aligned} \tag{37}$$

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

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_R - \alpha\psi_{SP} + 2\alpha\psi \leq \lambda [\varepsilon_R(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] . \\
&\Leftrightarrow \psi \geq \psi_4 \equiv -\psi_{SP} \left\{ \chi_\psi + \frac{\lambda(1 + \chi_\psi) + \chi_\psi \alpha_4 - \alpha}{2} \right\} .
\end{aligned} \tag{38}$$

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_P - \alpha\psi_{SC} + 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] . \\
&\Leftrightarrow \varepsilon \leq \varepsilon_4 \equiv -\varepsilon_R \left\{ \chi_\varepsilon - \left(\frac{\alpha_4 - \alpha}{\alpha_4} \right) \frac{\lambda(1 + \chi_\varepsilon) + \chi_\varepsilon}{2} \right\} .
\end{aligned} \tag{39}$$

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

$$\begin{aligned}
\Delta_{\tilde{\varepsilon}}^{\varepsilon\psi} &\leq \Delta_{\tilde{\psi}}^{\varepsilon\psi} \Leftrightarrow -\varepsilon_P + \alpha\psi_{SC} - 2\alpha\psi + 2\varepsilon \leq \lambda [\varepsilon_R(1 + \chi_\varepsilon) - \alpha\psi_{SP}(1 + \chi_\psi)] . \\
&\Leftrightarrow \varepsilon \leq \alpha(\psi - \psi_{SC}) + \varepsilon_4 .
\end{aligned} \tag{40}$$

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

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

$$\min(\varepsilon_1, \varepsilon_2, -\varepsilon_3, -\varepsilon_4) > 0. \tag{41}$$

By inspecting thresholds (27), (31), (35) and (39) it is immediate to see, given our assumption $\chi_\varepsilon, \chi_\psi < 1$, that condition (41) is implied by $\min(\varepsilon_2, -\varepsilon_4) > 0$, which is in turn implied by:

$$\lambda < \frac{\chi_\varepsilon}{1 + \chi_\varepsilon}. \tag{42}$$

Condition (42) is therefore sufficient to guarantee that some types feature $d = \tilde{\psi}$ when $\alpha < \underline{\alpha}$.

On the other hand, when $\alpha < \underline{\alpha}$ identification universally occurs along the class dimension,

namely $d = \tilde{\varepsilon}$ for all types provided:

$$\max(\varepsilon_1, \varepsilon_2, -\varepsilon_3, -\varepsilon_4) < 0, \quad (43)$$

which is implied by:

$$\frac{1 - \lambda(1 + \chi_\varepsilon)}{2} + \frac{\lambda(1 + \chi_\varepsilon) + 1}{2} \frac{\alpha}{\alpha_1} < 0. \quad (44)$$

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

$$-\lambda \left[(1 + \chi_\varepsilon) - \alpha \frac{\psi_{SP}}{\varepsilon_R} (1 + \chi_\psi) \right] + 1 + \alpha < 0, \quad (45)$$

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

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

$$\min(\psi_1, -\psi_2, \psi_3, -\psi_4) > 0. \quad (46)$$

By inspecting thresholds (26), (30), (34) and (38) one can see, given our assumption $\chi_\varepsilon, \chi_\psi < 1$, that condition (46) is implied by $\min(\psi_3, -\psi_4) > 0$, which is in turn implied by:

$$\lambda < \frac{\chi_\psi}{1 + \chi_\psi}. \quad (47)$$

Condition (47) is therefore sufficient to guarantee that some types feature $d = \tilde{\varepsilon}$ when $\alpha > \bar{\alpha}$.

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

$$\max(\psi_1, -\psi_2, \psi_3, -\psi_4) < 0, \quad (48)$$

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

$$\frac{1 - \lambda(1 + \chi_\psi)}{2} + \frac{\lambda(1 + \chi_\psi) + 1}{2} \frac{\alpha_1}{\alpha} < 0. \quad (49)$$

By plugging the expression for α_1 this becomes:

$$-\lambda \left[(1 + \chi_\psi) - \frac{\varepsilon_R}{\psi_{SP}} \frac{(1 + \chi_\varepsilon)}{\alpha} \right] + \frac{1}{2} \left(1 + \frac{\varepsilon_R}{\alpha \psi_{SP}} \right) < 0, \quad (50)$$

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

By conditions (42), (44), (47) and (49) we therefore conclude that across the two regimes $\alpha < \underline{\alpha}$, $\alpha > \bar{\alpha}$ identification switches from being fully $d = \tilde{\varepsilon}$ to being fully $d = \tilde{\psi}$ provided:

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

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

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

After rearranging $\underline{\alpha}$ and $\bar{\alpha}$ it is easy to see that they can be expressed as in Equation (17).

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 $h_x = \frac{\bar{x}}{\bar{x} + \underline{x}} \frac{1}{\underline{x}}$ and $l = \frac{\underline{x}}{\bar{x} + \underline{x}} \frac{1}{\underline{x}}$, so that:

$$\frac{l_x}{h_x} = \chi_x^2. \quad (51)$$

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

Consider the case $\alpha < \underline{\alpha}$, in which class identification is prevalent. We need to compute and compare population shares $(\pi_P, \pi_R, \pi_{SC}, \pi_{SP})$ 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 first 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 $\pi_R(SC) = (\bar{\varepsilon} - \varepsilon_2)(0 + \underline{\psi})l_\varepsilon h_\psi + A_R(SC)$, 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:

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

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

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

In light of the definitions of 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:

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

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

Compare now the regions $A_P(SC)$ and $A_R(SC)$. In light of the previous result, for the poor-identified SC types 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 $2A_P(SC) = \alpha(\chi_\psi\psi_{SP})^2 h_\varepsilon h_\psi > 2A_R(SC) = \alpha(\chi_\psi\psi_{SP})^2 l_\varepsilon h_\psi$. When $A_P(SC)$ is a trapezoid and $A_R(SC)$ is a triangle, we have that:

$$A_P(SC) > A_R(SC) \Leftrightarrow \left(2\chi_\psi\psi_{SP} + \frac{\varepsilon_4}{\alpha}\right)(-\varepsilon_4) > \alpha\chi_\psi^2\psi_{SP}^2\chi_\varepsilon^2. \quad (52)$$

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 (52) 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 $\pi_P(SP) = (\varepsilon_3 + \underline{\varepsilon})\bar{\psi}h_\varepsilon l_\psi + A_P(SP)$, and $\pi_R(SP) = (\bar{\varepsilon} - \varepsilon_1)\bar{\psi}l_\varepsilon l_\psi + A_R(SP)$, 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:

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

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

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

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 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 $\hat{\varepsilon}_3 = \alpha\psi_{SP} + \varepsilon_3 < 0$ and $\hat{\varepsilon}_1 = -\alpha\psi_{SP} + \varepsilon_1 > 0$ 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:

$$A_P(SP) > A_R(SP) \Leftrightarrow \left(2\psi_{SP} + \frac{\varepsilon_3}{\alpha}\right)(-\varepsilon_3) > \alpha\psi_{SP}^2 \chi_\varepsilon^2. \quad (53)$$

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 $\hat{\varepsilon}_3 > 0$ we have $\varepsilon_3/\alpha > -\psi_{SP}$, so that (53) 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 $\pi_{SP}(P) = -\varepsilon_3(\bar{\psi} - \psi_{SP})h_\varepsilon l_\psi + A_{SP}(P)$, 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:

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

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

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

which is easily verified because $\chi_\psi, \chi_\varepsilon < 1$. Compare now the regions $A_{SC}(P)$ and $A_{SP}(P)$. In light of the previous result, for the socially conservative identified P types 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 $\hat{\varepsilon}_4 = \alpha\psi_{SP} + \varepsilon_4 > 0$ and $\hat{\varepsilon}_3 = \alpha\psi_{SP} + \varepsilon_3 > 0$. 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 $\hat{\varepsilon}_4 < 0$ and $\hat{\varepsilon}_3 > 0$. In this case, we have that:

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

which also holds because $-\varepsilon_4 > \alpha\chi_\psi\psi_{SC}$ and because $\alpha\psi_{SC} > -\varepsilon_3$. This implies that $A_{SC}(P) - A_{SP}(P) > 0$ 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 $\pi_{SP}(R) = \varepsilon_1(\bar{\psi} - \psi_{SP})l_\varepsilon l_\psi + A_{SC}(R)$, 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:

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

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

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

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

Compare now the regions $A_{SC}(R)$ and $A_{SP}(R)$. In light of the previous result, for the socially conservative -identified R types 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:

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

which also holds because $\varepsilon_2 > \alpha\chi_\psi\psi_{SP}$ and because $\alpha\psi_{SP} > \varepsilon_1$. This implies that $A_{SC}(R) - A_{SP}(R) > 0$ 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 $\alpha > \bar{\alpha}$, 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:

$$\begin{aligned} |\varepsilon_P - \varepsilon_R| &> |\varepsilon_{SC} - \varepsilon_{SP}| > 0, \\ |\psi_{SC} - \psi_{SP}| &> |\psi_P - \psi_R| > 0, \end{aligned} \tag{54}$$

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 $\lambda \rightarrow \infty$, so that contrast among groups in the unique driver of identification. By Equation (16), 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, $E^\theta(\tilde{\psi}|\varepsilon, \psi, P) = \psi + \theta(\psi_P - \psi_R)$. If civil rights become more important and α rises above $\bar{\alpha}$, 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 (54) 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 $\lambda \rightarrow \infty$ everyone identifies along the same dimension. Let τ_d^*, q_d^* denote the equilibrium policies when identification is along dimension $d = \tilde{\varepsilon}, \tilde{\psi}$. Repeating the same steps as above, we get:

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

which implies the following:

Corollary 2 *Suppose that $\pi_P, \pi_{SC} > 1/2$. If condition (54) is satisfied and $\pi_P \approx \pi_{SC}$, then $\tau_{\tilde{\varepsilon}}^* > \tau_{\tilde{\psi}}^*$ and $q_{\tilde{\psi}}^* < q_{\tilde{\varepsilon}}^*$. Moreover, $\tau_{\tilde{\psi}}^* \gtrless \tau^\circ$ and $q_{\tilde{\varepsilon}}^* \lesseqgtr q^\circ$ 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:

$$\mathbb{E}^\theta(\tilde{\varepsilon} | \varepsilon, \psi, G) = \varepsilon + \theta [\mathbb{E}(\varepsilon | G) - \mathbb{E}(\varepsilon | \overline{G})], \quad (55)$$

$$\mathbb{E}^\theta(\tilde{\psi} | \varepsilon, \psi, G) = \psi + \theta [\mathbb{E}(\psi | G) - \mathbb{E}(\psi | \overline{G})]. \quad (56)$$

■

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

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

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 $\alpha \leq \bar{\alpha}$ to $\alpha \geq \bar{\alpha}$. We do so in both classes of identification equilibria. ■

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

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

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

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

$$(\pi_P - \pi_R)|_{\alpha \leq \underline{\alpha}} \geq \hat{\pi}_P - \hat{\pi}_R = [(\varepsilon_4 + \underline{\varepsilon})\underline{\psi} + (\varepsilon_3 + \underline{\varepsilon})\bar{\psi}] h_\varepsilon - [(\bar{\varepsilon} - \varepsilon_2)\underline{\psi} + (\bar{\varepsilon} - \varepsilon_1)\bar{\psi}] l_\varepsilon, \quad (57)$$

where $\hat{\pi}_P - \hat{\pi}_R$ 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 (57). After some algebra, we can show that this is equal to:

$$\hat{\pi}_P - \hat{\pi}_R = \hat{\pi}_0 + \alpha \psi_{SP}^2 h_\varepsilon (1 - \chi_\varepsilon^2) (1 - \chi_\psi^2) (1 + \lambda), \quad (58)$$

where

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

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

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

$$(\pi_P - \pi_R)|_{\alpha \geq \bar{\alpha}} \leq \pi_P^* - \pi_R^* = [-\psi_4 \underline{\varepsilon} + \psi_3 \underline{\varepsilon}] h_\varepsilon - [-\psi_2 \bar{\varepsilon} + \psi_1 \bar{\varepsilon}] l_\varepsilon, \quad (59)$$

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

We now show that, in fact, these trapezoids feature a majority of poor voters, so adding them to the count offers an upper bound of the poor-rich edge for $\alpha \geq \bar{\alpha}$. Consider first the set of socially conservative 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 preliminaries therefore establish the upper bound of Equation (59), which can be written as:

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

where

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

so that for $\alpha \geq \bar{\alpha}$ the upper bound on 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 $\hat{\pi}_0 > \pi_\infty^*$, so that across the extreme cases $\alpha = 0$ and $\alpha \rightarrow \infty$, 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:

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

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

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

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

$$K \equiv \varepsilon_R - \varepsilon_P, \quad (60)$$

where $\varepsilon_R = \mathbb{E}(\varepsilon | \varepsilon > \hat{\varepsilon})$ and $\varepsilon_P \equiv \mathbb{E}(\varepsilon | \varepsilon \leq \hat{\varepsilon})$. Income is uncorrelated with culture and trade attitudes, so (60) only involves income differences between rich and poor. ■

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

$$\gamma(\eta_{Gl} - \eta_N) + \alpha\rho(\eta_{Gl} - \eta_N) = (\gamma + \alpha\rho) K, \quad (61)$$

where γ is the importance, α that of culture (all relative to taxes). Contrast $(\eta_{Gl} - \eta_N)$ is equal to K in (60) due to the fact that the marginal distribution of η is identical to that of ε .

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

$$(\gamma\rho + \alpha) K. \quad (62)$$

Relative to class identification in (60), conflict along income diminishes, but conflict along the other dimensions widens. Relative to trade based identification in (61) conflict along trade dampens, but conflict along culture amplifies. By finding the minimum distances among Equations (60), (61), and (62) readily yields the proposition.

Proof of Proposition 7. Straightforward. ■

9.2 Evidence on the United States

Cultural Conflict

The data used to create Figure 4 are publicly available on the *Pew Research Center* website. Specifically, we use data from the following surveys: *June 2018 Political Survey*, *January 2017 Political Survey*, *December 2015 Political Survey*, *December 2014 Political Survey*, *March 2012 Political Survey*, *December 2011 Political Survey*, *February 2010 Political Survey*, *February 2009 Political and Economic Survey*, *January 2008 Political Survey*, *September 2007 Political Survey*, *January 2006 News Interest Index*, *January 2005 News Interest Index*, *July 2004 Foreign Policy and Party Images*, *April 2003 Iraq Poll*, *February 2001 News Interest Index*. All such surveys are conducted on nationally representative samples of US adults aged 18 or more, with size ranging from 1303 individuals in 2010 to 2009 individuals in 2004. Survey weights are used to enhance representativeness.

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

The rest of subsection 6.1.1 uses data from the *American National Election Studies* (ANES), waves 1998, 2000, 2004, 2008, 2012 and 2016. We use the version of the variables available in the Cumulative Dataset of December 2018, and complement such information with data from the yearly releases when required. All the surveys are representative of the national population aged 18 or more and, when computing aggregates, we use individual survey weights. Yearly sample sizes range from roughly 1200 individuals in 2004 to about 5900 individuals in 2012. Below we describe the questions used in the analysis and how the variables based on these questions are defined.

Desired Size of Government “Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel that it is important for the government to provide many more services even if it means an increase in spending. Where would you place yourself on this scale?” Answers are given on a seven-point scale, and recoded so that the variable is increasing in respondents’ desired size of government. We rescale the variable between 0 and 1.

Desired Immigration Levels “Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be [1. increased a lot; 2. increased a little; 3 left the same as it is now; 4 decreased a little; 5. decreased a lot]?” Answers are in a scale from 1 to 5, following the order in which they appear in the question. We reverse the scale and rescale the variable between 0 and 1.

Abortion Policy “There has been some discussion about abortion during recent years. Which one of the opinions on this page best agrees with your view? You can just tell me the number of the opinion you choose. [1. By law, abortion should never be permitted; 2. The law should permit abortion only in case of rape, incest, or when the woman’s life is in danger; 3. The law should permit abortion for reasons other than rape, incest, or danger to the woman’s life, but only after the need for the abortion has been clearly established; 4. By law, a woman should always be able to obtain an abortion as a matter of personal choice]”. We rescale the variable between 0 and 1.

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

Social Class "There’s been some talk these days about different social classes. Most people say they belong either to the middle class or the working class. Do you ever think of yourself as belonging in one of these classes? Which one?" We aggregate answers "Lower Class (Volunteered)", "Average Working", "Working" and "Upper Working" in the macro category "Working Class"; answers "Lower Middle" and "Average Middle" in the macro category "Middle Class"; and "Upper Middle" and "Upper (Volunteered)" in "Upper Middle/Upper Class".

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

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

Income Self-reported family income in the year before the survey. It should include income from all sources, including salaries, wages, pensions, Social Security, dividends, interest, and all other income. The original variable consists of roughly 30 income brackets, which we replace by their mid-points. The Cumulative Data File contains a further variable, which classifies observations in 7 year-specific quantiles. Such variable is used to condition on income in Figure A2.

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

Race Self-identified race. Respondents’ self-identified race is used to restrict the analysis in Figure 5 and Figure A1 to the subsample of white adults.

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

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

Figure A2 has been constructed as follows. i) We estimated the residuals of the questions on Desired Size of Government and Desired Immigration by conditioning on income and education, on the pooled data from 1998 to 2016. ii) From the distribution of the estimated

residuals(pooling all waves together), we have computed the threshold values corresponding to the top and bottom 5% of the residuals. iii) Figure A2 plots the frequency of respondents that, on each wave, fall in these extreme ranges. The figure looks very similar if in step (ii) the thresholds are defined as corresponding to the top and bottom 10%.

Exposure to Import Competition

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

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

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

$$\Delta IP_c = \sum_{m \in M} \frac{L_{m,c,00}}{L_{c,00}} \times \frac{I_{m,t_2} - I_{m,00}}{Y_{m,91} + I_{m,91} - X_{m,91}} \quad (63)$$

where the first term in summation is the share of manufacturing industry m in *total* employment of CZ c , while the second term is the increase in US imports from China of products typical of m between 2000 and year t_2 , standardized by m 's market size in 1991 (i.e, prior to the boom in China's exports). Since the change in penetration is likely to be endogenous, imports are instrumented as in Acemoglu et al (2016), in a way similar to Autor et al (2013). The instrument is obtained by replacing $(I_{m,t_2} - I_{m,00})$ with $(I_{m,t_2}^{EU} - I_{m,00}^{EU})$, namely the increase of Chinese imports in eight European countries over the same period, and all the other terms in (1) with their values in 1988.

Data on bilateral imports are downloaded from the UN Comtrade database in HS-6 product classification. In particular, we obtain data on imports from China for the US as well as

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

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

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

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

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

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

Aversion to Immigrants. We extract the first polychoric principal component from two questions: “What do you think the U.S. government should do about immigration? Grant legal status to all illegal immigrants who have held jobs and paid taxes for at least 3 years, and not been convicted of any felony crimes. [1. Yes; 2. No]” and “What do you think the

U.S. government should do about immigration? Increase the number of border patrols on the US-Mexican border. [1. Yes; 2. No]”. Migrant Aversion is the resulting first principal component, rescaled between 0 and 1. Higher values indicate more anti-migrant views.

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

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

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

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

Age Self reported age. We also include its square in order to account for non-linear relations often found when dealing with subjective dependent variables.

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

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

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

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

where i denotes the individual, c the CZ and $t = 1, 2$ the year, $y_{i,c,t}$ measures individual attitudes in year t , 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 individual covariates (gender, race, educational attainments, age and age squared), Z_c is a vector of covariates referring to the CZ in the year 2000.²⁷ 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), 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 exercise. Estimation is by 2SLS and standard errors are clustered at the CZ level.

²⁷As in Autor et al. (2017), the vector Z_c includes the manufacturing share of employment, the offshorability and routine task indexes of Autor and Dorn (2013) and the interaction between the county vote share for GW Bush in the presidential election and a dummy for Republican victory in that county, all variables measured in 2000. Inclusion of these variables is important for identification, given the nature of the instrument defined in the previous footnote. As in Autor et al. (2017), Z_c also includes the CZ average of the dependent variable measured at $t = 1$, $\bar{y}_{c,1}$, to allow for mean reversion of attitudes over time.

9.2.1 Evidence on France

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

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

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.

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

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

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

Immigration Enriching “The presence of immigrants is a source of cultural enrichment. [1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]”. We reverse the scale so that higher values indicate stronger agreement with the statement. To

increase the sample size, we add an intermediate category consisting of those who don't take a side. The resulting variable ranges between 1 and 5.

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

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

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

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

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

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

Same-Sex Adoption "It is normal that homosexuals can adopt children". [1. Strongly agree; 2. Somewhat agree; 3. Somewhat disagree; 4. Strongly disagree]". We reverse the scale so that higher values indicate stronger agreement with the statement. To increase sample size, we add an intermediate category consisting of those who don't take a side. The resulting variable ranges between 1 and 5.

Figure 8 was constructed as follows. First we estimate the residuals after conditioning the three measures described above on income and education (defined below), and standardize such residuals on the pooled sample of 2013 and 2017, so that all three variables have the

same variance. Clusters are then defined by applying Ward’s minimum variance method on the standardized residuals and stopping the hierarchical partitioning algorithm at 2 clusters. We run the algorithm separately for 2013 and 2017. Finally, the residuals of “Economy and Redistribution” and “Immigration and Globalization” are plotted in Figure 8, with different tones for the two clusters.

The three first principal components “Economy and Redistribution”, “Immigration and Globalization” and “Social Progressiveness”, are used as dependent variables in the regressions displayed in Table 2. They have been rescaled between 0 and 1 to facilitate the interpretation of regression coefficients. Below, we describe the independent variables used in the regressions.

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

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

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

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

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

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

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

Region NUTS 1 region dummy variables.

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