

Social Polarization and Partisan Voting in Representative Democracies

Dominik Duell, Justin Valasek

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl

www.cesifo-group.org/wp

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp

Social Polarization and Partisan Voting in Representative Democracies

Abstract

While scholars and pundits alike have expressed concern regarding increasing social polarization based on partisan identity, there has been little analysis of how social polarization impacts voting. In this paper, we incorporate social identity into a principal-agent model of political representation and characterize the influence of social polarization on partisan voting. We show that social identity has an indirect effect on voting through voters' beliefs regarding the ex post decision of political representatives on top of a direct effect through an expressive channel. We conduct a laboratory experiment designed to identify the relative effect of the two channels. We find that social polarization causes partisan voting, and that up to fifty-five percent of partisan voting is due to the indirect effect of social identity.

JEL-Codes: D720.

Keywords: social identity, partisan voting, social polarization, political polarization.

Dominik Duell
University of Essex
Colchester / United Kingdom
dominik.duell@essex.ac.uk

Justin Valasek
WZB Berlin
Berlin / Germany
justin.valasek@wzb.eu

Thanks to Charlotte Cavaille, Steffen Huck, Rachel Kranton, Anselm Rink, Arturas Rozenas, and Paul Seabright for their valuable comments and suggestions. Support through the ANR - Labex IAST and WZB Berlin is also gratefully acknowledged.

1 Introduction

While partisan identity has long been the strongest predictor of American political behavior, recent research indicates that partisan political identity is increasingly taking on a social dimension, resulting in an in-group/out-group mentality that is comparable in strength to racial identity (Iyengar and Westwood, 2015). This social polarization along partisan lines has caused pundits and academics alike to express concern that American politics has entered a new era of partisan tribalism. Such concerns are well-founded—a large body of literature has documented that social identity affects distributional preferences and political decisions (Charness, Rigotti and Rustichini, 2007; Chen and Li, 2009; Shayo, 2009; Klor and Shayo, 2010) and directly impacts voting behavior by creating a preference (or even norm) to vote for in-group candidates (Green, Palmquist and Schickler, 2002; Bassi, Morton and Williams, 2011). In addition to a direct effect, however, social polarization may also impact voter’s actions indirectly by influencing their beliefs regarding the actions political representatives will take once in office.

This indirect effect can be clearly illustrated using the principal-agent model of political representation: Political representatives often face a trade-off between policies that are in the best interest of the general population, and policies that favor a particular group. Accordingly, in an environment where political representatives cannot commit to decisions prior to being elected, voters must anticipate how prospective representatives choose between a utilitarian objective and group-specific interests before deciding whom to vote for. In such a situation, identity cues play an important, indirect role in the voter’s decision: since social identity has been shown to influence individuals’ distributional preferences, identity can provide voters with an informative signal of political representatives’ (candidates) *ex post* policy choices.

In this paper, we explore the impact of social polarization on partisan voting in a principal-agent model of political representation. First, we develop a formal theory that incorporates the group identity model of Charness, Rigotti and Rustichini (2007) into a simple principal-agent model of elections. In particular, we characterize the trade-off voters face between selecting political representatives based on identity and selecting based on quality. Our framework accounts for both the direct (expressive) and indirect (instrumental) impact of social identity on voter behavior, and allows us to develop a strategy to independently identify the two channels of impact.¹ Second, we conduct a laboratory experiment that mirrors our theoretical setup and find that, in an election environment with social polarization, over sixty percent of subjects vote for the in-group candidate even when the out-group candidate is of higher quality. Moreover, we demonstrate that up to fifty-five percent of in-group voting can be attributed to the indirect impact of social identity, as voters (correctly) anticipate that political representatives take policy decisions that are biased towards their respective in-groups. These findings have important

¹The existing literature describes the expressive impact of social polarization on partisan voting and partisanship-guided political participation as driven by expressive motivations through rising partisan loyalty or increased concern for in-group status and welfare, and the indirect impact as driven by instrumental motivations to support policies and ideologies shared by those voters who identify with the same party (Hamlin and Jennings, 2011; Mason, 2015; Huddy, Mason and Aarøe, 2015). We also consider the direct channel to be based on an expressive motive, since it is independent of the voter’s expected policy payoffs. The indirect channel, however, is instrumental in the sense that it impacts voter choice through their expected policy payoffs.

implications regarding the impact of social polarization on the electoral process: we show that both the direct and indirect impact of social polarization decrease electoral selection based on ability, as political representatives are more likely to be selected based on their group affiliation.

In our model, voters choose between two candidates via majority rule. Each candidate is characterized by membership in one of two identity-groups, and by a valence term (ability) that functions as a universal public good. Additionally, conditional on being elected, the candidate makes an ex post policy choice in a three-point policy space (left, center, right). Since candidates only receive benefits from holding office, their choice of policy is a function of their preferences over voters' payoffs: candidates can choose a centrist policy to maximize aggregate payoffs, or an extreme policy to favor a particular partisan group. Voters' payoffs are a function of both the ability of the winning candidate and the policy this candidate sets when in office. Importantly, voters' policy preferences are partisan, in the sense that voters' group identities are correlated with the location of their ideal point in the policy space. Therefore, while all voters prefer a candidate with higher valence *ceteris paribus*, instrumentally, they will only favor the co-partisan candidate to the extent that they expect candidates to select a partisan policy. Additionally, to account for the direct effect of social identity on voting behavior, as identified in Bassi, Morton and Williams (2011), voters may receive an expressive payoff of voting for the in-group candidate.

In addition to characterizing the relationship between social polarization and partisan voting, our model also illustrates the problem of separately identifying the direct and indirect effect: if we empirically observe partisan voting, this could be due to either the direct effect (expressive), the indirect effect (instrumental), or a combination of the two. Therefore, to identify the instrumental link we rely on our models prediction that the degree to which social polarization influences voting behavior is a function of the magnitude of polarization in voters' policy preferences. Intuitively, when policy preferences are homogeneous, voters will expect the candidates to adopt a centrist policy if elected and hence have a dominant instrumental incentive to vote for the higher-valence candidate. As policy preferences become more polarized, however, voters will expect candidates to take partisan policy positions to cater to the interests of their in-group, in which case partisan identity becomes the dominant incentive when selecting between the candidates. Since expressive motives to vote for the co-partisan candidate are not conditional on the underlying degree of polarization in voters' policy preferences, the prediction of a positive correlation between policy polarization and partisan voting provides a clear test of the indirect effect of social polarization.

While our model is a simplified setting, it captures important features of political competition that are affected by social identity and social polarization. Importantly, ex post policy discretion implies that voters face uncertainty regarding which policy the candidates will select once in office. Absent identity cues, voters expect all candidates to maximize aggregate utility and choose a centrist policy. In a setting with identity divisions, however, the group identity model predicts that candidates will favor policy positions that disproportionately benefit the in-group. This implies that voters will interpret identity cues as a signal that the co-partisan candidate will select policies that are consistent with the political values and norms of the group, and hence rationally respond to these cues by voting in a partisan manner.

In the experiment, both candidates and voters belong to one of two identity-groups. We induce social polarization either by a standard minimal-group intervention or by appealing to a pre-existing identity, both of which have been shown to result in an in-group preference in a controlled experiment (Tajfel and Billig, 1974; Goette, Huffman and Meier, 2006; Chen and Li, 2009; Landa and Duell, 2015).² Additionally, identity groups correspond to voters' ideal points in the three-point policy space; specifically, we precisely control the degree of policy preference polarization by changing the degree of correlation between group membership and ideal policy points: in the case of no policy polarization, all voters have ideal points at the center; in the case of full polarization in voters' policy preference, one group is located at the left extreme while the other is at the right extreme.

Our experiment robustly confirms the prediction of a positive correlation between policy polarization and partisan voting – whether minimal groups or natural identities – as the degree of correlation between group identity and extreme ideal policy preference increases, voters increasingly vote for their co-partisan candidate. The positive relationship between policy polarization and partisan voting allows us to conclude that in-group voting is more than just an expressive phenomena – the strategic response in voting behavior to the underlying degree of policy polarization suggests that a voter's behavior is sensitive to their beliefs regarding the degree of in-group favoritism displayed by the candidates. Moreover, our experiment shows that partisan voting is not purely driven by a subset of voters who always vote along partisan lines: only 14 percent of voters vote for their co-partisan candidate when the cost (in ability difference) is high and policy polarization is low, while 61 percent vote for their co-partisan candidate when the cost is low and policy polarization is high.

Our paper's makes several important contributions to existing literatures. Dickson and Scheve (2006) and Bassi, Morton and Williams (2011) introduce social identity into models of voting. However, both these papers consider voting over fixed policy alternatives – in contrast, we formally characterize the effect of social identity in a principle-agent model of voting. This allows for an indirect impact of social identity, as social identity impacts voter behavior by influencing their beliefs regarding the ex post actions of political representatives.

We extend the experimental literature in economics that focuses on preferences-driven effects of social identity, such as heightened concern for group welfare, norms or status (among others, Chen and Li, 2009) by documenting an indirect impact of social identity in the sense that it influences beliefs regarding others' behavior.

We also speak to the literature in political psychology and political behavior, which tries to clarify the relationship between identity-driven expressive and instrumental motivations behind vote choice. So far, it is argued that expressively motivated identity-driven behavior is more fundamental than instrumentally motivated choices and issue positions follow from social identity affiliations, e.g. partisanship or ethnicity, and not vice versa (Huddy, Mason and Aarøe, 2015;

²Throughout this paper we refer to *group identity* and *social identity* interchangeably. We acknowledge that the former only requires individuals subjective awareness of group membership but may not rise to the level of being a social identity while the latter goes beyond awareness of membership and demands the individual to attach value and emotional significance to the membership (Tajfel, 1981). A laboratory experiment like ours is certainly able to induce a group identity but may fail to instantiate the latter. Following protocols established in the literature, however, gives us reason to believe that we also induce elements associated with social identity.

Huddy, Bankert and Davies, 2018). Our theory-driven experiment allows us to separately identify the direct (expressive) and indirect (instrumental) impact of social polarization on partisan voting, and we show that both have a significant impact on voter behavior.

Finally, our study generates some novel empirical results regarding the comparison between minimal and naturally-occurring group identity. Similar to work in economics (Eckel and Grossman, 2005; Goette, Huffman and Meier, 2012), we compare the impact of a weak group identity with priming existing (political) identities within the same controlled laboratory experiment. However, we do not find large differences in distributional decisions between the minimal group and political identity treatment. Interestingly, we do find a positive correlation between the subjects' self-reported strength of political identity and their implied beliefs regarding the partisan bias of political representatives. Lastly, in contrast to most of the recent literature on social polarization, we clearly document the negative welfare impact of social polarization: it shifts political competition from ability to identity, which results in the election of political representatives with low ability.

2 A Principal-Agent Model of Elections with Social Identity

Here we introduce a simple formal structure that reflects our experimental design, and allows us to detail the interaction between social polarization and identity-contingent (partisan) political behavior. While the model informs our experimental strategy for separately identifying the direct and indirect impact of social polarization on partisan voting, readers may also skip straight to Section 2.2 for an overview of the theoretical findings.

AGENTS: There are n agents, denoted by the index set $N = \{1, \dots, n\}$, with n even and greater than two. Agents either belong to (identity) group A or group B . Abusing notation, we define group membership from the perspective of agent i when convenient; that is, i is a member of the in-group, denoted by set $I = \{j | j \in A \text{ if } i \in A \text{ else } j \in B\}$, while all other agents, j , are either in I or the out-group, denoted by set $I^- = N \setminus I$. Each identity group has an equal number of agents ($|A| = |B|$).

ACTIONS AND PAYOFFS: One agent in each group is a candidate; we denote these individuals by c^A and c^B . In addition to group membership, each candidate is endowed with a valence term, or ability, denoted by α^A and α^B . Candidates receive payoffs of x^w if they win the election and x^l if they lose the election, where $x^l < x^w$ (note that candidates only receive payoffs from holding office). After the election, the winning candidate implements a vector of policy choices, $\mathbf{p} = (p_l, p_m, p_r)$, over an ordered three-point policy space $\{l, m, r\}$. Each policy choice is represented by $p_k \in [0, 1]$ and available policy choices are constrained to the set of \mathbf{p} that satisfy $p_l + p_m + p_r \leq 1$.

Agents who are not candidates are voters and, after observing the candidates' group membership and abilities, submit a vote, v_i , for c^A or c^B (no abstention). The winner is chosen by majority rule, and the winner affects voters' payoffs through the following two channels:

1. [Policy] voter i 's payoffs are a function of the policy choice of the winning candidate, \mathbf{p}^w , and the voter's ideal point $p_i \in \{l, m, r\}$.

2. [Valence] voter i 's payoffs are strictly increasing in the winning candidate's ability, α^w . Formally, voter payoffs are as follows:

$$x_i = \alpha^w + v(\mathbf{p}^w, p_i), \quad (1)$$

where p_i is the ideal point of voter i .

Since we are considering a discrete policy space, we characterize policy payoffs a simple linear function of the ideal point and the amount allocated to each p_k :

$$v(\mathbf{p}^w, p_i) = \sum_{k=1}^k v_k(p_k, p_i), \quad (2)$$

where:

$$v_k(p_k, p_i) = \begin{cases} p_k & \text{if } k = p_i, \\ \frac{1}{2}p_k & \text{if } k, p_i \text{ contiguous,} \\ 0 & \text{if } k, p_i \text{ non-contiguous.} \end{cases}$$

That is, voters receive 1 unit of payoff for every unit of p_k placed at their ideal point, and 1/2 units of payoff for every unit of p_k placed at a point contiguous to their ideal point.

Consistent with our motivation, we are concerned with the case where policy is partisan, in the sense that policy preferences are correlated with identity. For simplicity, we consider the case in which policy preferences are stochastic: voters' ideal points are unknown prior to the election, but the distribution from which ideal points are drawn is common knowledge. Formally, for $i \in A$, p_i is drawn from $\{l, m\}$ and for $i \in B$, p_i is drawn from $\{m, r\}$. Additionally:

$$\Pr(p_i = l | i \in A) = \Pr(p_i = r | i \in B) = q. \quad (3)$$

This structure implies that each voter in group A (B) has the same *expected* policy position (ex ante symmetry). The assumption of stochastic policy preferences is not substantive with respect to the formal model; however, it simplifies the experimental analysis substantially, since group membership correlates perfectly with expected policy preferences. Moreover, q provides a measure of the polarization of voters' policy preferences, where $q = 1$ corresponds to perfect polarization in policy preferences.

UTILITY: In addition to own payoffs, we consider the case where agents have distributional preferences and may receive expressive payoffs based on their actions. However, to simplify the analysis, we assume that distributional preferences are only relevant for the candidates, and that voting for the in-group candidate is the only action that generates an expressive payoff. (Including distributional preferences for voters and an expressive payoff for the candidate does not qualitatively change our results.) Accordingly, the utility function of voters takes the following form:

$$u_i^v(x_i, v_i, e_i) = x_i + e_i \mathbb{1}(v_i = c^I), \quad (4)$$

where e_i is the expressive payoff for voting for the in-group candidate; e_i is heterogeneous and drawn from a uniform distribution, $U[0, \bar{e}]$ with $\bar{e} \geq 0$.

The utility function of the candidates is equal to:

$$u_i^c(x_i, \mathbf{x}^I, \mathbf{x}^{I^-}) = x_i + \delta g(\mathbf{x}^I, \mathbf{x}^{I^-}), \quad (5)$$

where \mathbf{x}^I represents the set of payoffs of agents in group I and $g(\mathbf{x}^I, \mathbf{x}^{I^-})$ represents the agent's distributional preferences. Note that regardless of the weight agents place on their distributional preferences (δ), $g(\mathbf{x}^I, \mathbf{x}^{I^-})$ still plays a crucial role to the equilibrium of the model: since candidates do not receive policy payoffs, their choice of \mathbf{p}^w is a function of their distributional preferences only.

Following the literature on minimal groups in social psychology, Charness, Rigotti and Rustichini (2007); Chen and Li (2009); Goette, Huffman and Meier (2012) document that even minimal group frames can significantly skew distributional preferences to favor payoffs for in-group members. As in Chen and Li, we formalize the group identity model by allowing for distributional preferences that are a function of group membership:

$$g(\mathbf{x}^I, \mathbf{x}^{I^-}) = \lambda \sum_{j \in I \setminus i} \sqrt{x_j} + (1 - \lambda) \sum_{j \in I^-} \sqrt{x_j}, \quad (6)$$

where $\lambda \in [0.5, 1]$.

However, given that our aim is to test the impact of group identity, we must also define an appropriate benchmark for comparison. A natural candidate for distributional preferences is social efficiency: as highlighted in Charness and Rabin (2002), efficiency concerns can explain many experimental data. Therefore, we also consider a benchmark case of distributional preferences for ex ante efficiency:

$$g(\mathbf{x}^I, \mathbf{x}^{I^-}) = \sum_{j \neq i} \sqrt{x_j}. \quad (7)$$

In what follows, we will characterize the predictions under both the *Identity* and *Benchmark* (efficiency) models.

TIMING: The timing of the game is as follows

1. Candidates c^A, c^B are drawn and their abilities, $\{\alpha^A, \alpha^B\}$, are publicly revealed.
2. Voters draw expressive payoffs, e_i , and simultaneously submit votes, v_i , for c^A or c^B .
3. The winning candidate (by simple majority) chooses \mathbf{p}^w .
4. Voter policy preferences, p_i , are drawn and payoffs, $\{x_i(p_i, \mathbf{p}^w, \alpha^w)\}$, realize.

EQUILIBRIUM AND WELFARE: The equilibrium concept is SPNE. That is, an equilibrium, $\{\mathbf{v}; \mathbf{p}^A, \mathbf{p}^B\}$, maximizes the candidates' distributional preferences and, given $\{\alpha^A, \alpha^B\}$ and $\{\mathbf{p}^A, \mathbf{p}^B\}$, v_i maximizes $E[u_i^v(x_i, v_i, e_i) | \alpha_w, \mathbf{p}^w]$ for each i . We impose the selection criteria that, when they are indifferent, candidates choose a centrist policy and voters vote for their co-partisan candidate; these assumptions are for convenience only, and are not substantive.

³To account for fairness considerations, we assume that distributional utility is a concave function of other's expected payoffs (see Fehr and Schmidt, 1999). The predictions of the model are qualitatively similar with linear utility.

We consider the welfare criterion of aggregate expected payoffs: i.e. the first-best solution maximizes aggregate expected payoffs, $\sum^N E[x_i]$.⁴

2.1 Analysis

We begin by characterizing the outcome that maximizes social efficiency.

Lemma 1 (Efficiency)

Social efficiency is maximized when candidates choose centrist policies, $\mathbf{p}^A = \mathbf{p}^B = \{0, 1, 0\}$, and all voters vote for the highest-valence candidate, $v_i = c^k$ for all i if and only if $\alpha_k \geq \alpha_{k'}$.

First, note that a centrist policy maximizes aggregate expected payoffs for any q , since payoffs are “linear” in policy. Second, given that both candidates choose the same policy, efficiency is maximized by selecting the candidate with the highest valence (social efficiency is always neutral with respect to candidate payoffs). Formal proofs of all results can be found in the SI.

CANDIDATES’ POLICY CHOICE: Following backward induction, we begin with the candidates choice of policy, $\{\mathbf{p}^A, \mathbf{p}^B\}$. Since candidates do not have access to commitment and do not choose policy until after they are elected, the chosen policy has no direct or indirect impact on the candidates’ payoffs. Therefore, the winning candidate will choose \mathbf{p}^w to maximize their distributional preferences.

The following propositions partially characterize the equilibrium choices of $\{\mathbf{p}^A, \mathbf{p}^B\}$ under the Benchmark and Identity models.

Lemma 2 (Policy choices: Benchmark model)

If agents’ distributional preferences are characterized by efficiency then both candidates will choose centrist policies in equilibrium, $\mathbf{p}^A = \mathbf{p}^B = \{0, 1, 0\}$.

Intuitively, Lemma 2 follows directly from Lemma 1.

Lemma 3 (Policy choices: Identity model)

If agents’ distributional preferences are characterized by group identity, then both candidates choose policies that are unique and weakly asymmetric, in the sense that $p^l \geq p^r$ for \mathbf{p}^A and $p^l \leq p^r$ for \mathbf{p}^B .

Lemma 3 stems from the inter-group conflict over the partisan policy space: Under the group identity model, candidates put a higher weight on the payoffs of their group members, and hence will take policy decisions that favor the partisan position of their group. That is, while under the Benchmark model the candidate’s group identity is irrelevant and the only distinguishing characteristic is their relative valence, under the Identity model group identity is an important predictor of the decisions the candidates will take when in office.

The next result details the comparative statics of the candidates’ policy positions in the Identity model, and will be key to our strategy for identifying the instrumental impact of affective polarization. However, instead of detailing the specific policy positions of the candidates, it will

⁴For simplicity, we consider the welfare criterion of aggregate payoffs rather than aggregate utility, which would include expressive utility. However, the resulting first-best election and policy outcome would remain the same as long as $n|\alpha_k, \alpha_{k'}| > \bar{e}$.

be more helpful to characterize the expected policy payoffs of the voters given the equilibrium policies $\{\mathbf{p}^A, \mathbf{p}^B\}$. Accordingly, we define Δ^x as the difference in expected policy payoffs between the two candidates for a voter with group identity I :

$$\Delta^x = E[x_i | \mathbf{p}^I] - E[x_i | \mathbf{p}^{I^c}]$$

Note that Δ^x is well-defined since the equilibrium policies of the candidates are unique.

This definition allows us to formulate the following proposition:

Proposition 1 (Comparative statics of the Identity model)

For $\lambda > 0.5$ there exists q^ such that for $q \geq q^*$, Δ^x is strictly positive and $\partial\Delta^x/\partial q$ and $\partial\Delta^x/\partial\lambda$ are strictly greater than zero.*

The comparative statics of the model with respect to q are also illustrated in Figure 1. The positive relationship between the candidates' partisan policy-bias and polarization in policy preferences is due to the fact that as q increases, the marginal benefit of partisan policy to the in-group increases, while the marginal cost of partisan policy to the out-group stays constant due to the linearity of payoffs.

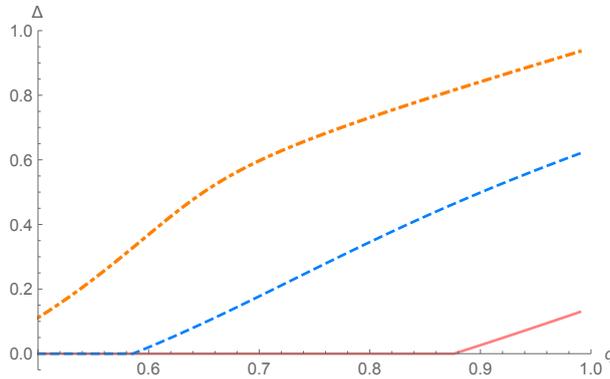


Figure 1: This graph shows the relative policy payoffs for electing the in-group candidate (Δ^x) as a function of the polarization in policy preferences (q), for different values of λ : 0.95 (orange, dot-dashed), 0.75 (blue, dashed), and 0.55 (red, solid).

VOTERS' BEHAVIOR:

Having detailed candidates' equilibrium behavior under the two different behavioral models, we turn to the main object of interest: voters' decisions. First, note that when their in-group candidate has a lower relative ability, voters may face a trade-off between voting expressively and voting instrumentally. In this case, the probability of influencing the outcome of the election becomes a relevant factor in the voting decision. Rather than explicitly considering the endogenous probability of being pivotal, we make the simplifying assumption that voters perceive their chances of influencing the outcome as invariant.

Assumption 1

Voters perceive their probability of being pivotal as constant and equal to \bar{p} .

While this assumption may be reasonable for large voting populations, it may bias the predictions of the model in small/medium groups. Therefore, we will directly account for this bias

when generating predictions and comparing results across experimental treatments.

Take Δ^α to be equal to the relative ability of the *out-group* candidate:

$$\Delta^\alpha = \alpha^{I^-} - \alpha^I.$$

The following proposition characterizes voting behavior under the benchmark model.

Proposition 2 (Voting: Benchmark model)

If agents' distributional preferences are characterized by efficiency, then voters will vote for the in-group candidate if and only if:

$$e_i \geq \bar{p}\Delta^\alpha. \tag{8}$$

Since the candidates will both choose a centrist policy, the voter's decision becomes a simple calculus of comparing the expressive payoff for voting for the in-group candidate and the relative ability of the out-group candidate, weighed by the perceived probability of influencing the election outcome.

As implied by Lemma 3 and as illustrated in the following proposition, the voting calculus becomes more complicated when agents have distributional preferences that favor the in-group.

Proposition 3 (Voting: Identity model)

If agents' distributional preferences are characterized by group identity, then voters will vote for the in-group candidate if and only if:

$$e_i + \bar{p}(E[x_i|\mathbf{p}^I] - E[x_i|\mathbf{p}^{I^-}]) \equiv e_i + \bar{p}\Delta^x \geq \bar{p}\Delta^\alpha. \tag{9}$$

Proposition 3 shows that, relative to the Benchmark model, the voting rule in the Identity model incorporates the perceived policy bias of the candidates, giving voters an instrumental incentive to vote for their in-group candidate.

Next, we define $\tilde{\Delta}_i^\alpha$ as the maximum value of Δ^α such that voter i prefers to vote for the in-group candidate:

$$\tilde{\Delta}_i^\alpha = \begin{cases} e_i/\bar{p} & \text{under Benchmark model,} \\ e_i/\bar{p} + \Delta^x & \text{under Identity model.} \end{cases}$$

In the following corollary, we utilize this definition to compare the comparative statics of voting behavior under the Benchmark and Identity models.

Corollary 1 (Voting: comparative statics)

- (i) Under the Benchmark model, $\tilde{\Delta}_i^\alpha$ is constant for all q .
- (ii) Under the Identity model, $\tilde{\Delta}_i^\alpha$ is weakly increasing in q .
- (iii) For $q = 0$, $\tilde{\Delta}_i^\alpha = e_i/\bar{p}$ under both the Benchmark and Identity models.

2.2 Summary of theoretical findings and identification strategy

While our model is a simplified setting, it captures features of political competition that are affected by social identities and polarization. Importantly, ex post policy discretion implies that

voters face uncertainty regarding which policy the candidates will select once in office. In such a setting, voters look to cues that signal candidates’ preferences, and hence are informative for predicting the candidates’ policy choices. Absent cues, voters might expect all candidates to maximize aggregate utility and choose a centrist policy. In a setting with identity division, however, the group identity model predicts that candidates will favor policy positions that disproportionately benefit the in-group. This implies that voters will interpret identity cues as a signal that the co-partisan candidate will select policies that are consistent with the political values and norms of the group, and hence rationally respond to these cues by voting in a partisan manner.

The formal theory that we present in this section illustrates a novel insight regarding the relationship between social and policy polarization. Namely, the impact of social polarization on partisan voting is a function of the degree of underlying polarization in policy preferences (see Corollary 1 (ii)). Importantly, this relationship is driven by the indirect, instrumental impact of social polarization: Intuitively, when policy preferences are homogeneous, there is little scope for choosing a policy that favors the in-group, and voters will expect the candidates to take a centrist policy if elected. Therefore, voters will prioritize the ability dimension when voting. As policy preferences polarize, however, voters will expect candidates to take partisan policy positions. In this case, partisan identity becomes the dominant concern when selecting between the candidates.

In addition to characterizing the direct (expressive) and indirect (instrumental) impact of partisan identity on voters’ decisions, the theory guides our experimental approach for empirically distinguishing between these two channels of influence: our identification strategy follows from the following two insights detailed in Corollary 1:

1. Given no polarization in voters’ policy preference, in-group voting is driven solely by expressive motives (direct impact).
2. Since expressive payoffs are constant, any increase in in-group voting as preference polarization increases is driven solely by instrumental motives (indirect impact).

First, (1) allows us to identify the level of expressive partisan voting by measuring the proportion of individuals who vote for an in-group candidate with lower relative ability when policy preferences are not polarized. Second, (2) allows us to identify the level of instrumental partisan voting for any other degree of preference polarization by accounting for the level of expressive partisan voting identified in (1).

3 Experimental Design

Our experimental design mirrors the theoretical framework previewed above and features two stages.

The *group inducement stage* happens at the beginning of each session. In our *Baseline-treatment*, subjects are randomly assigned into one of two groups of equal size, “Group A” or “Group B,” after they receive instructions for the voting game but before the voting game commences. We exogenously induce affective preferences over these groups by this standard minimal-group intervention because it has been shown to result in an in-group preference (Tajfel

and Billig, 1974; Goette, Huffman and Meier, 2006; Chen and Li, 2009; Landa and Duell, 2015) but also precludes that group membership is systematically correlated with other subject characteristics. We discuss the *Bike vs Car*- and the *Dem vs Rep*-treatments, featuring endogenous sorting into groups based on natural identities, in Section 4.4.

The *voting game stage* implements the structure and payoffs as laid out in Section 2 and utilizes the strategy method. That is, each subject makes decisions in the role of a *candidate* and in the role of a *voter* for all potential distributions of voters' ideal points and all possible combinations of candidate abilities. Each subject faces decision situations in which polarization in voters' policy preferences is absent (probability a voter holds extreme preferences, q , is 0), some degree of polarization (probabilities of extreme preferences are .25, .5, or .75), or perfect polarization (probability a voter holds extreme preference is 1). The ability of candidates is either low, medium, or high.

Identity groups correspond to voters's ideal points in the three-point policy space; specifically, we precisely control the degree of polarization in voters' policy preferences by changing the degree of correlation between voters' group membership and their ideal policy points: in the case of no polarization in policy preferences, all voters have ideal points at the center; in the case of full polarization, one group is located at the left while the other is at the right.

As candidates, subjects decide how to allocate up to 10 tokens to the three positions of a preference space, *Left*, *Center*, or *Right*. As voters, subjects make a choice between two candidates: Candidate A or Candidate B.

Again, the decision environment is characterized by the probability with which voters are at the extreme (0, .25, .5, .75, 1); this probability is assigned without replacement to a block of 9 rounds for each subject. At the beginning of this block, subjects make their allocation decision in the role of a candidate and then make 9 voting decisions between the candidates while the ability of candidates varies. The order of decision blocks and of vote-pairings within the block is randomized; therefore, subjects face decision environments in different orders.⁵ After the voting game, subjects also play one round of a dictator game and answer a questionnaire about basic demographics and the choices they made in the experiment.

In our experiment, subjects make decisions as candidate and voter but we are mostly interested in voter's choices. Acting as candidates, making an allocation decision in a given decision environment, helps subjects to form beliefs about what candidates may do, a belief that will then inform their choices as voters. The behavior of voters we approximate outside of the laboratory is characterized by exactly such uncertainty about what candidates are going to do once in office. Also, the implementation of how allocation decisions are made makes them a measure of in-group favoritism. Subjects' allocation choices as candidate can not be directly motivated by re-election concerns because voters do not learn about candidates' specific allocations. In this way, the allocation decision is an expression of subjects' distributional preferences as modeled in Section 2.

⁵An exact overview over the decision environment for each round of the experiment can be found in Section C.1 in the SI.

3.1 Payoffs

Subjects are paid depending on their and other subjects' choices in one randomly chosen decision situation of the voting game (and their choices in the dictator game). One subject from each identity group is chosen to be a candidate and is assigned an ability. Then, one distribution of voters' ideal points, q , is selected. Next, the subjects actual voting decisions for this distribution and set of candidate abilities are used to determine the "winning" candidate: the candidate with the largest vote share is the winning candidate and receives 15 tokens; the losing candidate receives 5 tokens. The winning candidate's token allocation, finally, determines the payoffs of the subjects not chosen as candidates. Each subject is assigned an ideal point equal to the partisan extreme with probability q . Corresponding to the model, if assigned an extreme ideal point, A-voters (B-voters) receive 1 token for each 1 token allocated to Left (Right) and .5 tokens for each 1 token allocated to the Center. Subjects assigned an ideal point of Center receive 1 token for each token allocated to center, and .5 tokens for each token allocated to Right or Left. Additionally, all voters receive 2 tokens if the winning candidate has a "low" ability, 3 tokens if "average" ability, and 5 tokens if "high" ability. Subjects received 5 Euro (7 Dollars) show-up fee plus the tokens they earned at an exchange rate of 60 cents for 1 token.

3.2 Hypotheses

In our experiment, subjects make decisions both as candidate and voter. Since subjects cast votes for all candidate-ability pairs, relative candidate ability serves as measure of how subjects trade-off between the candidates' ability and their identity as a function of the underlying preference distribution. As detailed in the Theory section, the Identity Model predicts that from an instrumental perspective the subjects should favor the in-group candidate to the extent that they expect candidates to choose a partisan policy.⁶ Therefore, our first hypothesis distinguishes between the Benchmark model, which predicts that voters will always select the candidate with higher ability, and the Identity model, which predicts that voters will favor their co-partisan candidate. Before stating the hypothesis, we first introduce the following term:

Definition 1 (Partisan Voting)

We define "Partisan Voting" as voting for the in-group candidate when the in-group candidate has a lower ability relative to the out-group candidate.

In other words, we only refer to voting for the in-group candidate as "partisan" when it is costly in ability terms.

Hypothesis 1 (No partisan voting)

Subjects will vote for whichever candidate has a higher level of ability (Benchmark vs Identity Model).

⁶For each distribution of ideal points, subjects take decisions as a candidate before taking decisions as a voter. While it is possible that the act of choosing policies as a candidate affects subjects' beliefs over the actions of the other subjects (in their role as candidates), we kept the order constant to preserve comparability across subjects and rounds.

Falsifying this hypothesis, however, is not sufficient to identify whether partisan voting is due to direct expressive or indirect, beliefs-driven instrumental concerns.

We measure the baseline level of expressive voting by eliciting subjects' votes between the candidates in a setting with no policy polarization. When all voters have an ideal point at the center of the policy distribution ($q = 0$) then all voters receive equal policy payoffs, regardless of the token allocation of the winning candidate. Therefore, the group identity model predicts that both candidates will choose equivalent policies, and voters have no instrumental incentive to vote for their co-partisan candidate. This benchmark level of expressive voting is constant across the levels of policy polarization.⁷ In accordance with the theoretical predictions, however, the level of instrumental partisan voting will increase with the level of policy polarization.

This gives us our second hypothesis:

Hypothesis 2 (Identifying Instrumental Voting)

The aggregate level of in-group voting is constant across ideal point distributions.

Note that our experiment is designed to precisely identify instrumental voting. For expressive voting, our measure may be biased. It represents an upper bound since, for example, subject errors may be classified as expressive voting.

Lastly, the identity model predicts that the level of partisan voting is an increasing function of the level of in-group affect (λ). In Section 4.4, we detail the results of several natural-identity treatments that are aimed to create variation in the level of affect across the identity groups:

Hypothesis 3 (Increasing Affect)

The aggregate level of in-group voting is constant across identity groups.

3.3 Session and summary statistics

In 7 sessions, with 24 subjects each (one with 26), we collect, for each subject, 45 observations as voter and 5 observations as candidate. In total, we collect observations on 170 subjects with a total of 7650 voter-round and 850 candidate-round observations. Given that subjects make decisions in the strategy method we have as many independent observations as subjects in the experiment. Subjects earning range from 7.7 to 20 Euro, average session earnings range from 12.9 to 18 Euros.⁸ We ran 4 sessions in the laboratory at Technical University Berlin (2 sessions for the baseline treatment and 2 sessions for the *Bike vs Car*-treatment) and 3 sessions in the laboratory at Florida State University (FSU; 1 session for the baseline treatment and 2 sessions for the *Dem vs Rep*-treatment).

For voters, the variable of interest is their *voting decision*: their choice of either candidate A or B. For candidates, we record their *allocation decision* of, in sum, up to 10 tokens to the positions left, center, and right. *Decisions situation* are characterized by the probability of the voter being positioned at the extreme (and not the center), a probability that is either 0, .25, .5,

⁷That is, our identifying assumption is that the level of expressive voting (and errors) is constant across ideal-point distributions. Also, as detailed in the Theory section, we assume that the probability of being pivotal is constant, or non-decreasing, in q – we show that this assumption holds in the following section.

⁸Table B.1 in the appendix provides an overview of these statistics.

.75, or 1). Also, the two candidates face by voters are assigned an *ability* of either low, medium, or high.⁹

4 Results

4.1 Partisan voting under social polarization

We start our investigation by assessing the prevalence of partisan voting in the baseline treatment, where we induce group identities using a minimal group intervention. For this subsection, we report the results of the two sessions run in Berlin together with one session run at FSU.¹⁰

We begin by documenting subjects' behavior for decision situations in which the probability that voters are located at the extreme position, q , is equal to one; that is, all A-voters have an ideal point of Left and B-voters have an ideal point of Right. This decision-environment maximizes the degree of policy conflict between the two groups, full polarization in policy preferences, and is equivalent to a zero-sum "divide the dollar" game. We find a strong bias among subjects to cast their vote for a candidate of their group: the average rate of in-group voting is .81 (.76,.85).¹¹ Specifically, we see that voters overwhelmingly choose a candidate of their own group when this candidate is assigned the same or a higher ability than the candidate of the other group, as to be expected. When the own candidate is of lower ability than the other candidate, the in-group candidate is still preferred, that is, voters engage in *partisan voting*. They elect the in-group candidate at an ability difference that amounts to a loss of 1 or 2 tokens about 60% of the time and in more than 40% of the decisions at a loss of 3 tokens. Figure 2 clearly demonstrates that the finding of partisan voting is independent of the difference in ability between a candidate who shares a group membership with voters and those who do not. In summary,

Result 1 (Partisan voting)

subjects disproportionately vote for their co-partisan candidate even when there is a negative ability difference between in- and out-group candidates.

Result 1 leads us to reject the *no partisan voting hypothesis* by showing that subjects vote in accordance with the identity model, and are willing to vote for their co-partisan candidate, even when this candidate has a lower relative ability.

⁹Table B.2 in the appendix gives the summary statistics on *voting decision* and *allocation decision* by treatment.

¹⁰There is no difference in behavior between laboratories: it is $-.02(-.09, .05)$ in in-group voting, $-.01(-.16, .15)$ in partisan voting, and $-.14(-.75, .46)$ in allocations to the in-group. Also see regression model 1 in Table B.6 and B.7 in the SI.

¹¹Here and throughout, we show 95%-confidence bound computed based on a subject-level clustered bootstrap. The results in this section are also robust to separating in-group voting by group membership (A or B), see Figure B.1 in the Appendix.

Figure 2: Rate of in-group voting in the baseline treatment with full polarization in policy preferences by cost of ability difference

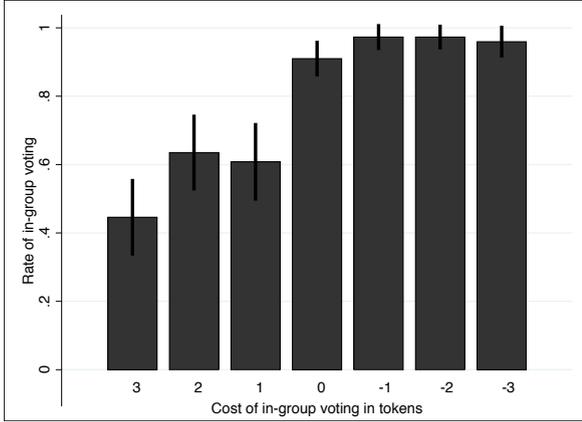
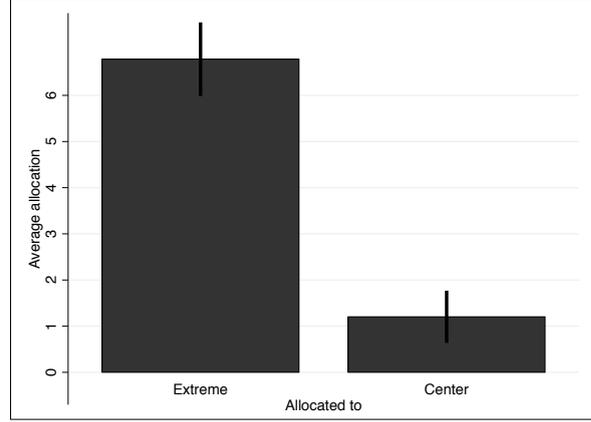


Figure 3: Allocation to partisan extreme (Left for Candidate A and Right for Candidate B) and center in the baseline treatment with full polarization in policy preferences



We also assess whether partisan voting is consistent with voters’ beliefs about what candidates are most likely to do or whether it is only psychologically sustained (i.e., in-group favoritism/out-group discrimination). Subjects’ allocation choices as candidates are one measure of such beliefs. We define *partisan allocations* as candidate allocations to the extreme that have the potential, given the distribution of voters, to favor in-group voters: A-candidates who allocate tokens to Left and B-candidates who allocate tokens to Right whenever the probability that voters are located at the extreme is higher than the one that they are located at the center. Candidates allocate heavily to the position where their own groups’ voters are most likely to reside. They allocate on average, 5.54 (5.11, 5.97) tokens to the extreme (Left and Right) but only 2.30 (1.97, 2.62) to the Center. Figure 3 illustrates this in-group biased pattern.

Subjects show similar patterns of in-group favoring behavior as candidate and as voter. Observing such in-group favoritism in form of partisan voting as well as partisan allocations could still be interpreted as *expressive choice* but voters in-group favoring vote choices may also reflect their expectations about candidates’ behavior. Expectations that would indeed be correct given candidates’ allocation choices.

4.2 Expressively and instrumentally motivated partisan voting under social polarization

We established the existence of partisan voting at a level of policy preference polarization among voters which would have called for no partisan bias assuming a rational agent; but is this finding enough to conclude behavior we induce is expressive? When partisan voting as well as partisan allocations vary with polarization in voters’ policy preferences in similar ways, what seems to be group biased behavior cannot be exclusively called an expressive choice. In the experiment, we vary the probability with which voters are located at the extreme of the policy space (Left for A-voters and Right for B-voters) from all voters in the Center (probability 0) – no polarization in policy preferences – to all A-voters located Left and all B-voters located Right (probability 1) – perfect polarization in policy preferences. Figure 4 demonstrates an increase in partisan

voting with increasing polarization in policy preferences.

Figure 4: Rate of partisan voting in the baseline treatment by polarization in voters' policy preference

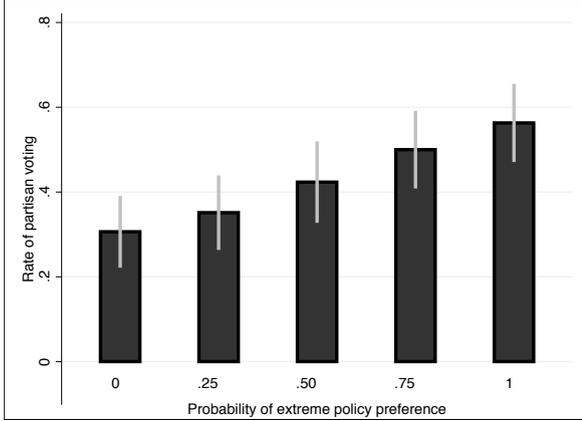
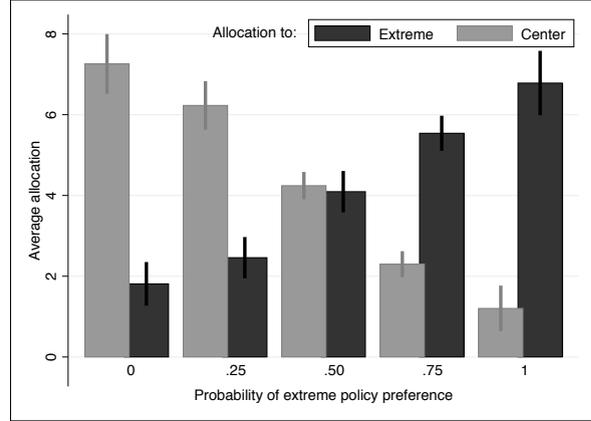


Figure 5: Allocation to the partisan extreme (Left for Candidate A and Right for Candidate B) and center in the baseline treatment by polarization in voters' policy preference



In particular, voters choose their in-group candidate when all voters are located at the center with probability 1. When the probability that voters are located at the extreme is 1, however, voters pick the in-group candidate at a rate of .56 (.46,.66). The marginal effect of increasing the probability of voters being located at the extreme positions by .25, increases voters' propensity for a partisan vote by 6.6 (4.2, 9.0) %.¹²

The following result summarizes these findings:

Result 2 (Instrumental partisan voting)

The aggregate level of partisan voting is increasing in the level of polarization of voters' policy preferences (q).

This result confirms our main hypothesis (Hypothesis 2 – Identifying Instrumental Voting) and shows that subjects instrumentally respond to the expectation of candidates' biased policy choices. This is evidence of an indirect, beliefs-driven effect of social identity.

Partisan voting for the case of no polarization in voters' policy preferences ($q = 0$) provides a benchmark of the level of expressive partisan voting – as seen in Figure 4, comparing the cases of no polarization in voters' policy preferences ($q = 0$) and full polarization ($q = 1$), the level of socially-costly partisan voting increases by over 70 percent – an increase that can be attributed to the instrumental impact of social polarization.

Partisan voting is also not purely driven by a subset of voters who always vote along partisan lines: only 14 percent of voters vote for their co-partisan candidate when the cost (in ability difference) is high and polarization in voters' policy preferences is low, while 61 percent vote for

¹²Estimates of the marginal effects are computed based on the regression in Table B.4 in the SI. The effect shown is the average marginal effect of a change in probability of voters being positioned on the extreme of .25 when the other variables are hold at their mean.

their co-partisan candidate when the cost is low and polarization in voters' policy preferences is high.¹³

4.3 Robustness of instrumentally motivated partisan voting under social polarization

Clean identification of instrumentally-driven voting is threatened by a range of behavioral patterns we may find in our data: variation of vote choice with voters' perceived pivotality, differences in motivations behind voters' decisions at different levels of ability (i.e., the relationship between policy-component and ability-component in voters' utility function is multiplicative not additive), or vote choice that does not reflect correct beliefs about candidates' allocation decisions.

In our model, we assume that agents perceive their probability of influencing the election outcome as constant. While our experiment consists of a medium-sized election (22-24 voters), it is possible that subjects take into account the strategic response of other subjects when forming their belief about their probability of being pivotal. Accordingly, subjects may anticipate that their probability of being pivotal is a function of the distribution of ideal points (q), and may be more likely to vote expressively when the probability of being pivotal is low. This could affect our identification strategy if the probability of being pivotal is negatively correlated with q . Note, however, that the probability of being pivotal is increasing with the degree of partisan voting, which implies that the probability of being pivotal is positively correlated with q . Therefore, the observed voting behavior is consistent with our interpretation of partisan voting for $q = 0$ as an *upper bound* on expressive voting.

We also consider the sensitivity of partisan voting to the degree of difference in candidate ability. Figure 6 shows the level of partisan voting when the ability difference between in-group and out-group candidate is "low" or "high". While, as predicted, partisan voting does decrease as the ability difference between the candidates increases, this decrease appears to be solely due to a decrease in expressive partisan voting. The level of partisan voting decreases substantially for $q = 0$ (expressive voting), while the degree to which partisan voting increases with q (instrumental voting) remains unaffected.¹⁴

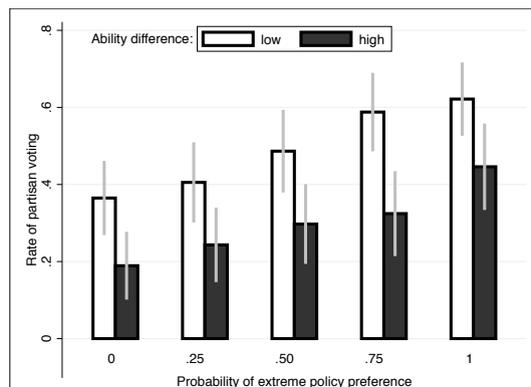
Finally, we consider whether instrumental partisan voting is rational based on the allocation decisions of the candidates. We see that variation in partisan voting is mirrored in the candidates' average allocation choices. When the probability of voters being located at the extreme (Left for A-voters and Right for B-voters) increases, allocation to the extreme position where the own group is more and more likely to reside increases (See Figure 5). Average allocation of tokens to the extreme position in a decision situation with a probability of 1 that voters are at the extreme is 6.78 (5.99, 7.58) while it is only 1.81 (1.27, 2.35) with a probability of 0. Average allocation to the Center is 1.20 (.64, 1.77) and 7.23 (6.52, 8.00), respectively. The marginal effect of a change in the probability that voters are located at the extreme of .25 leads to rise of

¹³We provide more tests to demonstrate the robustness of our findings to individual-level idiosyncrasies in the Section B.6 of the SI.

¹⁴This difference in partisan voting also shows in a significant coefficient on a high negative ability difference-dummy in regression model 2 in Tables B.6 and B.7 in the SI.

allocation to the extreme of 1.31 (1.09, 1.53) tokens but to a decline in allocation to the center of 1.61 (1.85, 1.37).¹⁵

Figure 6: Partisan voting when the ability difference between candidates is low (difference in candidate ability in-group/out-group: low ability vs medium ability or medium ability vs high ability) or high (difference in candidate ability in-group/out-group: low ability vs high ability)



On average, when compared to the voting behavior that would maximize individual payoffs, partisan voting is too low for high degrees of polarization in voters’ policy preferences, and too high for low degrees of polarization in voters’ policy preferences – i.e. subjects’ behavior is consistent with the model’s predictions regarding the comparative statics, but are off on the levels.

4.4 Robustness of expressively and instrumentally motivated partisan voting under social polarization to variation in group affect

The results of the baseline treatment confirm the predictions of the Identity model with respect to partisan voting and the polarization of policy preferences, and identifies both expressive and instrumental partisan voting. In this subsection, we report the results of two treatments designed to vary the *intensity* of partisan affect (λ).

To vary the level of partisan affect, we modify our baseline design to divide subjects into groups based on their reply to a question designed to prime a natural identity. For the FSU sample, subjects were allocated to groups based on whether they reported to feel closer to Democrats or Republicans.¹⁶ For the Berlin sample, subjects were allocated to groups based on whether they reported that they are more likely to use their bike than car¹⁷ While this may seem a strange choice of an identity-group, a pre-experiment survey of the experimental pool in Berlin showed that students rank the car/bike-divide as being more important than religious or political affiliations and that subjects are evenly split between car and bike.

To place the results from FSU in context, the sessions were run in June of 2016, shortly after Donald Trump had secured the nomination as the Republican presidential candidate. Given the

¹⁵Marginal effects are computed based on the regression shown in Table B.5 in the SI.

¹⁶We ran this treatment at FSU since questionnaires showed that their subject pool was roughly evenly divided between Democrats and Republicans – we also inform subjects of this fact before they select a group.

¹⁷In German: “Sagen Sie uns bitte, ob Sie öfters Ihr Fahrrad oder Auto benutzen?”

contentious nature of the 2016 presidential election, we had the *ex ante* expectation that the level of partisan affect induced by subjects party identity would be significantly greater than a minimal group intervention.

As shown in Figure 7, however, patterns of partisan voting are nearly identical for party identity, car/bike-identity, and minimal group treatments. Looking at the allocation decisions of candidates, we see that this null finding is due to a treatment failure: as detailed in the experimental design section, subjects' allocation decisions when $q = 1$ gives us a measure of in-group affect, since in this case the candidate allocation decision is equivalent to a zero-sum "divide-the-dollar" game. Subjects' allocation is slightly *less* biased towards the in-group for the party identity treatment on average (this difference, however, is not significant). Therefore, given the non-increase in the partisan bias of the candidates' allocations, the non-increase in partisan voting is still perfectly in line with the theoretical predictions.

Figure 7: Rate of partisan voting by policy preference polarization and treatment

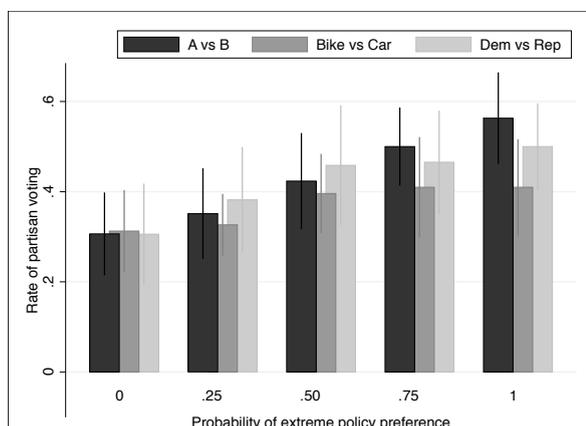
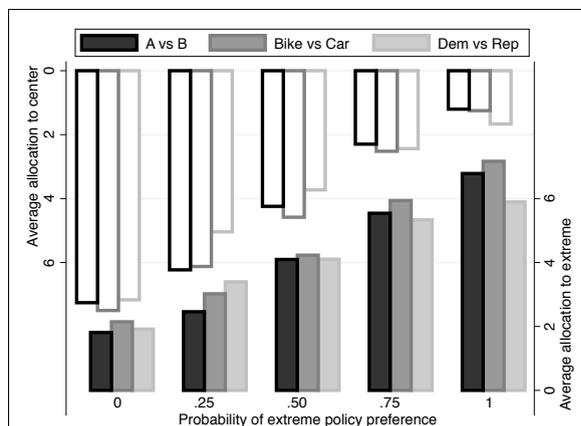


Figure 8: Rate of allocation to extreme and center by policy preference polarization and treatment



Lastly, while our priming identity treatments fail to implement a higher level of affect than our minimal group treatment, subjects, in the party identity treatment, were asked to self-report the strength of their party-identification in an post-experiment questionnaire. We use the variation in the strength of party identity within the subject pool to investigate the impact of higher affect. Our treatments were not explicitly designed to test the impact of higher affect using within-session variation thus, this *ex-post* analysis should be considered as exploratory.

Figure 9 shows that our measure of expressive partisan voting (partisan voting at $q = 0$) features no significant difference between subjects who reported a strong or weak party identification, albeit with a small sample size. At larger values of q , however, we see a significant increase in the level of partisan voting for strongly partisan subjects.¹⁸ This suggests an increase in instrumental partisan voting that is driven by higher expectations of partisan bias among

¹⁸In regression model 4 in Tables B.6 and B.7 in the SI, we see that the interaction of the strong partisan dummy and q is positive and significant. However, as seen in model 5, the interaction is not significant if we include the strong partisan dummy *and* the interaction – this could be due to the small sample of strong partisans, or due to the non-monotonic relationship between partisan voting and q exhibited by strong partisans (see Figure 9).

strongly partisan subjects. Interestingly, we do not observe an increase in partisan bias in allocation decisions for either subjects with strong or with weak party identification (see Figure 10).

Figure 9: Rate of partisan voting in the baseline treatment (FSU sample) by policy preference polarization and self-reported strength of party identification

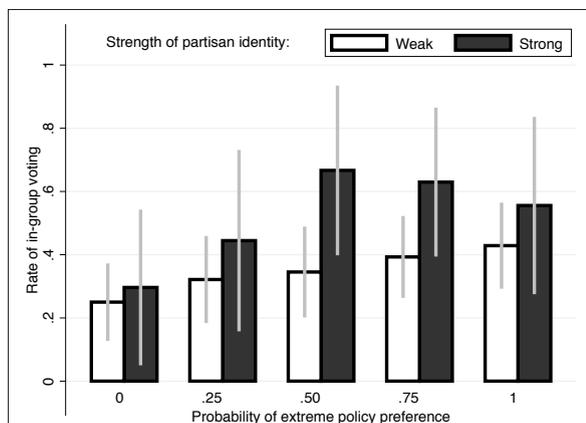
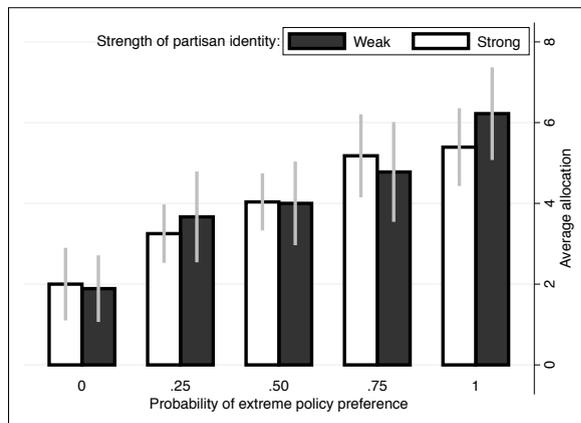


Figure 10: Allocation to partisan extreme (Left for Candidate A and Right for Candidate B) and center in the baseline treatment (FSU sample) by policy preference polarization and self-reported strength of party identification



Summarizing,

Result 3

Our experiment gives only weak evidence for an increase in partisan voting with stronger affect but robustly identifies expressively and instrumentally-driven motivations behind partisan voting.

This result, however, comes with the caveat that our experiment failed to induce stronger partisan affect *between* treatments.¹⁹

5 Discussion and conclusion

In this paper, we utilize formal theory to explicitly characterize the mechanisms by which social polarization impacts voting behavior, and laboratory experiments to cleanly identify and measure the impact of the mechanisms characterized by the theory. In this section, we discuss the robustness of the identified effects of social polarization on partisan voting and the portability of our stylized findings to an external setting.

Several elements of our theory-driven experimental design allow for identification of the effect of social polarization on partisan voting as well as the separation of expressive and instrumental motivations behind vote choice.

First, we fix the number and type of salient identity dimension. Certainly, multiple identities may compete for attention in real world political competition. Many studies operationalize

¹⁹The SI holds results from yet another treatment we implemented where subjects were induced with multiple-identities to gauge the effect of potentially different intensity in social identity on behavior. We also find weak evidence for an increase in partisan voting with stronger affect in this treatment

salience by self-reported importance-scales while we, for clarity, bring subjects into a decision situation in which only one identity matters. In our study, when political polarization is set to a certain level, it potentially represent several scenarios with different levels of political polarization on other dimensions. This feature of our design could be seen as particularly problematic given that one definition of political polarization is exactly that preferences in multiple dimensions or identities in multiple dimensions are aligned. We opt for this reduced form of a one-dimensional political space with one, potentially salient, identity in our baseline to add multiple identities later. In fact, we implement a multiple identity-treatment (2 identities) and vary whether they align or contrast in their correlation with policy preferences. Our results hold in this set-up as well (See Section B.5 in the SI).

Second, our parsimonious setting also allows us to experimentally separate group identity from political preference. In the laboratory, we can exactly specify voters' utility derived from a policy and candidates' ability. The distribution of ideal points correlates with group membership but it is still randomly assigned (i.e., some members of a particular identity group are assigned a more extreme ideal point and some are assigned a centrist ideal point). In this way, we define what a voter is to gain from a particular policy and, knowing that, estimate the voters' valuation of an expressive choice. Work based on surveys or unincentivized experiments face more challenges to separate policy utility, ideology-driven choices, and voting based on identity motivations (Mason, 2015), while it is clearly necessary to separate the policy-dimension from the psychological identity component of partisanship and/or ideology to understand the determinants of vote choice (Malka and Lelkes, 2010; Devine, 2015; Huddy, Mason and Aarøe, 2015).

Third, subjects in our experiment are also perfectly informed about the level of polarization in voters' policy preferences, while outside the laboratory voters tend to exaggerate the level of polarization (Westfall et al., 2015; Levendusky and Malhotra, 2015; Sherman, Nelson and Ross, 2003) or have heterogeneous beliefs about the level of polarization adding noise to estimating identity effects in observational data. Additionally, our subjects' behavior in the baseline treatment will not be driven by familiarity with the labels given random assignment of group affiliation.²⁰

Fourth, any interaction where partisanship labels are provided will trigger individuals' beliefs about how in-group and out-group members may behave. In certain decision situations, different expectations about one partisan group may act on the outcome variable in a way that, on average, the treatment effects cancel out. Say, we want to know how learning candidate's partisanship affects voter's assessment of future performance in office. A Republican candidate, for example, may be seen by a portion of the electorate as being less likely than a Democratic candidate to attempt to follow through on a particular policy promise, but as more able to actually deliver on the promise given an attempt. Both are reasonable assumptions about actual expectations, but a treatment that increases the salience of partisan labels could result in no change in vote choice, in the aggregate, given differing beliefs among the electorate. Starting from minimal groups and studying how such minimal identities, that are randomly assigned,

²⁰Broockman and Ryan (2016) show that familiarity with candidates alters behavior contingent on partisanship.

shape voters' decision, can help to correctly predict behavior more precisely once particular beliefs come into play.

Finally, for building all counterfactuals to causally identify the effect of social polarization on partisan voting, using hypothetical real-world scenarios would provide another viable research design (see e.g. Rogowski and Sutherland (2015)). Such design, however, suffers from the possibility that some scenarios are more likely or more believable than others. Voters may see polarization as the default and their responses in such situations may simply be different from behavior in a situation without polarization based on familiarity of the decision-situation. Keeping context at a minimum, as it is done in the baseline of our experiment, avoids this heterogeneous treatment effect.

A common concern with laboratory experiments is the question of whether the observed mechanism travel outside the laboratory? A range of features of our experiment aim to ensure ecological validity of electoral decisions and, ultimately, are implemented to enable the mechanism we observe within the experiment to travel outside the laboratory.

First, we randomly assign policy preferences but allow subjects to tie the preferences in this abstract game to policy preferences or an ideology subjects know from outside of the laboratory: the left-right dimension. While we aim to invoke the left-right dimension in subjects' mind, their assigned group membership and ideal point is not tied to differences in underlying endowment. In other words, our political dimension of competition is not correlated with income in any way. Still, decision-making should travel, at least, partially to every political behavior that is conceptualized by voters to happen in a one dimensional space.

Second, while we do not experimentally prime inter-group competition or create group threats, we bring subjects into a situation where groups compete – in the sense that certain decisions benefit the in-group over the out-group – and a group's identity and status may be seen as threatened. Inter-group competition, group threats, definition and re-definition of group boundaries, and status concerns have shown to invoke emotions directed towards the in-group and against the out-group even with minimal groups (Brewer, 1999). Arguably, the degree of emotions we instill is a lower bound of what we would observe in actual political competition.

Finally, and most importantly, we compare behavior with minimal groups to behavior when naturally occurring identities are primed. In particular, we allow context from non-political identities (German sample) and partisan identities (American sample) into the experiment. Extant research in behavioral economics has shown that when moving from interactions with minimal groups to actions with stronger, more contextualized group identities, patterns of beliefs and behavior change in degree but not in direction of group effects (Eckel and Grossman, 2005; Chen and Li, 2009; Goette, Huffman and Meier, 2006; Bernhard, Fehr and Fischbacher, 2006; McLeish and Oxoby, 2011). We extend this work to voting in a socially polarized political environment and find that our results are somewhat stronger when looking at partisan identities than when studying minimal groups, but no change in direction of the established treatment effect occurs.

Conclusion

In this paper we characterize and measure the impact of social polarization on partisan voting. In particular, we show that social identity has an indirect effect on voting through voters beliefs regarding the ex post decisions of political representatives on top of a direct effect through an expressive channel. Our laboratory experiment largely confirms the predictions of our formal model, and shows that both the direct and indirect impact of social identity significantly affect voter behavior. Moreover, we show that the impact of social polarization on voting behavior is strongest when there is large underlying polarization in voters' policy preferences.

Our findings suggest that polarization in voters' policy preferences is not sufficient to cause partisan behavior alone, that is, a society divided into two groups with starkly divergent policy preferences but with low levels of social polarization may still support a relatively consensual, non-partisan and efficient political system. Instead, socially-costly partisan behavior will occur when both policy and social polarization are high. In this case, as a society becomes polarized on a social dimension, voters come to expect a higher degree of partisanship from elected officials, which causes them to rationally respond by voting along partisan lines, leading to the selection of candidates with lower average levels of valence.

This narrative is particularly problematic when applied to the US context, where evidence suggests that social polarization occurs along an explicitly political dimension (Republican/Democrat). In this case, a self-reinforcing cycle may arise: Social polarization leads to the perception of increasing divergence in the policy platforms of the two parties in turn, this may cause partisan identities to strengthen, leading to an even greater degree of social polarization and a further increase in partisan behavior.

References

- Bassi, Anna, Rebecca B Morton and Kenneth C Williams. 2011. "The effects of identities, incentives, and information on voting." *The Journal of Politics* 73(2):558–571.
- Bernhard, Helen, Ernst Fehr and Urs Fischbacher. 2006. "Group affiliation and altruistic norm enforcement." *American Economic Review* 96(2):217–221.
- Brewer, Marilynn. 1999. "The Psychology of Prejudice: Ingroup Love or Outgroup Hate?" *Journal of Social Issues* 55(3):429–44.
- Broockman, David E and Timothy J Ryan. 2016. "Preaching To The Choir: Americans Prefer Communicating To Copartisan Elected Officials." *American Journal of Political Science* 60(4):1093–1107.
- Charness, Gary, Luca Rigotti and Aldo Rustichini. 2007. "Individual Behavior and Group Membership." *American Economic Review* 97(4):1340–52.
- Charness, Gary and Matthew Rabin. 2002. "Understanding social preferences with simple tests." *The Quarterly Journal of Economics* 117(3):817–869.
- Chen, Yan and Sherry Li. 2009. "Group Identity and Social Preferences." *American Economic Review* 99(1):431–57.
- Devine, Christopher J. 2015. "Ideological social identity: Psychological attachment to ideological in-groups as a political phenomenon and a behavioral influence." *Political Behavior* 37(3):509–535.
- Dickson, Eric and Kenneth Scheve. 2006. "Social Identity, Political Speech, and Electoral Competition." *Journal of Theoretical Politics* 18(1):5–39.
- Eckel, Catherine and Philip Grossman. 2005. "Managing Diversity by Creating Team Identity." *Journal of Economic Behavior & Organization* 58:371–392.
- Fehr, Ernst and Klaus Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114(3):817–68.
- Goette, Lorenz, David Huffman and Stephan Meier. 2006. "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups." *American Economic Review* 96(2):212–6.
- Goette, Lorenz, David Huffman and Stephan Meier. 2012. "The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups." *American Economic Journal: Microeconomics* 4(1):101–15.
- Green, Donald, Bradley Palmquist and Eric Schickler. 2002. *Partisan Hearts and Minds*. New Haven: Yale University Press.
- Hamlin, Alan and Colin Jennings. 2011. "Expressive political behaviour: Foundations, scope and implications." *British Journal of Political Science* 41(03):645–670.
- Huddy, Leonie, Alexa Bankert and Caitlin L Davies. 2018. "Expressive vs. Instrumental Partisanship in Multi-Party European Systems." *Advances in Political Psychology* .
- Huddy, Leonie, Lilliana Mason and Lene Aarøe. 2015. "Expressive partisanship: Campaign involvement, political emotion, and partisan identity." *American Political Science Review* 109(01):1–17.

- Iyengar, Shanto and Sean J Westwood. 2015. "Fear and loathing across party lines: New evidence on group polarization." *American Journal of Political Science* 59(3):690–707.
- Klor, Esteban and Moses Shayo. 2010. "Social Identity and Preferences over Redistribution." *Journal of Public Economics* 94(3):269–78.
- Landa, Dimitri and Dominik Duell. 2015. "Social Identity and Electoral Accountability." *American Journal of Political Science* 59(3):671–89.
- Levendusky, Matthew S and Neil Malhotra. 2015. "(Mis) perceptions of Partisan Polarization in the American Public." *Public Opinion Quarterly* 80(S1):378–391.
- Malka, Ariel and Yphtach Lelkes. 2010. "More than ideology: Conservative–liberal identity and receptivity to political cues." *Social Justice Research* 23(2-3):156–188.
- Mason, Lilliana. 2015. "'I disrespectfully agree': The differential effects of partisan sorting on social and issue polarization." *American Journal of Political Science* 59(1):128–145.
- McLeish, Kendra N and Robert J Oxoby. 2011. "Social interactions and the salience of social identity." *Journal of Economic Psychology* 32(1):172–178.
- Rogowski, Jon C and Joseph L Sutherland. 2015. "How Ideology Fuels Affective Polarization." *Political Behavior* pp. 1–24.
- Shayo, Moses. 2009. "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution." *American Political Science Review* 103(2):147–74.
- Sherman, David K, Leif D Nelson and Lee D Ross. 2003. "Naï Realism and Affirmative Action: Adversaries are More Similar Than They Think." *Basic and Applied Social Psychology* 25(4):275–289.
- Tajfel, Henri. 1981. *Human Groups and Social Categories*. Cambridge: Cambridge University Press.
- Tajfel, Henri and Michael Billig. 1974. "Familiarity and Categorization in Intergroup Behavior." *Journal of Experimental Social Psychology* 10:159–70.
- Westfall, Jacob, Leaf Van Boven, John R Chambers and Charles M Judd. 2015. "Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide." *Perspectives on Psychological Science* 10(2):145–158.

Supporting Information

A Proofs for Section 2

Proof of Lemma 1:

First, note that the candidates' aggregate payoffs are constant and equal to $x^l + x^w$. Next, note that $\max_{\mathbf{p}} \sum^N v(\mathbf{p}, p_i) = \{0, 1, 0\}$ since a “policy unit” placed at p^m generates aggregate payoffs of $n(1 - q) + nq$, while a policy unit placed at p^l or p^r generates aggregate payoffs of $1/2n(1 - q) + nq \leq n(1 - q) + nq$. Lastly, given that valence is a public good, aggregate payoffs are higher when the candidate with $\alpha_k \geq \alpha_{k'}$ wins the election. Together, this implies that maximal aggregate payoffs are achieved when both candidates choose centrist policies, and all voters vote for the highest-valence candidate. ■

Proof of Lemma 2:

Since the equilibrium concept is SPNE, by backward induction, the winning candidate will choose the policy that maximizes their social preferences, $g(\mathbf{x}^I, \mathbf{x}^{I-}) = \sum_{j \neq i} \sqrt{x_j}$. And since payoffs are separable, this is equivalent to the following maximization problem, $\max_{\mathbf{p}} \sum^N \sqrt{E[v(\mathbf{p}, p_i)]}$, which is equal to $\{0, 1, 0\}$ by the proof of Lemma 1. ■

Proof of Lemma 3:

As in the above proof, the winning candidate will choose the policy that maximizes their social preferences, $g(\mathbf{x}^I, \mathbf{x}^{I-})$, which results in the following maximization problem:

$$\max_{\mathbf{p}} \left(\lambda \sum^I E[v(\mathbf{p}^w, p_i)]^{\frac{1}{2}} + (1 - \lambda) \sum^{I-} E[v(\mathbf{p}^w, p_i)]^{\frac{1}{2}} \right).$$

For simplicity, we assume c^A wins the election and sets $\mathbf{p}^w = \mathbf{p}^A$ (\mathbf{p}^B is symmetric).

To prove the result, we focus on the marginal utility that c^A receives from shifting a unit of policy to p^l , given $\mathbf{p}^A = \{0, 1, 0\}$. Note that:

$$E[v(\mathbf{p}^w, p_i)] = q(p^l + 0.5p^m) + (1 - q)(0.5p^l + p^m + 0.5p^r),$$

for $i \in A$, and

$$E[v(\mathbf{p}^w, p_i)] = (1 - q)(0.5p^l + p^m + 0.5p^r) + q(0.5p^m + p^r),$$

for $i \in B$. Therefore, the relative marginal utility of partisan policy at $\mathbf{p}^A = \{0, 1, 0\}$ is equal to:

$$\begin{aligned} & \lambda(n/2 - 1) \frac{1}{2} E[v(\{0, 1, 0\}, p_i) | i \in A]^{-\frac{1}{2}} \left[\frac{1}{2}q - \frac{1}{2}(1 - q) \right] \\ & + (1 - \lambda)(n/2 - 1) \frac{1}{2} E[v(\{0, 1, 0\}, p_i) | i \in B]^{-\frac{1}{2}} \left[-\frac{1}{2}(1 - q) + -\frac{1}{2}q \right], \end{aligned} \quad (10)$$

which is positive iff:

$$\lambda \left[\frac{1}{2}q - \frac{1}{2}(1 - q) \right] > (1 - \lambda) \left[\frac{1}{2}(1 - q) + \frac{1}{2}q \right] \Rightarrow \lambda(2q - 1) > 1 - \lambda. \quad (11)$$

(Note that the same calculation for p^r yields the equation $(1 - \lambda)(2q - 1) > \lambda$, which implies that $p^r = 0$, since the marginal relative utility of placing a unit of policy at the outgroup extreme is always negative.) Therefore, $\mathbf{p}^A = \{0, 1, 0\}$ if Equation 11 does not hold, and $\mathbf{p}^A = \{p^l, p^m, 0\}$ with $p^l > 0$ if Equation 11 holds. ■

Proof of Proposition 1:

The proof follows from Expression 10 and Equation 11 in the proof of Lemma 3 above. First, fixing λ , there exists q^* such that $\lambda(2q - 1) = 1 - \lambda$, since $\lambda > 0.5$. Specifically:

$$q^* = \frac{1 - \lambda}{2\lambda} + \frac{1}{2}.$$

For, $q < q^*$, Equation 11 does not hold, and as shown in the proof of Lemma 3, the unique equilibrium policies for both candidates set $\mathbf{p}^I = \{0, 1, 0\}$; for $q > q^*$, however, both candidates will allocate a strictly positive amount to ingroup extreme. For convenience, assume c^A wins the election. In this case, c^A will set p^l, p^m such that the relative marginal utility of partisan policy (Expression 10) is equal to zero. That is:

$$\frac{E[v(\{p^l, p^m, 0\}, p_i)|i \in A]}{E[v(\{p^l, p^m, 0\}, p_i)|i \in B]} = \left(\frac{\lambda(2q - 1)}{(1 - \lambda)} \right)^2 \quad (12)$$

Both the RHS and LHS of Equation 12 are continuous in q, λ and p^l . Moreover, for $q > q^*$, the LHS is increasing in p^l since $E[v(\{p^l, p^m, 0\}, p_i)|i \in A]$ is increasing in p^l and $E[v(\{p^l, p^m, 0\}, p_i)|i \in B]$ is decreasing in p^l , which implies a unique crossing.

For $\partial\Delta^x/\partial\lambda$, note that the RHS of Equation 12 is increasing in λ while the LHS is constant. Therefore, the equilibrium value of p^l is increasing in λ , which implies that $E[v(\{p^l, p^m, 0\}, p_i)|i \in A] - E[v(\{p^l, p^m, 0\}, p_i)|i \in B]$ is increasing in λ as well.

Second, the case of $\partial\Delta^x/\partial q$ is a bit more complex since both the RHS and LHS of Equation 12 are functions of q . However, the result can be proved directly. First, we introduce the following simplified notation: $E[v^I] = E[v(\{p^l, p^m, 0\}, p_i)|i \in I]$ and $X = (\lambda(2q - 1)/(1 - \lambda))^2$.

Consider a discrete increase in q to $q' = q + \Delta q$. We wish to show that $E[v^A]' - E[v^B]' > E[v^A] - E[v^B]$ or, equivalently, that $\Delta E[v^A] > \Delta E[v^B]$. Using this notation and rearranging Equation 12, we get:

$$E[v^A]' + (X + \Delta X)E[v^B]' = 0 = E[v^A] + XE[v^B].$$

And since X is increasing in q , which implies that $\Delta X > 0$, we get:

$$E[v^A]' + XE[v^B]' > E[v^A] + XE[v^B].$$

Rearranging this equation gives:

$$\Delta E[v^A] + X\Delta E[v^B] > 0 \Rightarrow \Delta E[v^A] > X\Delta E[v^B].$$

Note that since $X > 1$ this expression implies that $\Delta E[v^A] > \Delta E[v^B]$, which proves the result for a discrete change. Lastly, since all expressions are continuous in q , this result holds as $\Delta q \rightarrow 0$. ■

Proof of Proposition 2:

Lemma 2 shows that under the Benchmark model, both candidates will select $\mathbf{p}^w = \{0, 1, 0\}$. Since both candidates select the same policy, voter i receives the following relative expected utility for voting for the ingroup candidate:

$$e_i + \bar{p}(\alpha^I - \alpha^{i^-}),$$

which is positive iff $e_i \geq \bar{p}\Delta^\alpha$. ■

Proof of Proposition 3:

Lemma 3 shows that under the Identity model, both candidates may select policies that favor their partisan ingroup. Therefore, voter i receives the following relative expected utility for

voting for the ingroup candidate:

$$e_i + \bar{p}(E[x_i|\mathbf{p}^I] - E[x_i|\mathbf{p}^{I^-}] + \alpha^I - \alpha^{i^-}),$$

which is positive iff $e_i + \bar{p}\Delta^x \geq \bar{p}\Delta^\alpha$. ■

Proof of Corollary 1:

(i) Follows directly from Proposition 2: Under the Benchmark model the candidates choose the same policy for all q , which implies that voters have a constant incentive to vote for the ingroup candidate.

(ii) In contrast to (i), Proposition 1 shows that the incentive to vote for the ingroup candidate is increasing in q , since candidates choose more partisan policy for higher q .

(iii) Proposition 1 shows that when $q = 0$, there is no instrumental incentive to vote for the ingroup candidate under the Identity model since $\Delta^x = 0$. ■

B Statistical appendix

B.1 Session statistics

Table B.1: Session statistics

Session	Treatment								
	Single ID						Multi ID		
	A vs B			Bike vs Car		Dem vs Rep		1	2
	1	2	3	1	2	1	2	1	2
Number of subjects	24	24	26	24	24	24	24	24	24
Number of observations									
of voters	1080	1080	1170	1080	1080	1080	1080	864	864
of candidates	120	120	130	120	120	120	120	96	96
Average earnings (in Euro)	13.9	12.9	13.4	14.5	12.9	13.3	14.3	18.0	16.6

B.2 Summary statistics

Table B.2: Summary statistics

Variable	Single ID			
	A vs B	Bike vs Car	Dem vs Rep	Multi-ID
<i>Share of subjects in A</i>	.50 (.50)			.50 (.50)
<i>Share of subjects in Bike</i>		.52 (.50)		.65 (.48)
<i>Share of subjects in Dem</i>			.69 (.46)	
<i>voting decision</i>				
vote for A over B	.52 (.50)			
vote for Bike over Car		.55 (.50)		
vote for Dem vs Rep			.61 (.49)	
vote for A-Bike over B-Car				.55 (.50)
vote for A-Car over B-Bike				.41 (.49)
<i>allocation decision</i>				
Left	2.84 (2.87)	2.89 (2.91)	3.65 (2.87)	
Center	4.33 (3.24)	4.40 (3.25)	4.01 (2.92)	
Right	2.68 (2.67)	2.67 (3.10)	2.05 (2.02)	
Left-Up				2.76 (2.87)
Right-Up				2.21 (2.49)
Center				5.31 (4.35)
Right-Down				2.26 (2.42)
Left-Down				2.51 (2.71)

Table B.3: Difference in rate of in-group voting (=partisan voting) and in average in-group allocations (=partisan allocation) between A vs B-treatment and Bike vs Car- as well as Dem vs Rep-treatment.

A vs B vs		Partisan voting	Partisan allocation	
			to extreme	to center
Bike vs Car	0	-.01 (-.13,.12)	-.34 (-1.35,.68)	-.24 (-1.41,.93)
	.25	.02 (-.09,.14)	-.56 (-1.31,.18)	.10 (-.85,1.06)
	.5	.03 (-.13,.19)	-.13 (-1.10,.83)	-.34 (-1.12,.44)
	.75	.09 (-.07,.25)	-.40 (-1.29,.50)	-.22 (-.80,.35)
	1	.15 (-.00,.31)	-.38 (-1.52,.76)	-.05 (-.85,.75)
	Overall	.06 (-.06,.18)	-.36 (-.95,.23)	-.15 (-.58,.28)
Dem vs Rep	0	.00 (-.13,.14)	-.11 (-.98,.77)	.09 (-1.11,1.29)
	.25	-.03 (-.16,.10)	-.94 (-1.68,-.20)	1.19 (.41,1.97)
	.5	-.03 (-.18,.11)	-.01 (-.86,.84)	.51 (-.23,1.26)
	.75	.03 (-.10,.17)	.21 (-.75,1.16)	-.14 (-.83,.55)
	1	.06 (-.07,.20)	.89 (-.36,2.14)	-.46 (-1.44,.52)
	Overall	.00 (-.11,.13)	.01 (-.58,.59)	.24 (-.16,.63)

B.3 In-group voting, partisan voting, allocation decision regressions

Table B.4: Logistical regression of in-group voting and partisan voting on probability of extreme policy preferences (q), order in which q is shown to subjects, and round of play in the baseline treatment; errors are clustered on the subject-level

VARIABLES	In-group voting	Partisan voting
q	0.68*** (0.140)	1.10*** (0.208)
Order of q	-0.08 (0.135)	-0.25 (0.199)
Round	0.01 (0.015)	0.03 (0.022)
Constant	0.85*** (0.138)	-0.74*** (0.246)
Observations	3,330	1,110

Subject-level clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table B.5: Linear least squares regression of allocation to in-group extreme and allocation to center on the probability of the extreme policy preferences (q) and round of play in the baseline treatment. Note, the round of play variable already captures variation in the order of probabilities; errors are clustered on the subject-level

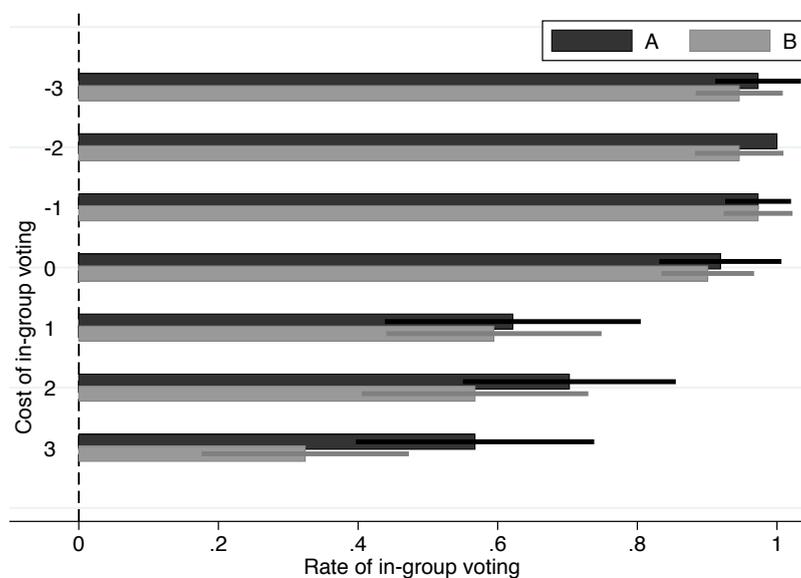
VARIABLES	Allocation to in-group extreme	Allocation to center
q	5.23*** (0.419)	-6.44*** (0.465)
Round	-0.01 (0.008)	0.01 (0.009)
Constant	1.70*** (0.258)	7.25*** (0.321)
Observations	370	370
R-squared	0.395	0.489

Subject-level clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

B.4 Robustness of findings to difference in groups A vs B

Figure B.1: Rate of in-group voting in the A vs B-treatment with polarization in policy preferences of 1 (full polarization) by in-group ability difference separated by group membership in A vs B



B.5 Robustness of findings in a multi-identity treatment

Our finding of instrumental motivations for group identity-driven behavior may be undermined by the weakness of identities induced in the laboratory: for any expressive motivation to surface, a stronger, more contextualized identity may be needed. We implement the Bike vs Car-, Dem vs Rep-, and multi-identity-treatments exactly for the purpose of inducing, even if only marginally, stronger identities. Partisan voting and partisan allocation emerges as well, as expected, when

Table B.6: Linear least squares regression of partisan voting on the probability of extreme policy preferences (q), the order in which q is shown to subjects, round of play, and other relevant variables: treatment, FSU vs Berlin sample, high negative ability difference, strong partisan identity (recovered in the exit survey from the FSU sample) on selected samples.

VARIABLES	(1) Baseline	(2) Baseline	(3) All treatments	(4) Dem vs Rep treatment	(5) Dem vs Rep treatment
q	0.2570*** (0.057)	0.2783*** (0.055)	0.2647*** (0.048)		0.1669** (0.064)
Bike vs Car			0.0187 (0.065)		
Dem vs Rep			0.0311 (0.071)		
Bike vs Car \times q			-0.1538** (0.073)		
Dem vs Rep \times q			-0.0745 (0.068)		
FSU sample	-0.0027 (0.092)				
FSU sample \times q	0.0196 (0.103)				
High negative ability difference		-0.1706*** (0.042)			
High negative ability difference \times q		-0.0419 (0.055)			
Strong partisan identity					0.1256 (0.140)
Strong partisan identity \times q				0.3058* (0.156)	0.0886 (0.089)
Order of q	-0.0597 (0.047)	-0.0370 (0.045)	-0.0279 (0.030)	0.0494 (0.057)	0.0404 (0.056)
Round	0.0067 (0.005)	0.0041 (0.005)	0.0030 (0.003)	-0.0065 (0.007)	-0.0055 (0.006)
Constant	0.3233*** (0.063)	0.3700*** (0.061)	0.3108*** (0.049)	0.3526*** (0.075)	0.2698*** (0.085)
Observations	1,110	1,110	2,550	555	555
R-squared	0.037	0.070	0.026	0.035	0.046

Subject-level clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table B.7: Logistic regression of partisan voting on the probability of extreme policy preferences (q), the order in which q is shown to subjects, round of play, and other relevant variables: treatment, FSU vs Berlin sample, high negative ability difference, strong partisan identity (recovered in the exit survey from the FSU sample) on selected samples.

VARIABLES	(1) Baseline	(2) Baseline	(3) All treatments	(4) Dem vs Rep treatment	(5) Dem vs Rep treatment
q	1.0728*** (0.248)	1.1391*** (0.236)	1.1040*** (0.207)		0.7449** (0.295)
Bike vs Car			0.0778 (0.295)		
Dem vs Rep			0.1376 (0.318)		
Bike vs Car \times q			-0.6268** (0.316)		
Dem vs Rep \times q			-0.3156 (0.295)		
FSU sample	-0.0125 (0.420)				
FSU sample \times q	0.0818 (0.451)				
High negative ability difference		-0.8530*** (0.215)			
High negative ability difference \times q		0.0149 (0.259)			
Strong partisan identity					0.5665 (0.598)
Strong partisan identity \times q				1.2630* (0.673)	0.2967 (0.393)
Order of q	-0.2517 (0.200)	-0.1589 (0.197)	-0.1175 (0.125)	0.2158 (0.249)	0.1804 (0.244)
Round	0.0282 (0.022)	0.0177 (0.022)	0.0127 (0.014)	-0.0284 (0.028)	-0.0246 (0.027)
Constant	-0.7372*** (0.281)	-0.5304** (0.261)	-0.7892*** (0.220)	-0.9918** (0.404)	-0.6097* (0.326)
Observations	1,110	1,110	2,550	555	555

Subject-level clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

inducing stronger identities but we do not see a large increase in the strength of these in-group biases in vote choices and allocation decisions.

A more valid test of whether subjects attach higher salience to a stronger identity may be to put two dimensions of group identities that supposedly differ in their strength side-by-side and let subjects choose which one they emphasize in their decision-making. We do so by letting subjects decide whether they want to put more emphasis on the distinction A vs B or Bike vs Car, as we chose to implement in our multi-identity treatment. In this treatment, subjects are characterized by a compound identity: A-Bike, B-Bike, A-Car, or B-Car. We find, first, while partisan voting for candidates who share the exact identity with the voter is larger than the one for those who share the A/B- or Bike/Car-identity, in-group allocations are smallest for this candidate. Second, and more importantly, there is, again, a (weak) increase in partisan voting and in partisan allocations with increasing policy preference polarization.

In this multi-identity treatment, voters are labelled A-Bike-, B-Bike-, A-Car-, or B-Car-voters according to their group assignment. Assignment to *Bike* or *Car* takes place before subjects learn the content of the voting game and assignment to A or B occurs after subjects received instructions for the voting game but before the voting game commences. Candidates are asked to allocate up to 16 tokens to the five positions of a preference space: four positions in the extreme corners, *Left-Up*, *Right-Up*, *Right-Down*, or *Left-Down*, and one position at the *Center*. Voters decide whether to vote for one of two candidates in two possible pairings of candidates: *A-Bike* vs *B-Car* or *A-Car* vs *B-Bike*. The probability with which voters are at the extreme of the preference space is either .25 or .75. Voters are, again, also told the ability and group membership of both of the candidates. The decision environment is described not only by the probability with which voters are at the extreme (.25 or .75) but also by which pairing of candidates they can choose from: A-Bike vs B-Car or A-Car vs B-Bike. Here, the probability level is fixed for 18 rounds while the pairing of candidate is the same for 9 rounds (while the level of ability varies of those 9 rounds). For payoffs, a decision situation is picked at random: it is represented by a probability, candidate abilities, and a candidate pairing. In this treatment, voters will receive one token for each token that was allocated to the extreme of the preference space associated with their group and half a token for each token that was allocated to a position in the preference space contiguous to the extreme. That is, e.g., a A-Bike-voter receives 1 token from each 1 token allocated to Left-Up and .5 token from each 1 token allocated to Left-Down, Right-Up, and Center. Similarly, A-Car-, B-Bike-, and B-Car-voters receive a full 1 token from each 1 token allocated to their extreme corner but only .5 token from each 1 token allocated to the corners contiguous to theirs and to the Center.

In 2 sessions with 24 subjects each, we collect, for each subject, 36 observations as voter and 4 observations as candidate. In total, we collect 1728 voter-round observations in the multi-identity treatment. The total number of candidate-round observations is 192.²¹

Figure B.2 and B.3 compare in-group voting for and in-group allocations of candidates who share the exact identity of the voter, those who share their A/B-identity, and those who share their Bike/Car-identity. First, while partisan voting for candidates who share the exact identity with the voter is larger than the one for those who share the A/B- or Bike/Car-identity, in-group allocations are smallest for this candidate.

²¹In session 1 of this treatment, voters choices were not recorded correctly in the payment file. We therefore paid the for subjects in the role of voters maximum possible earning of 16 tokens slightly increasing the average payoff of that session relative to the other session of that treatment. Subjects only learned at the moment of payout that this error occurred, in this way, their choices throughout the experiments could not have been affected.

Figure B.2: Rate of partisan voting by policy preference polarization and type of shared identity – exact, A/B, or Bike/Car-identity – for negative ability difference between in- and out-group candidate

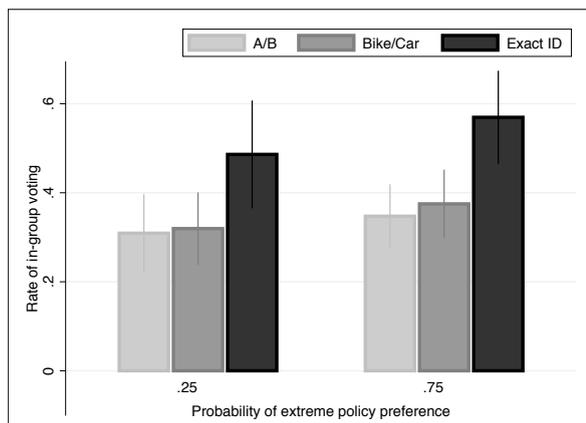
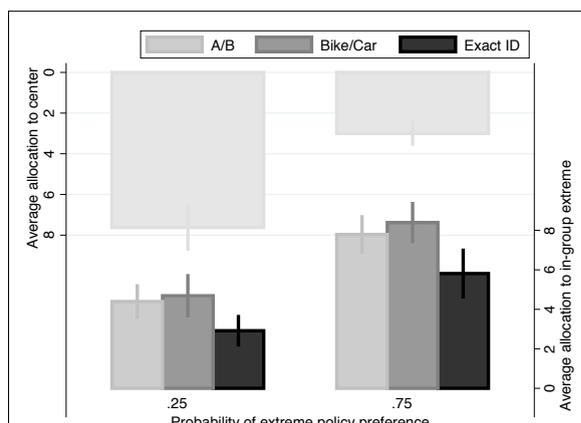


Figure B.3: Rate of allocation to extreme and center by policy preference polarization and type of shared identity – exact, A/B, or Bike/Car-identity



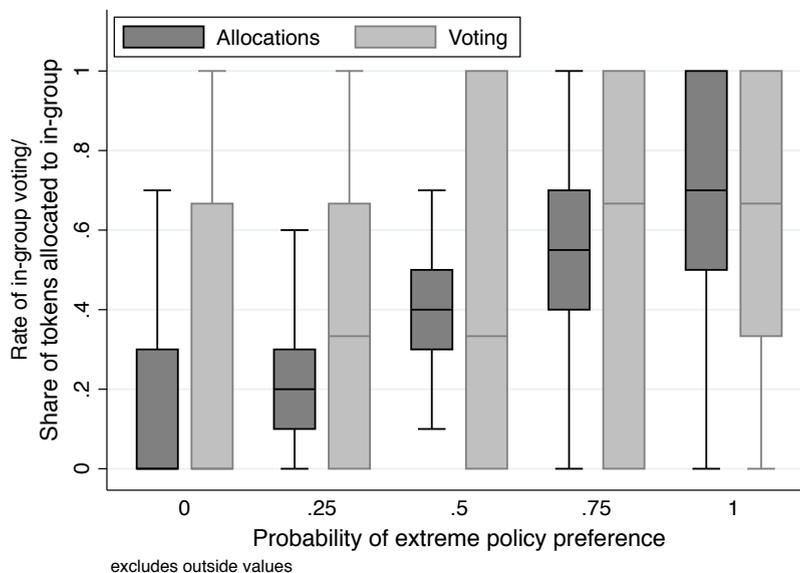
Second, and more importantly, there is, again, a (weak) increase in partisan voting and in partisan allocations with increasing policy preference polarization. The change in partisan voting, however, is not significantly different from zero, i.e., the marginal effect of raising the level of policy preference polarization from .25 to .75 heightens the rate of voting for the in-group candidate who exactly matches the voters identity by .03 % (-.03,.08). In contrast, more polarization in policy preferences systematically increase allocations to all types of in-group voters, i.e., those who share an exact, A/B, or Bike/Car-identity with the candidate. A change in polarization from .25 to .75 raises in-group allocations by 2.53 (1.66,3.41) of the candidates who share the exact identity, 3.11 (2.21,4.00) of those who share the A/B-identity, and by 3.33 (2.38,4.28) tokens of those who share the Bike/Car-identity. More polarization also decreases allocation to the center by 4.43 (3.29,5.57) tokens.

B.6 Robustness of findings to subject-level idiosyncrasies

We now turn to the question whether the observed patterns are robust to subject-level effects. In other words, are results driven by a small subset of subjects or do we observe a general tendency in behavior. We find, in the aggregate, that partisan voting is driven by instrumental motivations rooted in group identities; in-group voting varies with the level of polarization in policy preferences and is, in this pattern, consistent with candidates' in-group favoritism when making allocation decision To characterize subjects individual-level behavior, we investigate whether (1) the distribution of individual-level choices is rather narrow around the overall treatment mean, (2) they condition their choices as voters on their behavior as candidates and whether, and (3) the prevalent behavioral type is one whose choices are driven by instrumental motivations.

Figure B.4 shows the distribution of subject-level rates of partisan voting and subject-level average allocations to the in-group extreme (Left for candidates in Group A/Bike/Dem and Right for candidates in group B/Car/Rep). The spread of subjects' behavior is quite narrow for allocations and follows the aggregate pattern of increasing allocation to the extreme with increasing polarization. The overall average allocation as well as most of the weight in the distribution of subject's average allocation are moving upwards with increasing probability of the extreme policy preference.

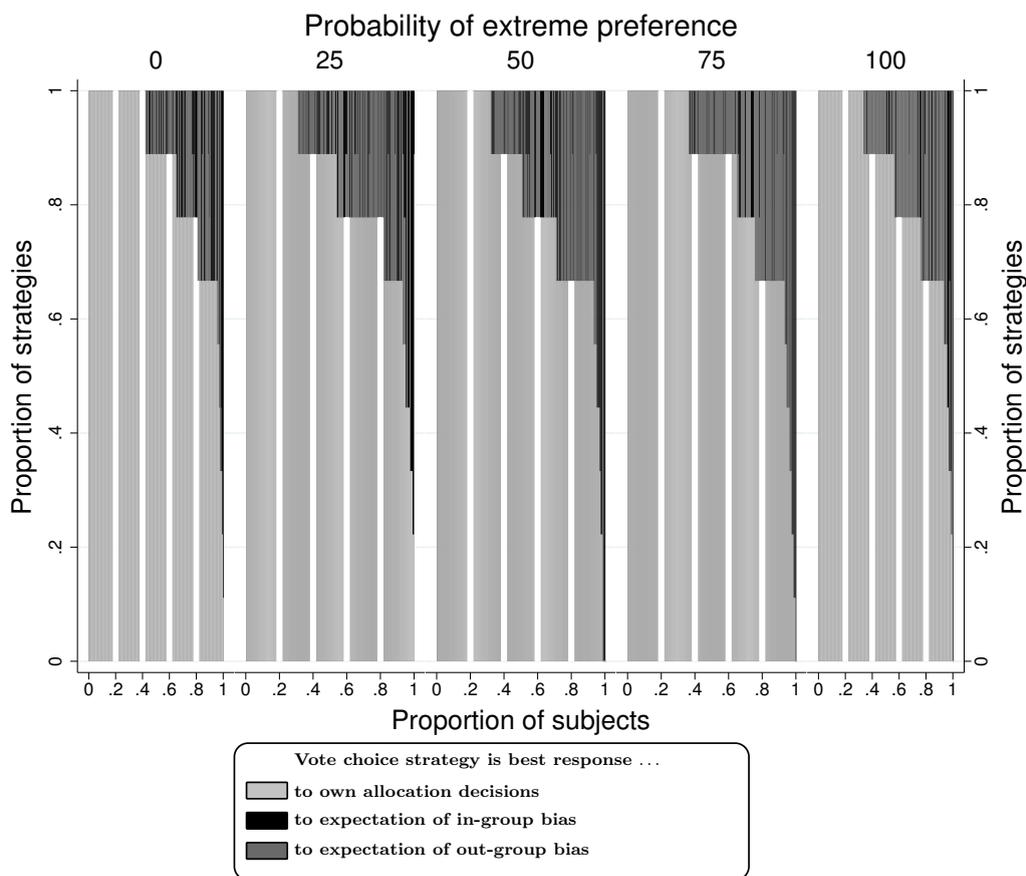
Figure B.4: Subject-level: partisan voting and average allocation of tokens to in-group extreme over the probability of extreme policy preference



The subject-level pattern for in-group voting is less clear. While the average over subject-level rates of in-group voting are clearly increasing with rising polarization, the spread of this subject-level quantity is certainly wider than what we observe for allocations.

Why is that and is there widespread misfit within subject between allocations and voting behavior? The answer is no. A good approximation of what voters believe about candidates' behavior is how they themselves chose to allocate. How do voters vote when we fill in their own allocation choices in each decision situation (characterized by a probability of extreme preferences and abilities assigned to each candidate)? Is their vote choice strategy a best response to their own allocation decisions paired with the ability difference between candidates? Figure B.5 shows that for most subjects it is. 30-40% of subjects always vote for an in-group (out-group) candidate whenever ability difference and their own allocations would give them higher utility than voting for the out-group (in-group) candidate. About 90% of subjects best respond in at least 2/3 of their choices. This is indicated by the dominance of the light gray area in Figure B.5.

Figure B.5: Subject-level: proportion of a vote choices strategy which is a best response to the subjects' own allocation decisions, proportion of in-group votes where expected utility from out-group vote was higher, and proportion of out-group votes where expected utility from in-group vote was higher. Each spike represents behavior of one subject.

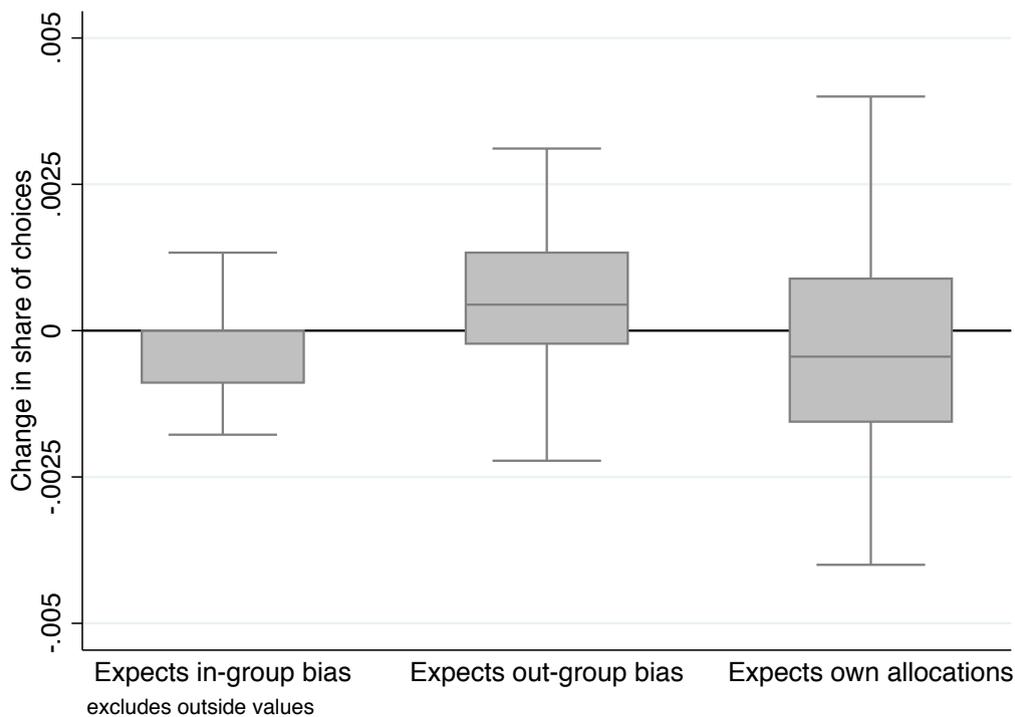


Between 60-70% of subjects sometimes vote for the in-group candidate even though they would obtain a lower payoff than voting for the out-group candidate substituting candidates' allocation decisions with their own allocations. Only a handful of subjects, though, does so in more than 1/3 of their choices. Such behavior describes a strategy that is best responding to the expectation that other subjects, in the role of a candidate, allocate more favorable to the in-group than oneself did. Similar numbers and interpretations are obtained for subjects who sometimes vote for the out-group candidate even though they would obtain a lower payoff than voting for the in-group candidate. Such strategy is optimal if the expectation is that other subjects are allocating less favorable to the in-group than oneself did.

Also, with increasing polarization, those subjects who do not best respond to their own allocations all the time, tend to vote for the out-group too often while for low probability of voters' preference being extreme. We see a mix of voters voting for the in-group too often and voters voting for the in-group not often enough, given their own allocation decisions accounting for ability differences.

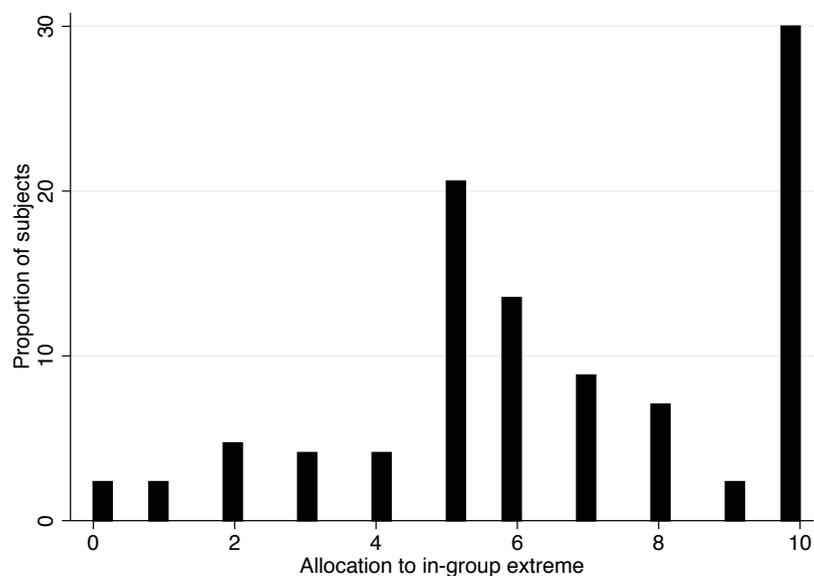
Figure B.6 shows that the profile of strategies within subject does not change with increasing polarization. For the bulk of subjects, increasing polarization by 25% increases the share of any of the strategies by less than 1%.

Figure B.6: Subject-level: distribution of estimated change in share of subject choice type (best responding to own allocations, should vote for in-group but votes for out-group, should vote for out-group but votes for in-group) when polarization increases. Estimate is coefficient on polarization (probability of extreme preference) of regression of share of subject choice type on polarization..



Further, a measure of subjects' underlying in-group bias is also their allocation decision when the probability that voters have extreme preferences is 1. In this case, distributing tokens between left, right, and center is a divide-the-dollar-game. Whatever a candidate places in tokens on that extreme associated with his or her group assignment, solely benefits his or her fellow group members in the role of voters. Figure B.7 gives the distribution of subjects' in-group bias score determined as their allocation to the extreme in-group position at full polarization.

Figure B.7: Distribution of subjects by in-group bias in allocations. In-group bias is measured as the allocation to the extreme in-group position when the probability of an extreme ideal point for voters is 1



In fact, that ability to express in-group favoritism in allocations, which increases with the level of polarization in voters' preferences, also has the effect that subjects, when making voting decisions, are more likely to best respond to their own allocations in the role of candidates. Each subject-level increase by one token in allocations to the in-group when voters' preferences are extreme with a probability of 1 leads to a increase in the rate at which that subject best response in voting to the own allocation choices by 1% (0.4%,1.7%). Partisan allocations, in other words, are most crystalized at full polarization.

C Experimental design

C.1 Decision environment by round

The placeholder in the table, e.g. the probabilities “1” to “5” in column 4 were filled with either 0, .25, .5, .75, or 1 for different subjects in random order.

Table C.1: Choices of subjects in the role of voters and candidates in the voting game by round. Probability of the extreme preference (0, .25, .5, .75, .1) are randomly assigned to each subject by block of 9 rounds each

Round	Voters	Candidates	Probability Block
1	Vote 1	Allocation 1	1
2	Vote 2	–	
⋮	⋮	–	
9	Vote 9	–	
10	Vote 10	Allocation 2	2
11	Vote 11	–	
⋮	⋮	–	
18	Vote 18	–	
19	Vote 19	Allocation 3	3
20	Vote 20	–	
⋮	⋮	–	
27	Vote 27	–	
28	Vote 28	Allocation 4	4
29	Vote 29	–	
⋮	⋮	–	
36	Vote 36	–	
37	Vote 37	Allocation 5	5
38	Vote 36	–	
⋮	⋮	–	
45	Vote 45	–	

C.2 Instructions: baseline treatment

Introduction

In this experiment you will make a series of choices. At the end of the experiment, you will be paid according to your choices and the choices of other participants. Pay close attention to the instructions because each of your decisions potentially affects your payoff from this experiment. This experiment has two parts. Your total earnings will consist of a show-up fee of 7 Dollars and your earnings from each of the two parts of the experiment. During the course of the experiment you will earn tokens, which will be exchanged into Dollars at the end of the experiment at a rate of

$$1 \text{ Token} = 60 \text{ Cent.}$$

We will start with a brief instruction period and Part 1 of the experiment. You will then receive instructions for Part 2 of the experiment and finish that part accordingly. Should you have questions while I read out these instructions, please raise your hand and after I have finished reading the instructions, I will come and assist you. Should you have questions during the

experiment, please raise your hand at any time.

Part 1

In part 1 of the experiment, you will make **50 decisions**: 45 in the role of a **Voter** and 5 in the role of a **Candidate**.

Assignment to Group A and Group B

At the beginning of the experiment you will be randomly assigned to either **Group A** or **Group B**. You will remain a member of this group until the end of the experiment; that is, until you have made all 50 decisions.

Decisions as Voter

As a voter you will make 45 decisions. In each decision, you will be asked whether you prefer **Candidate A, who is a member of Group A, or Candidate B, who is a member of Group B**.

While you are making your decisions, you will see the following information on the screen:

1. The level of **Ability** of Candidate A and Candidate B;
2. The probability with which your **Position** is either **Left, Center, or Right**.

The level of ability of Candidate A and Candidate B is either **Low, Average, or High**.

Importantly, if you are assigned to Group A, your position can only be **Left** or **Center** and if you are assigned to Group B, it can only be **Center** or **Right**.

Additionally, when you are assigned to Group A, the probability that your Position is Left, and not Center, in the voter decisions is either 100%, 75%, 50%, 25% or 0%. When you are assigned to Group B, the probability that your Position is Right, and not Center, in the voter decisions is either 100%, 75%, 50%, 25% or 0%.

To assist you, you will see those probabilities on your screen while you make your decision.

Decisions as Candidate

As a Candidate you will make 5 decisions. In each decision, you will be ask how you want **to allocate up to 10 Tokens to the positions Left, Center, or Right**. You may allocate all 10 Tokens to one position, allocate them in any way over two or three positions, allocate less than 10 Tokens, or do not allocate tokens at all.

While you are making your decisions as a Candidate, you will see on the screen **the probabilities with which the Voters are distributed over the three positions Left, Center, and Right**.

That is, for each candidate decision, you will learn whether the Voters in **Group A** are always on Position **Left** (100%) and never on **Center** (0%), mostly Left (75%) and rarely Center (25%), equally likely Left (50%) and Center (50%), rarely Left (25%) and mostly Center (75%), or never Left (0%) and always Center (100%). You will receive similar information about the probability distribution of Voters in **Group B**, differing only in the fact that those Voters are either on Position **Right** or Position **Center**.

To illustrate the distribution of probabilities of Voters in each decision, you will see them on

your screen while making your decisions as a Candidate.

Earnings in part 1 of the experiment

After you have made 45 decisions as a Voter and 5 decisions as a Candidate, your earnings will be calculated as follows:

1. One participant in Group A and one participant in Group B is randomly selected as Candidate A and Candidate B.
2. One of the distribution of probabilities of the Voters Positions is randomly chosen – every distribution is equally likely to be picked.
3. Either a low, average, or high ability is randomly assigned to the two participants who are selected as candidates – each level of ability is equally likely to be picked.
4. The choices of voters in the decision situation that corresponds to the randomly chosen levels of ability of the two candidates, and the randomly picked distribution of probabilities of the Voters Positions, is used to determine whether Candidate A or Candidate B has won the election.
5. The candidate who received a majority of votes of the participants who were not picked to be the candidates is the Winner of the Election; should both candidates receive the same number of votes, a fair coin is tossed to determine the winner.
6. The assigned ability and the allocation of tokens of the **Winner of the Election** now serves as basis for the earnings of the remaining participants. The two participants who were selected to be the candidates are paid according to the outcome of the election.

Should you be picked to be one of the Candidates for the purpose of determining earnings, your earnings will be

15 Tokens

if you are the Winner of the Election, but only

5 Tokens

if you are **not** the Winner of the Election.

If you are picked to be one of the Voters for the purpose of determining earnings, your earnings will be

As a Voter in Position Left:

Income based on the ability of the Winner of the Election + Tokens Left + (Tokens Center)/2

As a Voter in Position Center:

Income based on the ability of the Winner of the Election + Tokens Center + (Tokens Left + Tokens Right)/2

As a Voter in Position Right:

Income based on the ability of the Winner of the Election + Tokens Right + (Tokens Center)/2

The **Income based on the ability of the Winner of the Election** is **5 Tokens** when the assigned ability is **High**, is **3 Tokens** when the ability is **Average**, and is **2 Tokens** when the ability is **Low**.

Here is an example of how your earnings are calculated. After all participants have made all of their decisions, two participants, one from Group A and one from Group B, are randomly chosen to be Candidate A and Candidate B, respectively. Now, suppose Candidate A from Group A is randomly assigned a low ability and Candidate B from Group B is randomly assigned a high ability. Further, suppose the randomly chosen probability distribution specifies that the Position of Voters in Group A is Left with a probability of 75% and Center with a probability of 25%, and that the Position of Voters in Group B is Right with a probability of 75% and Center with a probability of 25%.

Additionally, assume that in the decision situation with this probability distribution, a Candidate A with low ability, and a Candidate B with high ability, a majority of Voters prefer Candidate B. Further suppose that the participant chosen to be Candidate B allocated 7 Tokens to Position Left, 2 to Position Center, and 1 to Position Right. In this way, the participant who was chosen to be Candidate B receives 15 Tokens as Winner of the Election, and the participant who was chosen to be Candidate A receives 5 Tokens for losing the election. Moreover, of the remaining participants, those chosen to be Voters in Position Left receive $5 + 7 + 2/2 = 13$ Tokens, in Position Center receive $5 + 2 + (7 + 1)/2 = 11$ Tokens, and in Position Right receive $5 + 1 + 2/2 = 7$ Tokens.

You will receive instructions for part 2 on your screen at the beginning of that part.

Again, your total earnings in this experiment will consist of a show-up fee of 7 Dollars and your earnings in part 1 and part 2 of the experiment.

C.3 Screenshot

Figure C.1: Voter decision screen

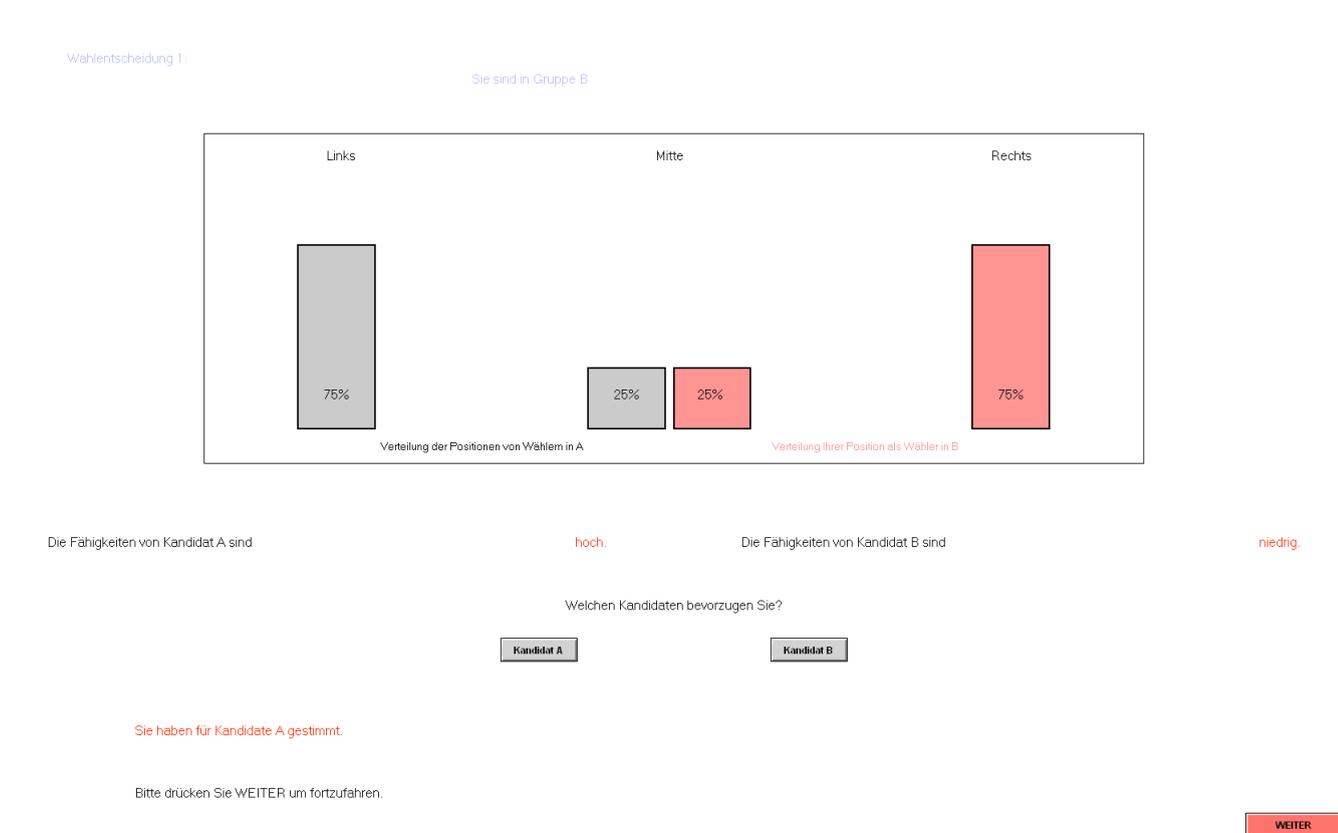
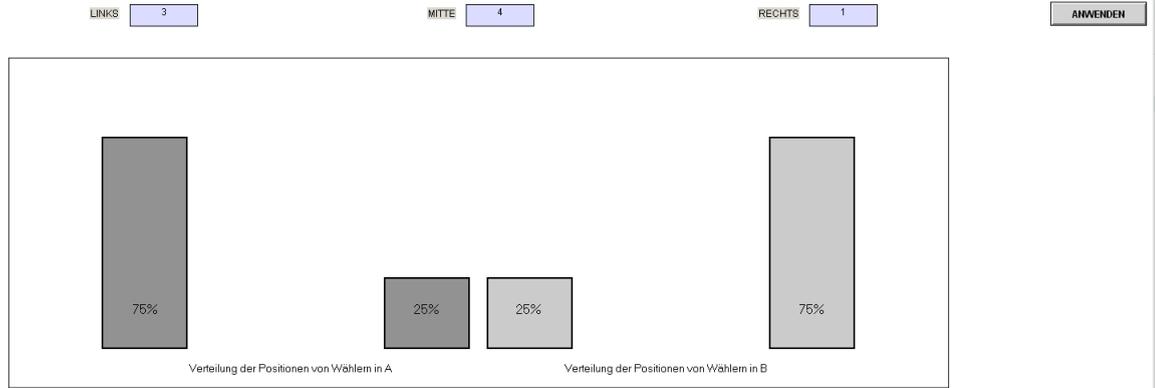


Figure C.2: Candidate decision screen

Verteilungsentscheidung 1:

Sie sind in Gruppe B

Wie möchten Sie die 10 Tokens verteilen?



Bitte drücken Sie WEITER um fortzufahren.