Tom Lane, Luis Miller,
Isabel Rodriguez

May 2023

The normative permissiveness of political partyism

CEDEX

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit http://www.nottingham.ac.uk/cedex for more information about the Centre or contact

Samantha Stapleford-Allen
Centre for Decision Research and Experimental Economics
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: +44 (0)115 74 86214
Samantha.Stapleford-Allen@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx

The normative permissiveness of political partyism

Tom Lane[a], Luis Miller[b1], Isabel Rodriguez[bc,]

[a] School of Economics, University of Nottingham Ningbo China, 199 Taikang Middle Road, Ningbo, China; [b] Spanish National Research Council (IPP-CSIC), C/ Albasanz 26, 28037 Madrid, Spain; [c] International Doctorate School of the Spanish National Distance Education University (EIDUNED).

Abstract

Political identity has become the strongest social divide within Western societies. This paper employs experiments to measure discrimination along multiple dimensions of social identity, and replicates previous findings showing the strongest discrimination against out-groups occurs in the political domain. Moreover, we explore a possible explanation for this phenomenon based upon social norms. We measure the social appropriateness of discrimination along each identity dimension. The ranking of dimensions by discrimination against out-groups reflects the extent to which such behaviour is normatively permissible, with the weakest anti-discrimination norms on the political dimension. Results are qualitatively similar in two European countries. We argue that, while norms sanctioning discrimination on other dimensions have developed historically, no such process has taken place in relation to political affiliation, bringing political identity to the fore and helping polarisation to flourish.

---

[1] To whom correspondence should be addressed: E-mail: luis.miller@csic.es

1. Introduction

Why has political identity become the most important social divide in the Twenty-First Century? A fast-growing literature in economics and political science has established the primacy of partyism over other relevant social divides, like race, ethnicity, religion, or national identity. Partyism – also labelled party affective polarization (Iyengar et al., 2019) – refers to any form of prejudice or hostility towards voters or supporters of opposing political parties (Sunstein 2015). Observational (mostly survey) and experimental studies on partyism in the United States and many European Countries abound (Iyengar et al., 2019; Iyengar and Westwood, 2015; Westwood et al., 2018). Such research shows people have stronger preferences to engage in economic discrimination against political out-groups than against outgroups along any other dimension of social identity.

Yet despite this proliferation of empirical studies, the underlying theoretical mechanisms explaining the prominence of partyism in contemporary Western societies remain unclear. Most studies resort to the human tendency for tribalism, an evolved predisposition to group conflict that makes us favour and be loyal to *our* groups and hostile to *other* groups with whom we compete (Clark et al., 2019). Unfortunately, while tribalism can explain partyism, on its own it would seem to equally predict such strong ingroup favouritism and outgroup hostility based on other social divides. To explain the dominance of party affective polarization, we need something else beyond our natural predisposition toward being tribal. Here, we argue that the missing piece in the puzzle of the currently exacerbated level of party affective polarization is the role of social norms.

Social norms are the unwritten 'grammar of society' (Bicchieri, 2005), the informal rules which dictate which behaviours should or should not be taken.[2] In recent years, there has been an explosion of interest in economics in social norms. A vast body of empirical evidence suggests that economic behaviour – across a wide range of domains – is guided and constrained by social norms (Allcott, 2011; Fehr and Schurtenberger, 2018; Gachter et al., 2013; Kimbrough and Vostroknutov, 2016; Köbis et al., 2022; Krupka and Weber, 2013; Krupka et al., 2017; Szekely et al., 2021). More specifically related to our paper's research question, there is evidence that the strength of economic *discrimination* is driven by the extent to which norms condone such behaviour (Barr et al., 2018; Coffman et al., 2021; Restrepo-Plaza and Fatas, 2022).

---

[2] See Elster (1989) and Ostrom (2000) for further definitions of social norms.

We argue that, historically, norms have developed which characterize as morally reprehensible discrimination against outgroups along many dimensions of social identity, especially race and gender (Gaertner and Dovidio, 2000). Anecdotal evidence for this comes from the social sanctions imposed upon public figures caught expressing racist or sexist remarks which might have gone unpunished in previous decades. These norms act as a constraint on discriminatory behaviour. Our conjecture, however, is that Western societies have not evolved corresponding social pressures or sanctions that mute disapproval of political opponents. In today's world, it is not uncommon to find left-wing opponents of the British government denouncing its supporters on Twitter as 'scum' or boasting that they have 'never kissed a Tory' (Cohen, 2016). While such actions would seem likely to trigger backlash if taken against outgroups on other domains, we argue that lax norms along the political identity domain have left partisans free to express animus and engage in discriminatory behaviour toward their opponents. The absence of strong norms regulating partyism may explain its development. Continued permissiveness of political hatred, at the same time as the development of norms against other types of hatred, may have redirected group animosity that would otherwise be expressed along other dimensions of identity, but where it is no longer acceptable, towards the political domain.

In this article, we provide empirical evidence to support our argument. This comes from a set of laboratory experiments designed to study the role of social norms in political and other types of social intergroup interactions. We focus specifically on interpersonal discrimination, which can be regarded as one manifestation of intergroup hostility and, on the political domain, one of the "dark consequences" of polarization (Finkel et al., 2020). Previous studies have reported discrimination based on political affiliation in the labour market, in college admissions, and in everyday economic interactions (Dimant, 2023; Iyengar et al., 2019; Michelitch, 2015). In our experiments, we employed a standard norm-elicitation task to establish the social appropriateness of economic discrimination along different dimensions of social identity, as well as an incentivized task to measure such discrimination on each dimension.

We hypothesize that, like in previous research, subjects will discriminate less on the basis of other social identities, such as religion, than on the basis of political party support. Moreover, our conjecture is that participants will perceive it to be correspondingly less socially inappropriate to discriminate in the political identity context. In our experiments run in both the UK and Spain, we divided participants into groups on the basis either of their support for political parties, their

religious affiliation, or a trivial characteristic (colour of the ball the participants randomly picked during the experiment) representing a 'minimal group' identity (Tajfel, 1970). The dimension of identity upon which we divided participants represents our treatment manipulation. Having formed these groups, we implemented a standard distribution game in which participants allocated money between ingroup and outgroup individuals, therein facilitating the measurement of discrimination in favour of the ingroup over outgroups. In each treatment, for every possible level of discrimination a participant could enact in the allocation task, we have a corresponding estimate from the norm-elicitation task of the perceived social appropriateness of this behaviour.

Results are in line with our hypotheses. We observe significantly stronger levels of discrimination based on political identity than on either of the other identity dimensions. We also find significant differences in norms across treatments: discrimination in favour of one's ingroup is perceived to be significantly less inappropriate when the groups are based on political parties than either of the other two identity dimensions. The relatively permissive norms regarding political discrimination emerge as a compelling explanation for the relative prominence of partyism, over other social divisions, that has been identified in the existing literature.

Our findings are drawn from experiments following a design similar to that introduced in a previous paper, Barr et al. (2018). That study measured discrimination between groups – and the social norms pertaining to such behaviour – based on either nationality or minimal-group identity. Like the current paper, Barr et al. found strongest discrimination on the identity dimension for which the social norms against discrimination were weakest. The purpose of their study was to identify social appropriateness as a potential moderator of discrimination in general, as well as to explain a surprising result in the previous literature: that discrimination in lab experiments tends to be stronger between minimal groups than between groups based on various types of natural identity (Lane, 2016). Our study expands upon this, finding that the ranking of discrimination across *three* identity dimensions matches that of the normative permissibility of discrimination, again identifying fairly strong levels of discrimination between minimal groups and showing that this non-intuitive result can be explained by the relatively weak norms against such discrimination.

The main purpose of our study, however, differs from Barr et al. in that we are focused primarily on explaining the strength of discrimination along one particular dimension: political party identity. The other two identity dimensions are included in our design as benchmarks against which

4

to compare both political discrimination and the social norms relating to such. That political discrimination and its normative permissiveness outrank the equivalent phenomena on *two* other identity dimensions – one representing a strong natural characteristic (religion) and another (minimal group identity) which has been found empirically to yield strong effects in practice – provides strong evidence for our argument that political party support is special among group identity dimensions. Note that our experiment does *not* show that social norms *favour* discrimination against political opponents; the norm-elicitation task reveals that, along all dimensions of identity, the most socially appropriate behaviour is to provide equal treatment to ingroups and outgroups. However, there is lesser agreement about this on the political dimension, and greater acceptance towards those who deviate from equality. Our results suggest that those who discriminate against political opponents may escape with only mild social sanctions, resulting in discriminatory actions taking on a more attractive cost-benefit profile than they would in other social arenas. Our findings are qualitatively similar in the two countries the experiments were run, lending credence to their generalisability across different political contexts.

2. Methods

2.1. Research Design

Our experiments expand upon the design of Barr et al. (2018). Like in their study, our approach relies upon measuring both discrimination between groups and the social norms pertaining to such discrimination. Between treatments, we exogenously vary which dimension of social identity the experimental groups are based upon. We employ a simple economic task to measure the extent to which subjects discriminate in favour of in-groups over out-groups on the dimension of identity relevant in their treatment. We also employ a separate norm-elicitation task to measure the perceived social appropriateness of discrimination against out-groups on the relevant identity dimension. This will enable us to test for treatment differences in the level of discrimination resulting from forming groups on the basis of different identity dimensions, and also for corresponding treatment differences in the perceived social appropriateness of discrimination. We can therefore explore whether differences in the strength of discrimination along different identity dimensions can be explained by equivalent differences in the strength of social norms prohibiting or allowing such discrimination.

In Barr et al., groups were based upon either nationality or an artificially induced identity. The current experiments differ in that our treatments employ different dimensions of social identity, with groups constructed on the basis of either religion, artificial identity, or support for political parties. We also depart from Barr et al. in that, along a given identity dimension, subjects are divided into four (rather than two) identity groups, and for members of each group we estimate three measurements of discrimination in favour of their in-group, separately against each of the three out-groups they face; correspondingly, we also separately elicit the social appropriateness they perceive of discrimination in favour of their in-group over each out-group. Separation of subjects into four groups more naturally reflected the real-world structure of the social identity dimensions employed in our experiment. Note, however, that we are not primarily interested in drawing comparisons between the behaviour of different groups within an identity dimension, but rather in making aggregate-level comparisons across dimensions, to which end the division of subjects into groups was an essential design feature.

We conducted two experiments: one in the United Kingdom, at the University of Nottingham, from December 2022 to January 2023, and a second one in Spain, at the University Jaume I in Castellón, in April 2023. Both experiments were pre-registered.[3] Below, we outline the design of the first experiment in the UK, and in the process highlight any deviations from this design in the implementation of the second experiment in Spain. Note that our study is not an attempt to conduct a controlled cross-cultural comparison between the UK and Spain. In particular, the religious and political identity dimensions are differently composed in each country with their own groups and local characteristics. Rather, we are interested in the results in each country in their own right, as well as in identifying whether they allow us to draw qualitatively similar conclusions.

2.2. Discrimination Measurement Task

To measure discrimination, we used a simple third-party allocator game. See Barr et al. (2018) for a discussion of the advantages of this task in our research context. In the game, a decision-maker is endowed with £16 (€16 in Spain) and required to divide it between two passive players, one

---

[3] The pre-registration can be found at https://www.socialscienceregistry.org/trials/10172 and https://www.socialscienceregistry.org/trials/11275. Ultimately, our execution of the experiments deviated slightly from the initial plan in that we were practically unable quite to attain the desired sample sizes and the number of sessions, but otherwise did not depart from it.

belonging to the decision-maker's in-group and the other belonging to an out-group.[4] The money can be split however the decision-maker prefers, as long as the amount given to each player is a multiple of two. The interaction is anonymous, and the decision-maker receives no information about either of the passive players except for their group identity.

In order to maximize sample sizes, we employed two elements of randomization. First, while only one third of subjects would be allocators and the other two thirds would be passive players, all subjects were required to make decisions in the role of allocator. It was made clear that, after the end of the experiment, it would be randomly determined which subjects had been assigned to the allocator role, and the decisions of those not assigned to it would be discarded. Secondly, while subjects were told that – if assigned to the role of allocator – they would definitely be matched with one in-group and one out-group player, they were not informed the specific group identity of the out-group player. This would be randomly determined after the allocator role assignments. In the meantime, subjects were required to commit to three allocation decisions, one for each of the possible groups the out-group player might belong to; when it was later determined which out-group an allocator had been matched with, the allocation they had committed to make if matched with a member of this out-group was automatically implemented, while their other two allocation decisions were discarded. Since this randomization approach was relatively complex, care was taken to write the instructions with sufficient detail and clarity to avoid confusion (our full instructions can be found in section I of the Online Appendix, and subjects were required to answer understanding test questions before they could proceed to their allocation decisions). The order of the three allocation decisions was also randomized.

2.3. Norm-elicitation Task

The social appropriateness of discrimination in the allocator game was measured using the well-established norm-elicitation method first introduced by Krupka and Weber (2013). Subjects were described the allocator game and required to evaluate the social appropriateness of each of the different allocations available to the decision-maker, by selecting on a four-point scale one of the following options: 'Very socially inappropriate', 'Somewhat socially inappropriate', 'Somewhat socially appropriate' or 'Very socially appropriate'.

---

[4] Although the decision-maker could not keep any money for their self, they knew they would receive a payment from the task of either £6, £8 or £10, randomly selected with equal probability.

When making evaluations, subjects were informed of the group identity of the decision-maker whose behaviour they were assessing, as well as that of both of the passive players the decision-maker was allocating between. Subjects only evaluated allocations made by decision-makers of their own group. They made three sets of evaluations, one for each of the groups the passive out-group player could belong to. Thus, the task separately reveals each subject's perception of the social appropriateness of any possible level of discrimination by members of their own group towards members of each of the other three groups along the relevant identity dimension in their treatment.

Subjects making the evaluations were the same as those who participated in the allocator games described in the task. Although there may be a concern in principle that participation in either part of the experiment might influence responses to the other, Barr et al. found an absence of evidence for this. Nevertheless, in the UK experiment we randomized across sessions whether subjects first partook in the allocator game or norm-elicitation task, so that we can also check for order effects in the current study.[5] The order in which subjects completed their three sets of evaluations was also randomized (within sessions), but subject to the constraint that the order of the three out-groups was the same for any given subject in the norm-elicitation task as it was in the allocator game.

Evaluations were incentivized. Each subject knew that, at the end of the experiment, one of the actions from one of the three sets of evaluations they had made, would be randomly selected. The subject's evaluation of this action would be compared with that of another randomly selected subject from the same treatment. The subject would receive a bonus of £8 if and only if their evaluation matched that of the person they were compared against. The norm-elicitation task, therefore, takes the form of a coordination game, where subjects are incentivized to provide the same evaluations as others. Importantly, subjects were told that the person to whom their evaluation would be compared would belong to their own identity group.

The incentives are thus designed to guide subjects to reveal perceptions of appropriateness as commonly agreed by members of their own identity group (i.e. group-specific social norms), rather

---

[5] Subjects were not informed about the nature of the second activity until they had completed the first one. However, those who did the norm-elicitation task first were, just like those who first played the allocator game, made aware at the start of the experiment about the four different identity groups that subjects in their session were divided between. Thus, the dimension of identity made relevant in their treatment had already been made salient to all subjects before they began whichever was their first task.

than personal opinions about morality. See Bicchieri (2005) for a discussion of the important distinction between personal opinions and social norms, the latter of which can be regarded as second-order beliefs about the appropriateness of behaviour. The way 'socially appropriate' behaviour is defined to subjects before they undertake a Krupka-Weber task is aimed at conveying to them this concept – in our experiment, the instructions told them to think of it as 'behaviour that you think most participants [of your identity group] in this experiment would agree is the "correct" thing to do.'

A possible concern is that the coordination incentives could lead subjects to report third – or higher–order beliefs. However, if subjects use salient focal points to coordinate, in the manner suggested by Schelling (1980), it is likely they will indeed report second-order beliefs since this is what they are asked to do, and answering the question truthfully seems by far the most salient available strategy. In principle, the Krupka-Weber method could produce responses wholly unrelated to social norms if subjects could find an alternative strategy to coordinate, but existing empirical research suggests this does not happen, even when other plausible focal points are made available (Fallucchi and Nosenzo, 2022). There is also evidence that the Krupka-Weber method provides norm estimates consistent with other methods where alternative coordination strategies are excluded by design (Bicchieri et al., 2022; Lane et al., 2023).[6]

2.4. Treatments

Our three treatments differed according to the dimension of identity that the groups in the experiment were formed on the basis of. This dimension of identity was made salient at the beginning of the experiment by announcing the four groups, along this dimension, that the subjects in the treatment were divided between.

In two of the treatments, the groups were determined prior to the experiment based on real-world identity characteristics. Of course, the labels of the groupings differ between the two experiments, on account of the differences between the two countries in which they were run. In the UK, in the *Religion* treatment, we invited subjects who were either *Christian*, *Hindu*, *Muslim,* or *non-religious*. In the *Politics* treatment, the invited subjects were supporters of one of the *Conservative*,

---

[6] For a methodological discussion of the Krupka-Weber method, see Nosenzo and Gorges (2020).

Green, *Labour* or *Liberal Democrat* parties.[7] In the Spanish experiment, in the *Religion* treatment, we invited subjects who were either *Catholic*, *Muslim*, *agnostic*, or *atheist*. In the *Politics* treatment, the invited subjects were supporters of one of the *PSOE* (Partido Socialista Obrero Español – Spanish Socialist Workers' Party), *PP* (Partido Popular – People's Party), Vox (Voice), and *Unidas Podemos* (United We Can) parties.[8] Meanwhile, in the *Artificial* treatment, group identity was created at the beginning of the experiment itself, using a similar method to Barr et al., which itself followed in the long tradition – stemming from Tajfel (1970) – of inducing 'minimal group' identity within an experiment. Upon entering the lab, subjects were instructed to blindly draw a ball from a bag. The colour of the ball – *blue*, *pink*, *red* or *yellow* – determined the subject's group.

We knew subjects' religious or political affiliation from a pre-survey we conducted in the weeks leading up to the lab experiment (see Online Appendix, section A). In the UK, political identity was derived from subjects' response to the question: 'If the next general election were held tomorrow, which party would you vote for?'[9] Given difficulties in Nottingham to attain a pool of eligible subject of sufficient size, in Castellón we used both the voting question and the question 'name the party that you feel closer to than any other'. The pre-survey was distributed to subjects at the Universities of Nottingham and Jaume I registered on ORSEE for participation in economic experiments. From the responses, we identified the four most popular political parties and religious identities to employ as our groups in these treatments (see section A in the Online Appendix). We assigned to the experiment only those who belonged to one of the four chosen groups for both identity dimensions; these subjects were then randomly assigned either to the *Religion* or *Politics* treatment.[10] This ensures that subjects in these two treatments are drawn from the same wider sample, and therefore any differences in the outcomes of the experiment should be driven by the

---

[7] The *Conservatives* are the governing right-wing party in the UK. *Labour* is the main, left-wing opposition. The *Liberal Democrats* and *Greens* are smaller parties (but relatively popular among student populations) occupying, respectively, centrist internationalist and environmental leftist positions.

[8] *PSOE* (centre-left) and *PP* (centre-right) are the two majoritarian Spanish parties. *Vox* is a relatively new far right party, currently third at the national level. *United We Can* is a left-wing coalition.

[9] Subjects were allowed to not select any party to this question if they did not consider any party worth voting for. Thus, we did not force political identity on subjects. In all treatments, subjects were reminded in the lab experiment of their group identity – this seemed particularly necessary in the *Politics* treatment, in case some subjects had forgotten their answer to the voting intention question.

[10] We would ideally have recruited subjects for the *Artificial* treatment also from the same pool of individuals identified by the pre-survey as eligible for participation in both of the other treatments. Unfortunately, however, this pool was not sufficiently large to fill three treatments. See sections A and B of the Appendix for further information on our recruitment approach and related methodological discussion.

treatment manipulation (i.e. dimension of identity) rather than demographic differences between the subjects assigned to each treatment. The most important results of our study should be considered the comparisons between the *Religion* and *Politics* treatments, with the *Artificial-Religion* and *Artificial-Politics* comparisons regarded as suggestive but not perfectly controlled.

2.5. Procedure and Sample

After participating in both the allocator game and norm-elicitation task, each subject was randomly paid their earnings from only one of them, as determined randomly by a coin toss at the end of the session (this rule was made clear to subjects from the outset). All subjects additionally received a show up fee of £4 (€3 in Spain).

The experiment was conducted using Z-Tree (Fischbacher, 2007). In the UK, we ran four sessions for the *Artificial* and *Politics* treatments, and three sessions for the *Religion* treatment; since we randomized across sessions whether the allocator game or norm-elicitation task was run first, we are able to check for order effects. Consistent with Barr et al., we do not find such effects (see section H in the Online Appendix) and will therefore pool our analysis in the next section across ordering conditions. In Spain, we conducted three sessions (one per treatment) and the norm elicitation task was always implemented after the decision task.

Sessions ranged in size from 13 to 29 subjects in Nottingham and from 57 to 60 in Castellon. Table 1 presents the total number of subjects in each treatment. Relative sizes of each identity group can be found in section C of the Online Appendix.[11] Across the two experiments, subjects were 57% female, 23 years old on average, and 11% from rural areas. According to the self-reported ideology variable, 45% are left-wing, 25% centre and 30% right-wing. All subjects in the UK, and the vast majority in Spain, were students. Individual characteristics are similarly distributed between experimental conditions.

---

[11] See section G of the Online Appendix for robustness checks related to some procedural errors made in the sessions in Nottingham.

Table 1: Descriptive characteristics of participants and treatment assignment, by experiment

	Total	**Nottingham, UK**	**Castellón, Spain**
	N=415	N=238	N=177
Age	22.8 (7.6)	21 (2.9)	25.2 (10.7)
Gender			
Male	171 (41.2%)	107 (45.0%)	64 (36.2%)
Female	238 (57.3%)	127 (53.4%)	111 (62.7%)
Other	6 (1.4%)	4 (1.7%)	2 (1.1%)
Background			
Rural	46 (11.1%)	30 (12.6%)	16 (9.0%)
Urban	205 (49.4%)	121 (50.8%)	84 (47.5%)
Mixed	164 (39.5%)	87 (36.6%)	77 (43.5%)
Ideology			
Left	178 (42.9%)	104 (43.7%)	74 (41.8%)
Centre	96 (23.1%)	60 (25.2%)	36 (20.3%)
Right	113 (27.2%)	46 (19.3%)	67 (37.9%)
Missing	28 (6.7%)	28 (11.8%)	0 (0.0%)
Treatment			
Artificial	143 (34.5%)	83 (34.9%)	60 (33.9%)
Politics	145 (34.9%)	85 (35.7%)	60 (33.9%)
Religion	127 (30.6%)	70 (29.4%)	57 (32.2%)

Notes: Distribution of gender, background and ideology of participants. Age is presented as a continuous variable, including the mean and standard deviation. Ideology is measured using a standard 0-10 left-right survey question. *Left* is defined as 0-4, *Centre* as 5, and *Right* as 6-10. Treatment assignment is also included, by location of the experiments.

2.6. Predictions

Following the literature on the primacy of partyism (Westwood et al. 2018) and the previous behavioural study of Barr et al. (2018), we derive two core predictions:

Prediction 1: the social appropriateness of in-group favouritism will be ranked across treatments as following:

Politics > Artificial > Religion

Prediction 2: participants will behave according to social norms when presented with the allocation task.

3. Results

3.1. Nottingham, UK

We start by examining the results from the norm-elicitation task, in which individuals assess the social appropriateness of the range of allocation decisions. There is a clear consensus in recognizing the equal-split choice – allocating 8 monetary units to each participant – as the most appropriate behaviour: 87.27% of participants rate this decision as "very socially appropriate". Yet, variations across treatments can be observed. Table 2 presents the distribution of assessments of social appropriateness given to every possible outcome of the allocation task, by treatment. This pools all assessments made by subjects of each group, regarding allocations towards each of the out-groups. The mean ratings are computed assigning evenly-spaced values of −1 for the "very socially inappropriate" rating, −0.33 in the case of "somewhat socially inappropriate", 0.33 for "somewhat socially appropriate" and 1 for "very socially appropriate", following Krupka & Weber (2013) and Barr et al. (2018).

In the *Politics* treatment, 80.6% of responses rated the equal-split choice as "very socially appropriate", in contrast to 91% in both the *Artificial* and *Religion* treatments. In all treatments, social appropriateness decays rapidly as choices deviate from the equal-split standard toward more discriminatory choices. However, that decline is faster for the *Religion* treatment. The smallest deviation possible from the equal choice is extensively more censured in the *Religion* treatment than in the *Politics* case: 52% of responses deem the (10,6) allocation – giving £10 to the in-group member and £6 to the out-group member – to be either "somewhat socially inappropriate" or "very socially inappropriate", in contrast to 25% in the *Politics* treatment.

In the case of *Religion*, as in the case of the *Artificial* treatment, the inappropriateness of discriminatory choices is largely symmetric, meaning participants evaluate favouring the in-group participant and favouring the out-group participant alike. The exception would be the extreme allocations. 91.9% of responses identify it as "very socially inappropriate" to give all the money to the in-group participant, versus 82.9% that assign that evaluation to the (0,16) split – giving all the money to the out-group participant. Hence, in the *Religion* framework, discriminating in favour of a participant of your group is regarded as equally socially inappropriate, or even more so, than favouring other groups. For the *Politics* framework, the opposite is true. Evaluations in this

Table 2: Distribution of social appropriateness ratings, by treatment (UK)

|	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
Politics treatment									
Very appropriate	13.0	11.1	10.3	20.2	80.6	3.6	0.8	0.4	0.4
Somewhat appropriate	4.7	11.9	31.6	54.9	14.6	43.9	10.3	4.7	4.0
Somewhat inappropriate	9.1	33.6	38.3	22.1	2.8	43.5	52.6	24.1	5.5
Very inappropriate	73.1	43.5	19.8	2.8	2.0	9.1	36.4	70.8	90.1
Mean rating	**-0.62**	**-0.40**	**-0.12**	**0.28**	**0.83**	**-0.05**	**-0.50**	**-0.77**	**-0.90**
Religion treatment									
Very appropriate	1.4	1.4	1.9	5.7	91.0	5.7	4.3	4.3	5.7
Somewhat appropriate	0.5	2.4	11.0	42.4	6.7	41.0	12.4	8.1	5.2
Somewhat inappropriate	6.2	26.2	37.6	26.2	0.5	28.1	31.0	20.5	6.2
Very inappropriate	91.9	70.0	49.5	25.7	1.9	25.2	52.4	67.1	82.9
Mean rating	**-0.92**	**-0.76**	**-0.56**	**-0.15**	**0.91**	**-0.15**	**-0.54**	**-0.67**	**-0.77**
Artificial treatment									
Very appropriate	6.0	6.8	1.6	6.0	90.9	7.1	6.0	6.8	9.1
Somewhat appropriate	2.8	4.8	16.3	65.5	4.4	54.8	9.5	3.6	1.2
Somewhat inappropriate	4.4	27.0	57.5	17.9	0.8	25.8	49.6	21.4	2.8
Very inappropriate	86.9	61.5	24.6	10.7	4.0	12.3	34.9	68.3	86.9
Mean rating	**-0.81**	**-0.62**	**-0.37**	**0.11**	**0.88**	**0.04**	**-0.42**	**-0.67**	**-0.78**

Notes: Percentage of responses corresponding to each of the available social appropriateness ratings for each of the nine possible decisions in the allocator game. Decisions range from allocating all the money to the in-group participant (16,0) to allocating all the money to the out-group participant (0,16). The modal evaluation for each outcome is highlighted. Mean ratings are computed assigning values of 1, 0.33, −0.33 and −1 to assessments of "very socially appropriate", "somewhat socially appropriate", "somewhat socially inappropriate" and "very socially inappropriate", respectively. Therefore, mean ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

treatment are not symmetric: the social appropriateness of discriminatory choices favouring the out-group declines faster than those favouring the in-group. In the extreme splits, allocating all the money to the in-group participant is regarded as less inappropriate on average that allocating all the money to the out-group. Between 10 and 20% of responses rate any discriminatory behaviour

favouring the political group of the allocator as "very appropriate", in contrast to almost none providing the same rating to behaviour favouring the out-group. Choosing to favour the participant with whom you share political affiliation is deemed less inappropriate than favouring the in-group member in any other treatment, or favouring the out-group member in the *Politics* treatment.

Figure 1: Mean evaluation of social appropriateness of allocation decisions, by treatment (UK)
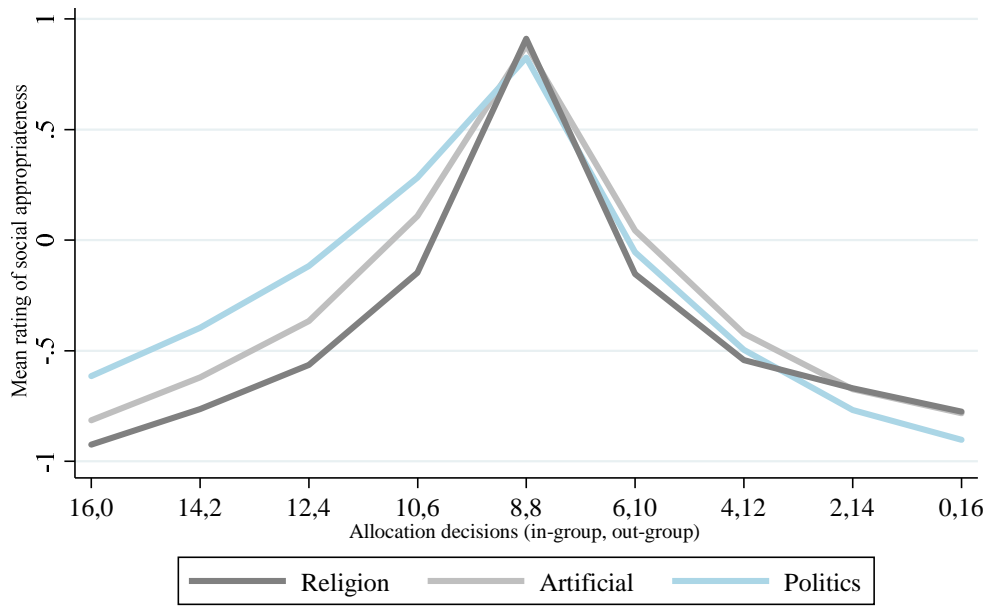
Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

Figure 1 plots the mean ratings of all allocation outcomes by treatment. It is illuminating that for all choices that favour the in-group participants, ratings signal less social inappropriateness in the *Politics* treatment than in the *Artificial* or *Religion* treatment. Table 3 displays the p-values of Fisher-Pitman permutation tests comparing the mean ratings of social appropriateness between treatments.[12] The ranking by treatment of social inappropriateness of favouring the in-group member – the *Religion* treatment having the most inappropriate discrimination and the *Politics* treatment yielding the lowest social inappropriateness levels – remains consistent and significant for all possible choices. In the case of the equal (8,8) split choice, the ranking inverts. Dividing

---

[12] As we perform a total of 54 tests, 9 by treatment for each of the two experiments, the reported p-values are corrected with the Benjamini-Hochberg False Discovery Rate method. This method accounts for the increasing probability of reporting a false result when performing multiple tests. The Benjamini-Hochberg method consists of sorting the p-values in ascending order, multiplying by the total number of tests performed and dividing by the rank number.

the money equally is regarded as significantly less appropriate in the *Politics* treatment than in the *Religion* treatment. For actions favouring the out-group, there is not such a clear ranking between treatments and differences are often not statistically significant. However, evaluations of extreme choices show significantly stronger inappropriateness of favouring the out-group member when they are a political opponent than in other treatments.

Table 3: P-values for permutation tests comparing mean ratings of social appropriateness across treatments (UK)

	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
P-values									
Politics vs. Religion	0.000	0.000	0.000	0.000	0.017	0.053	0.322	0.040	0.003
Politics vs. Artificial	0.001	0.000	0.000	0.000	0.134	0.040	0.119	0.046	0.008
Artificial vs. Religion	0.013	0.004	0.000	0.000	0.468	0.000	0.025	0.966	0.912

Notes: P-values obtained from Fisher-Pitman permutation tests for two independent samples, corrected using the Benjamini-Hochberg False Discovery Rate method.

Next, we focus on decisions made in the allocation tasks. Figure 2 reveals two broad tendencies quite clearly. First, in correspondence to the norm elicitation results, the modal distribution is the equal split, which accounts for 68% of allocations. Second, those that deviate do so mostly to favour the in-group member, with only 3.2% of allocations favouring the out-group member. On average, participants allocate £1.88 more to the partner with whom they share an identity. Even so, as was the case for the norm-elicitation task, there are significant differences across treatments. The in-group premium ranges from an average of £0.88 in the *Religion* treatment, with 85.2% of allocations producing an equal (8,8) split, to £3.07 in the case of the *Politics* treatment, where only 53.4% of allocations were the equal split. Results for the Artificial treatment lay in between: the in-group premium is on average £1.52, while 68.3% of allocations resulted in an equal split. Thus, the ranking of discrimination across the three treatments matches that of the perceived social appropriateness of discrimination discussed above.[13]

Table 4 presents a random effects model of the allocation decisions. The dependent variable is the level of discrimination in favour of the in-group participant, measured by subtracting the amount allocated to the out-group from the amount allocated to the in-group. Columns (1) and (2) of the

---

[13] See section D of the Online Appendix for a detailed description of the average allocation decisions across treatments and sub-groups.

regression table confirms the observed difference of more than two pounds for the in-group premium between the *Religion* and *Politics* treatments. That difference is statistically significant[14], as is the difference between the *Politics* and *Artificial* treatments. These results are robust to a set of control variables that account for individual characteristics. Alternative specifications including other individual characteristics such as ideology, party affiliation and religion do not alter the results, with no statistically significant effects of the additional control variables either. The added premium in the *Artificial* treatment (the baseline category in this model) is not statistically distinguishable from that of the *Religion* treatment.

Figure 2: Allocation decisions, by treatment (UK)
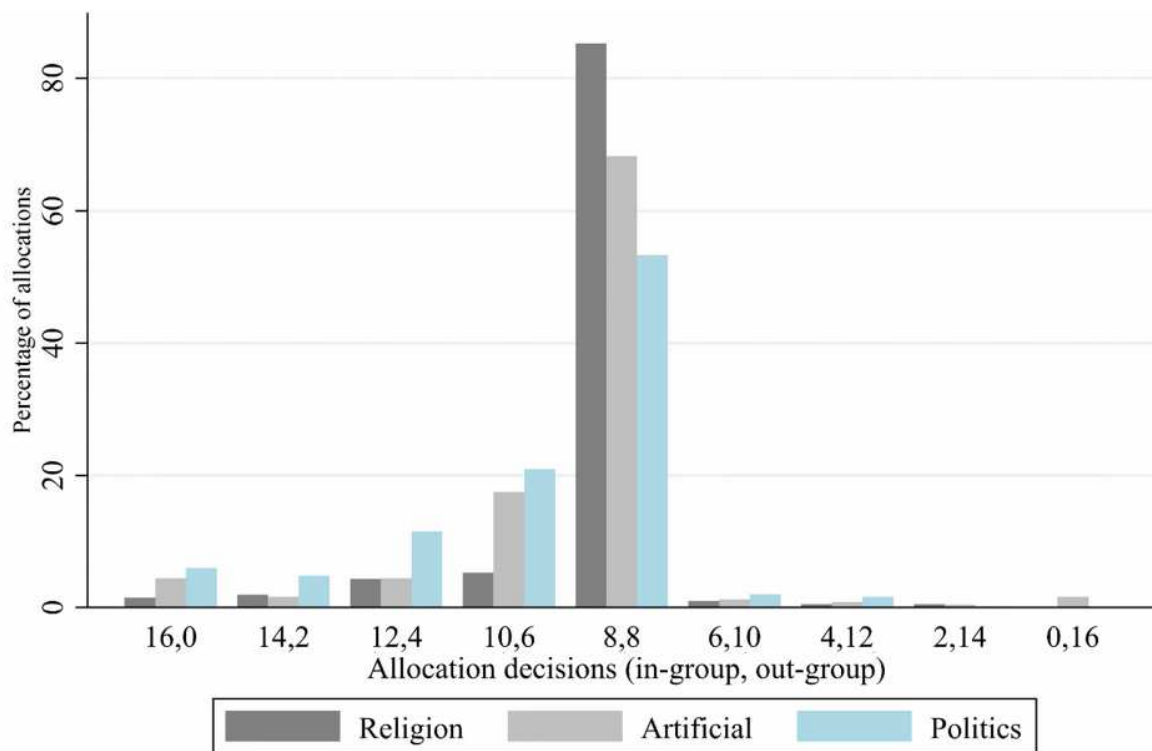
Note: Percentages of allocation decisions, by treatment. Each participant made three allocation decisions, one for each different possible affiliation of the out-group participant.

---

[14] A linear restriction test indicates that allocations in the *Politics* and *Religion* treatments are statistically different (*p < 0.001*).

Table 4: Regression analysis of allocation decisions

	Nottingham, UK		Castellón, Spain		Pooled	
	Dependent Variable: Amount allocated to in-group – amount allocated to out-group participant					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment:						
Religion	-0.648	-0.629	-0.07	-0.124	-0.648	-0.622
	(0.583)	(0.594)	(0.858)	(0.872)	(0.659)	(0.665)
Politics	1.543***	1.513***	2.356***	2.421***	1.543**	1.568**
	(0.554)	(0.575)	(0.847)	(0.853)	(0.627)	(0.64)
Spain					1.587**	1.72**
					(0.689)	(0.7)
Spain*Religion					0.577	0.513
					(1.001)	(1.013)
Spain*Politics					0.813	0.835
					(0.973)	(0.985)
Controls	✗	✓	✗	✓	✗	✓
Constant	1.524***	3.032	3.111***	5.7**	1.524***	2.283*
	(0.393)	(2.382)	(0.599)	(2.321)	(0.445)	(1.295)
Observations	712	712	531	531	1246	1246
R^2	0.041	0.047	0.038	0.051	0.075	0.079

Note: Results for random effects models. The reference category is the *Artificial* treatment. Control variables include: gender (binary variable, 1=female), age, year of university degree the participant is in, rural if they are originally from a rural area, and household income. Robust standard errors are in parentheses. Stars indicate significance level: *** $p<.01$, ** $p<.05$, * $p<.1$.

3.2. Castellón, Spain

In this section, we present the same results reported in the previous section, but this time for the experiment in Castellón, Spain. To avoid repetitions, we will only highlight the main effects and focus on the results that deviate from those obtained in Nottingham. Table 5 reports the mean social appropriateness ratings of the nine possible distributions. As in Nottingham, there is a high degree of agreement on the social appropriateness of the equal split in the *Religion* and the *Artificial* treatments, where this distribution is considered "very socially appropriate" in 91% and 83% of the responses. In the same vein, the inappropriateness of discriminatory choices in these treatments is symmetric and participants evaluate favouring the in-group and the out-group participants alike. Yet, the *Politics* treatment deviates quite substantially from this consensus and

Table 5: Distribution of social appropriateness ratings, by treatment (Spain)

|	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
Politics treatment									
Very appropriate	24,4	16,1	17,8	17,8	56,1	1,1	0,0	0,0	0,0
Somewhat appropriate	10,0	25,0	32,2	46,7	23,9	14,4	3,3	1,1	0,6
Somewhat inappropriate	10,0	16,7	21,7	23,3	14,4	46,7	27,2	11,1	3,9
Very inappropriate	55,6	42,2	28,3	12,2	5,6	37,8	69,4	87,8	95,6
Mean rating	**-0,31**	**-0,23**	**-0,07**	**0,13**	**0,54**	**-0,47**	**-0,77**	**-0,91**	**-0,97**
Religion treatment									
Very appropriate	2,9	4,1	3,5	3,5	90,6	4,7	3,5	1,2	0,6
Somewhat appropriate	1,2	1,2	7,6	38,0	7,0	24,0	3,5	3,5	0,0
Somewhat inappropriate	1,8	19,3	36,3	38,6	1,8	45,0	33,3	16,4	7,6
Very inappropriate	94,2	75,4	52,6	19,9	0,6	26,3	59,6	78,9	91,8
Mean rating	**-0,91**	**-0,77**	**-0,59**	**-0,17**	**0,92**	**-0,29**	**-0,66**	**-0,82**	**-0,94**
Artificial treatment									
Very appropriate	11,1	10,0	10,6	10,0	83,3	5,6	3,9	3,9	3,9
Somewhat appropriate	3,3	8,3	17,8	61,1	13,3	54,4	13,9	6,1	6,1
Somewhat inappropriate	7,8	26,7	53,3	26,1	1,1	28,3	47,8	24,4	8,9
Very inappropriate	77,8	55,0	18,3	2,8	2,2	11,7	34,4	65,6	81,1
Mean rating	**-0,68**	**-0,51**	**-0,20**	**0,19**	**0,85**	**0,03**	**-0,42**	**-0,68**	**-0,78**

Notes: Percentage of responses corresponding to each of the available social appropriateness ratings for each of the nine possible decisions in the allocator game. Decisions range from allocating all the money to the in-group participant (16,0) to allocating all the money to the out-group participant (0,16). The modal evaluation for each outcome is highlighted. Mean ratings are computed assigning values of 1, 0.33, −0.33 and −1 to assessments of "very socially appropriate", "somewhat socially appropriate", "somewhat socially inappropriate" and "very socially inappropriate", respectively. Therefore, mean ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

the symmetry of the evaluations. First, only 56% of this treatment's responses rate the equal split as "very socially appropriate". Second, only about half of the responses rate the extreme distributions favouring the in-group (16,0 and 14,2) as "very socially inappropriate". Third, there is no symmetry in the evaluations anymore, since the extreme distributions favouring the out-group

are clearly considered inappropriate, but those favouring the in-group are not. In summary, the *Politics* treatment in Castellón is the context in which discrimination is by far the most tolerated.

Figure 3 plots the mean ratings of all allocation outcomes by treatment in Castellón. The shape of the lines corresponding to the *Religion* and *Artificial* treatments match the Nottingham data reported in Figure 1. The *Politics* treatment is different in that the mean evaluation of the equal split is substantially lower than any other treatment. This is confirmed by the p-values of tests comparing the mean ratings of social appropriateness between treatments reported in Table 6. Interestingly, moderate and extreme discrimination against the outgroup (12,4; 14,2; 16;0) follows the same treatment ranking as in Nottingham, whereas small deviations toward the in-group are regarded as more socially appropriate in the *Artificial* than in the *Politics* treatment. On the right-hand side plotting the ratings of distributions favouring the out-group, *Politics* is consistently below *Religion* and *Artificial*, with only the exception of extreme discrimination against the in-group (0,16) in the *Religion* treatment. All in all, the lines reveal a clear pattern of tolerance over in-group favouritism and rejection of out-group advantage in the *Politics* treatment. Does this pattern of normative beliefs guide actual behaviour?

Figure 3: Mean evaluation of social appropriateness of allocation decisions, by treatment (Spain)
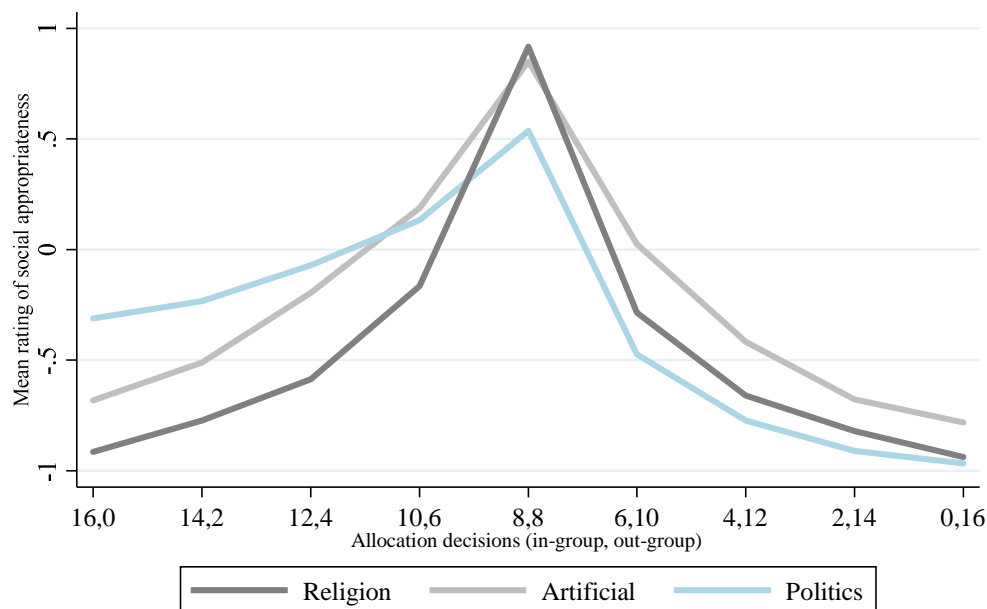
Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

Table 6: P-values for permutation tests comparing mean ratings of social appropriateness across treatments (Spain)

	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
	P-values								
Politics vs. Religion	0.000	0.000	0.000	0.000	0.000	0.001	0.017	0.012	0.182
Politics vs. Artificial	0.000	0.001	0.087	0.358	0.000	0.000	0.000	0.000	0.000
Artificial vs. Religion	0.000	0.000	0.000	0.000	0.072	0.000	0.000	0.008	0.000

Notes: P-values obtained from Fisher-Pitman permutation tests for two independent samples, corrected using the Benjamini-Hochberg False Discovery Rate method.

Figure 4 plots decisions made in the allocation tasks in Castellón. A first striking result is the large number of deviations from the equal split, as compared with our results in Nottingham and those reported in Barr et al. (2018). This is confirmed by the dummy variable *Spain*, and its interactions, in the pooled regression analysis on discrimination levels reported in the last two columns of Table 4. Second, the same ranking of treatments is replicated, *Politics* being the one in which more discrimination is observed, followed by the *Religion* and *Artificial* treatments. As in Nottingham, the *Artificial* and *Religion* treatments are not significantly different, but the *Politics* treatment generates significantly more in-group favouritism than the other two (see columns (3) and (4) of Table 4).[15] Finally, the crucial contribution of the Spanish experiment is that although there is a strong and significant difference between countries in the overall levels of discrimination, there are not significant differences in the treatment effects (interactions *Spain*Religion* and *Spain*Politics* in columns (5) and (6) of Table 4). Furthermore, in the Spanish experiment – just like in the UK – the ranking of treatments by discrimination reflects the ranking according to the social appropriateness of discrimination. This confers credibility to the argument that our findings on political identity are robust to changes in the social context. Regardless of a society's baseline tendency for discrimination, the political domain induces higher levels of discrimination than other dimensions of identity, which reflects the relatively weak norms it induces against such discrimination. Finally, an interesting result emerges that fosters our confidence on the independence of the normative and behavioural measures. Although in the *Religion* and *Artificial* treatments in Castellón there is high agreement on the social inappropriateness of in-group favouritism, participants in this treatment do discriminate somewhat frequently, as revealed by the

---

[15] Linear restriction tests indicate that allocations in the *Politics* and *Religion* treatments are statistically different (*p < 0.001*).

grey bars in Figure 4. This means that discriminatory behaviour is not fully determined by norms reported in the lab, but also reflect personal preferences which in some cases may be to violate them.

Figure 4: Allocation decisions, by treatment (Spain)
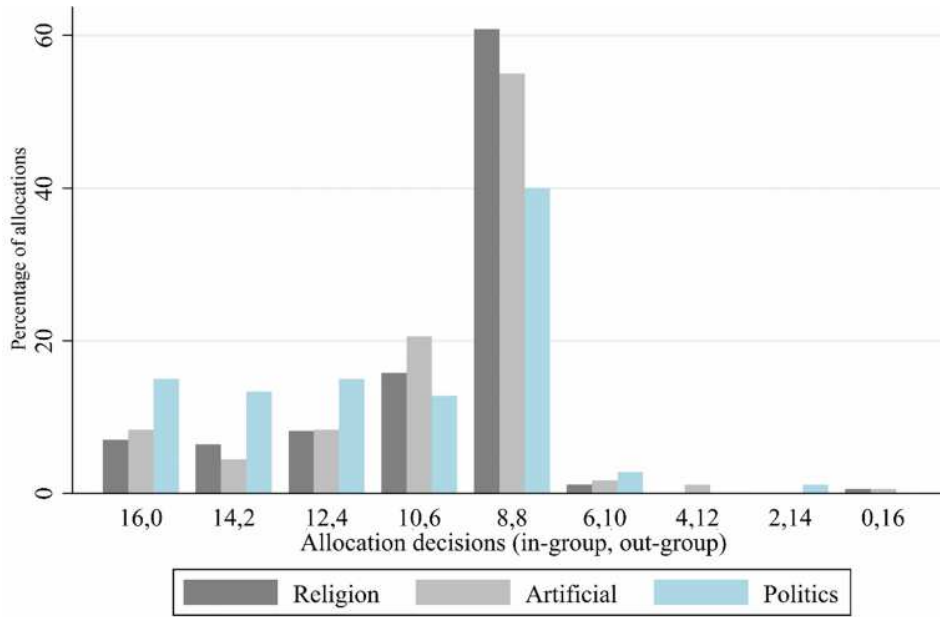
Note: Percentages of allocation decisions, by treatment. Each participant made three allocation decisions, one for each different possible affiliation of the out-group participant.

3.3. Subgroup analysis

Here, we take advantage of the four identity categories in the Politics treatment to study whether results are driven by some subgroup behaviour. In section E of the Online Appendix, we report such analysis for Nottingham and Castellón separately. In Spain, the conservatives (PP and VOX) discriminate more, and they seem to regard in-group favouritism as more socially appropriate on average. In the UK, conservatives receive the lowest amount from out-group allocators.[16] In Spain, the two extremist parties receive less and the two moderate parties more. This is just preliminary evidence about subgroup behaviour (the analysis was not pre-registered). Further research is needed to better understand discriminatory behaviour in multi-party systems.[17]

---

[16] This is consistent with previous research finding greater hostility from liberals toward conservatives in the UK (Lane, 2023).

[17] We conducted similar subgroup analysis for our Religion treatment, which we report in section F of the Online Appendix.

4. Discussion

Political sectarianism, extremism, and polarization are among the social phenomena about which citizens in Western democracies are more concerned (Finkel et al., 2020). These issues feature highly in the news and all forms of media, traditional or new. They are also among the most frequent social problems identified by citizens. Political polarization is becoming so severe that it is even bringing democracy itself and societal stability under threat as it undermines support for democratic norms (Kingzette et al., 2021). Descriptive studies in the social sciences defining these topics abound. However, there is a lack of studies attempting to address the causal processes underlying such social phenomena. In this paper, we explore one mechanism explaining the primacy of political polarization over other social divides: the role of social norms.

We follow a long tradition in social psychology and economics that goes back to the works of Tajfel (1970) and conduct experiments in which we use group identities and elicit discriminatory choices. More specifically, we base our study on the previous investigation of Barr et al. (2018) that introduced an experimental design to study norms and discriminatory behaviour in different social contexts. They compared artificial and national identity contexts and found strongest discrimination on the identity dimension for which the social norms against discrimination were weakest. We extend their design to study a larger set of identities (including religion and politics) and a larger set of identity categories (four rather than two).

Our main results are two. First, we replicate the finding that discrimination is largest when norms against it are weaker. Second, regardless of the experimental setting, most discrimination is observed in the political domain. We conducted the experiment in two locations that differed in their baseline tendency for discrimination, but always found by far the strongest discrimination in the *Politics* treatment, which also mirrored the fact that in both countries norms against discrimination were weakest in this domain. Additionally, our four-category identities allow us to perform subgroup analysis and study whether there is a subgroup that is driving the results. On average, conservatives seem to regard in-group favouritism as more socially appropriate.

Why is the political domain different? We argue that this domain may lack the sort of norms that prevent discrimination, prejudice, and hostility between social groups in society. Interactions across religion or nationality, also race, gender, and other social divides are constrained by social norms, but there seem not to be corresponding pressures or sanctions that mute disapproval of

political opponents. Partisans, therefore, can feel free to express animus and engage in discriminatory behaviour toward their opponents. The deeper question of *why* norms have not developed to counter political discrimination is beyond the scope of this paper, but we can offer the reader speculation. In a democratic system, people are free to select in and out of political groups. A person's political leanings are a reflection of their philosophical values and opinions about society. It may be regarded as acceptable to dislike a person because of their worldview, but not on the basis of a characteristic such as race or sex which is unchosen and fixed from birth.[18] This is an important matter for future research.

We believe our design sets the stage for international comparisons of the role of social norms on political discrimination and polarization. We have shown that our treatment effects remain qualitatively the same in two societies that differ in their baseline tendencies for discrimination and in their political ecosystems. A worthwhile endeavour would be to extend the analysis to a wider set of more different societies, including non-Western countries and those with less democratic political systems. The relative simplicity of our experimental design would allow it to be easily implemented in large-sample online studies. Finally, understanding the role of norms in political polarization may help devise interventions to return political conflict to the realm of ideas, reducing interpersonal hostilities. Once knowledge is established about existing norms, how they vary across contexts and the mechanisms behind their development and enforcement, one could envision and formulate social interventions that help break the vicious cycle of political group hostility (Dimant, 2023).

---

[18] It is interesting to note that religion, which we also employed in our experiment, is at root also a matter of belief, with people able to switch between affiliations. In practice, however, religion has evolved into a characteristic that is explicitly inherited from one's parents, and conversion is a rare event. As such, religion may be viewed less as a choice-based identity group like politics, and more akin to an immutable one like ethnicity.

Acknowledgments

References

Allcott, H., 2011. Social norms and energy conservation. J. Public Econ. 95 (9-10), 1082-1095.

Barr, A., Lane, T., Nosenzo, D., 2018. On the social inappropriateness of discrimination. J. Public Econ., 164, 153-164.

Bicchieri, C., 2005. The grammar of society: The nature and dynamics of social norms. Cambridge University Press.

Bicchieri, C., Dimant, E., Gächter, S., Nosenzo, D., 2022. Social proximity and the erosion of norm compliance. Game Econ. Behav. 132, 59-72.

Clark, C.J., Liu, B.S., Winegard, B.M., Ditto, P.H., 2019. Tribalism Is Human Nature. Curr. Dir. in Psychol. Sci. 28 (6), 587-592.

Coffman, K. B., Exley, C. L., Niederle, M., 2021. The role of beliefs in driving gender discrimination. Manage. Sci. 67(6), 3551-3569.

Cohen, N., 2016. Did better broadband make Americans more partisan? The Guardian. https://www.theguardian.com/commentisfree/2016/aug/27/donald-trump-fox-news-charlie-sykes-modern-debate

Dimant, E., 2023. Hate Trumps Love: The Impact of Political Polarization on Social Preferences. Manage. Sci.

Elster, J. 1989. Social norms and economic theory. J. Econ. Perspect. 3(4), 99-117.

Fallucchi, F., Nosenzo, D., 2022. The coordinating power of social norms. Exp. Econ. 25(1), 1-25.

Fehr, E., Schurtenberger, I., 2018. Normative foundations of human cooperation. Nat. Human Behav. 2(7), 458-468.

Finkel, E. J., et al., 2020. Political sectarianism in America. Science 370 (6516), 533-536

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10(2), 171-178.

Gächter, S., Nosenzo, D., Sefton, M, 2013. Peer effects in pro-social behavior: Social norms or social preferences? J. Eur. Econ. Assoc. 11(3), 548-573.

Gaertner, S. L., Dovidio, J. F., 2000. Reducing intergroup bias: The Common Ingroup Identity Model. The Psychology Press.

Iyengar, S., Westwood, S.J., 2015. Fear and Loathing across Party Lines: New Evidence on Group Polarization. Am Polit Sci Rev. 59 (3), 690-707.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., Westwood, S.J., 2019. The Origins and Consequences of Affective Polarization in the United States. Annu Rev of Polit Sci 22 (1), 129-146.

Kimbrough, E. O., Vostroknutov, A., 2016. Norms make preferences social. J. Eur. Econ. Assoc. 14(3), 608-638.

Kingzette, J., Druckman, J., Klar, S., Krupnikov, Y., Levendusky, M., Ryan, J. B., 2021. How Affective Polarization Undermines Support for Democratic Norms. Public Opin. Quart. 85 (2), 663–677

Köbis, N. C., Troost, M., Brandt, C. O., Soraperra, I., 2022. Social norms of corruption in the field: social nudges on posters can help to reduce bribery. Behav. Public Policy, 6(4), 597-624.

Krupka, E. L., Leider, S., Jiang, M., 2017. A meeting of the minds: informal agreements and social norms. Manage. Sci. 63(6), 1708-1729.

Krupka, E. L., Weber, R. A., 2013. Identifying social norms using coordination games: Why does dictator game sharing vary? J. Eur. Econ. Assoc. 11(3), 495-524.

Lane, T., 2016. Discrimination in the laboratory: A meta-analysis of economics experiments. Eur. Econ. Rev. 90, 375-402.

Lane, T., 2023. The strategic use of social identity. CeDEx Discussion Paper No. 2023-01.

Lane, T., Nosenzo, D., Sonderegger, S., 2023. Law and Norms: Empirical Evidence. Am. Econ. Rev.

Michelitch, K., 2015. Does Electoral Competition Exacerbate Interethnic or Interpartisan Economic Discrimination? Evidence from a Field Experiment in Market Price Bargaining. Am. Pol. Sci. Rev. 109(1), 43-61.

Nosenzo, D., Görges, L., 2020. Measuring social norms in economics: why it is important and how it is done. Analyse & Kritik 42(2), 285-312.

Ostrom, E, 2000. Collective action and the evolution of social norms. J. Econ. Perspect. 14(3), 137-158.

Restrepo-Plaza, L., Fatas, E., 2022. When ingroup favoritism is not the social norm a lab-in-the-field experiment with victims and non-victims of conflict in Colombia. J. Econ. Behav. Organ. 194, 363-383.

Schelling, T. C., 1980. The Strategy of Conflict. Harvard university press.

Sunstein, C.R., 2016. Partyism. Univ Chicago Leg For. 2015 (2), 1-26.

Szekely, A., Lipari, F., Antonioni, A., Paolucci, M., Sánchez, A., Tummolini, L., Andrighetto, G. 2021. Evidence from a long-term experiment that collective risks change social norms and promote cooperation. Nature Commun. 12(1), 5452.

Tajfel, H., 1970. Experiments in Intergroup Discrimination. Sci Am. 223 (5), 96-103.

Westwood, S.J., Iyengar, S., Walgrave, S., Leonisio, R., Miller, L., Strijbis, O., 2018. The Tie That Divides: Cross-National Evidence of the Primacy of Partyism. Eur J Polit Res. 57 (2), 333-354.

Online Appendix

Appendix A. Pre-survey results

For the first set of experiments in Nottingham (UK), the pre-survey was distributed to 3715 students registered on ORSEE, after which we received 678 responses and retained 385 eligible subjects. Participation in the survey was incentivised. Subjects were considered eligible for the lab experiment if we had information both on their partisanship and religious affiliation, and those both corresponded to the four most common groups. For the second set of experiments in Castellón (Spain), we maximised the pool of eligible subjects to 497 participants by considering answers to an additional pre-survey question for the identification of political partisanship, and by regarding atheists and agnostics as two separate religious groups. For similar reasons, in the UK, we combined Catholics and Protestants as one religious group, Christian.

Table A1: Pre-survey results

Nottingham, UK		Castellón, SPAIN	
Partisanship			
Conservative	71 (12.5%)	PP	194 (19.5%)
Green	75 (13.2%)	PSOE	222 (22.4%)
Labour	222 (39.2%)	UP	138 (13.9%)
Liberal-Democrats	59 (10.4%)	VOX	72 (7.3%)
Missing	140 (24.7%)	Other	210 (21.2%)
		Missing	157 (15.8%)
Religion			
Buddhist	8 (1.4%)	Agnostic	226 (22.8%)
Catholic	44 (7.8%)	Atheist	184 (18.5%)
Hindu	83 (14.6%)	Catholic	343 (34.6%)
Jewish	8 (1.4%)	Muslim	33 (3.3%)
Muslim	37 (6.5%)	Orthodox	15 (1.5%)
Protestant	52 (9.2%)	Evangelical	8 (0.8%)
Sikh	9 (1.6%)	Buddhist	6 (0.6%)
No religion	256 (45.1%)	Other	20 (2.0%)
Other	48 (8.5%)	Missing	158 (15.9%)
Prefer not to say	9 (1.6%)		
Missing	13 (2.3%)		
Total	**567**		**993**

Notes: **Religion** shows responses to the question *How would you define yourself in religious terms?* in the survey for both set of experiments. **Partisanship** shows responses to the question *If the next general election were held tomorrow, which party would you vote for?* and, in the case of Castellón (Spain) it includes as well responses *to Name this party that you feel closer to than any other* when participants admitted to feeling closer to one particular political party.

Appendix B. Demographic characteristics of participants, by treatment

As it was stated above, despite distributing the pre-survey for the first set of experiments in Nottingham to 3715 students registered on ORSEE, after receiving 678 responses, we retained 385 of eligible subjects, of whom 155 eventually participated in the *Religion* or *Politics* treatment. To fill the *Artificial* treatment, we therefore recruited directly from the wider ORSEE subject pool (excluding first those subjects who were assigned to the other treatments or who were not retained as eligible based on their responses to the pre-survey). Therefore, the most important results of our study should be considered the comparisons between the *Religion* and *Politics* treatments, with the *Artificial-Religion* and *Artificial-Politics* comparisons regarded as suggestive but not perfectly controlled. Like in the other treatments, we only recruited subjects whose nationality on ORSEE was registered as British (this was done in order to ensure political identity was meaningful, as international students may have been rather unfamiliar with British parties).

Table B1: Demographic characteristics of participants by treatment, in Nottingham (UK)

	Total N=238	Politics N=85	Religion N=70	Artificial N=83
Age	21.0 (2.9)	20.8 (3.0)	21.1 (3.2)	21.1 (2.6)
Gender				
Male	107 (45.0%)	30 (35.3%)	30 (42.9%)	47 (56.6%)
Female	127 (53.4%)	53 (62.4%)	39 (55.7%)	35 (42.2%)
Other	4 (1.7%)	2 (2.4%)	1 (1.4%)	1 (1.2%)
Background				
Rural	30 (12.6%)	14 (16.5%)	9 (12.9%)	7 (8.4%)
Urban	121 (50.8%)	41 (48.2%)	39 (55.7%)	41 (49.4%)
Mixed	87 (36.6%)	30 (35.3%)	22 (31.4%)	35 (42.2%)
Ideology				
Left	104 (43.7%)	37 (43.5%)	31 (44.3%)	36 (43.4%)
Centre	60 (25.2%)	15 (17.6%)	15 (21.4%)	30 (36.1%)
Right	46 (19.3%)	14 (16.5%)	15 (21.4%)	17 (20.5%)
Missing	28 (11.8%)	19 (22.4%)	9 (12.9%)	0 (0.0%)

Notes: Statistics for Age include mean and standard deviation.

Table B2: Demographic characteristics of participants by treatment, in Castellón (Spain)

	Total	Politics	Religion	Artificial
	N=177	N=60	N=57	N=60
Age	25.2 (10.7)	24.4 (8.7)	27.7 (14.1)	23.7 (8.3)
Gender				
Male	64 (36.2%)	20 (33.3%)	21 (36.8%)	23 (38.3%)
Female	111 (62.7%)	40 (66.7%)	36 (63.2%)	35 (58.3%)
Other	2 (1.1%)	0 (0.0%)	0 (0.0%)	2 (3.3%)
Background				
Rural	16 (9.0%)	7 (11.7%)	5 (8.8%)	4 (6.7%)
Urban	84 (47.5%)	24 (40.0%)	26 (45.6%)	34 (56.7%)
Mixed	77 (43.5%)	29 (48.3%)	26 (45.6%)	22 (36.7%)
Ideology				
Left	74 (41.8%)	23 (38.3%)	25 (43.9%)	26 (43.3%)
Centre	36 (20.3%)	9 (15.0%)	8 (14.0%)	19 (31.7%)
Right	67 (37.9%)	28 (46.7%)	24 (42.1%)	15 (25.0%)

Notes: Statistics for Age include mean and standard deviation.

Despite attaining a slightly larger pool of eligible subjects for the second set of experiments in Castellón, we decided to follow the same recruitment strategy as in the first set in Nottingham.

Appendix C. Distribution of sub-group identities across treatments

Table C1: Identity categories per treatment (Nottingham, UK)

Politics treatment				
Conservative	Green	Labour	Lib-dem	Total
15 (17.7%)	13 (15.3%)	38 (44.7%)	19 (22.4%)	85
Religion treatment				
Cristian	Hindu	Muslim	Non-religious	Total
15 (21.4%)	13 (18.6%)	4 (5.7%)	38 (54.3%)	70
Artificial treatment				
Blue	Pink	Red	Yellow	Total
17 (19.5%)	25 (28.7%)	27 (31.0%)	18 (20.7%)	87

Table C2: Identity categories per treatment (Castellón, Spain)

Politics treatment				
PP	PSOE	UP	VOX	Total
21 (35%)	20 (33.3%)	11 (18.3%)	8 (13.3%)	60
Religion treatment				
Agnostic	Atheist	Catholic	Muslim	Total
16 (28.1%)	10 (17.5%)	30 (52.6%)	1 (1.75%)	57
Artificial treatment				
Yellow	Blue	Red	Pink	Total
15 (25%)	15 (25%)	15 (25%)	15 (25%)	60

Appendix D. Average allocation decisions across treatments and sub-groups

Table D1: Average allocations by treatment

	Politics	**Religion**	**Artificial**
Mean allocation to **in-group** participant	£9.55	£8.44	£8.76
Mean allocation to **out-group** participant	£6.47	£7.56	£7.24
Difference	£3.07	£0.88	£1.52

Notes: Average pounds allocated by participants in the allocator game to the in-group and out-group member (and difference).

Table D2: Mean difference in pounds allocated to in-group vs. out-group participant (Nottingham, UK)

Politics treatment			
by identity of allocator			
Conservative	Green	Labour	Lib-dem
£3.64	£3.18	£3.08	£2.50
by identity of out-group participant			
Conservative	Green	Labour	Lib-dem
£5.14	£2.17	£2.17	£2.46
Religion treatment			
by identity of allocator			
Cristian	Hindu	Muslim	Non-religious
£0.80	£1.95	£0	£0.63
by identity of out-group participant			
Cristian	Hindu	Muslim	Non-religious
£1.02	£0.21	£1.21	£1.13
Artificial treatment			
by identity of allocator			
Blue	Pink	Red	Yellow
£2.04	£3.01	£0.67	£0.37
by identity of out-group participant			
Blue	Pink	Red	Yellow
£1.02	£1.38	£2.00	£1.76

Notes: Average difference in pounds allocated to the in-group and out-group, aggregated by allocator and identity of out-group member.

Table D3: Average allocations by treatment (Castellón, Spain)

	Politics	**Religion**	**Artificial**
Mean allocation to **in-group** participant	€10.73	€9.52	€9.56
Mean allocation to **out-group** participant	€5.27	€6.48	€6.44
Difference	€5.47	€3.04	€3.11

Notes: Average euros allocated by participants in the allocator game to the in-group and out-group member (and difference).

Table D4: Mean difference in pounds allocated to in-group vs. out-group participant (Castellón, Spain)

|			
Politics treatment			
by identity of allocator			
PP	PSOE	UP	VOX
€6.79	€4.73	€3.52	€6.50
by identity of out-group participant			
PP	PSOE	UP	VOX
€4.00	€4.60	€6.29	€6.46
Religion treatment			
by identity of allocator			
Agnostic	Atheist	Catholic	Muslim
€2.00	€2.27	€3.47	€14.67
by identity of out-group participant			
Agnostic	Atheist	Catholic	Muslim
€2.44	€3.91	€2.96	€2.79
Artificial treatment			
by identity of allocator			
Yellow	Blue	Red	Pink
€4.44	€1.87	€3.73	€2.40
by identity of out-group participant			
Yellow	Blue	Red	Pink
€3.38	€3.20	€3.02	€2.84

Notes: Average difference in pounds allocated to the in-group and out-group, aggregated by allocator and identity of out-group member.

Appendix E. Analysis by political identity in Politics treatment

E.1. Nottingham, UK

In this section, we examine closely the results within the Politics treatment, looking at differences between interactions across different political affiliation groups. For this discussion, it should be noted that sample sizes are small and vary across subgroups. Hence, any conclusions should be regarded with care.

Figure E1.1: Mean evaluation of social appropriateness of allocation decisions in the *Politics* treatment, by political identity of the **allocator**
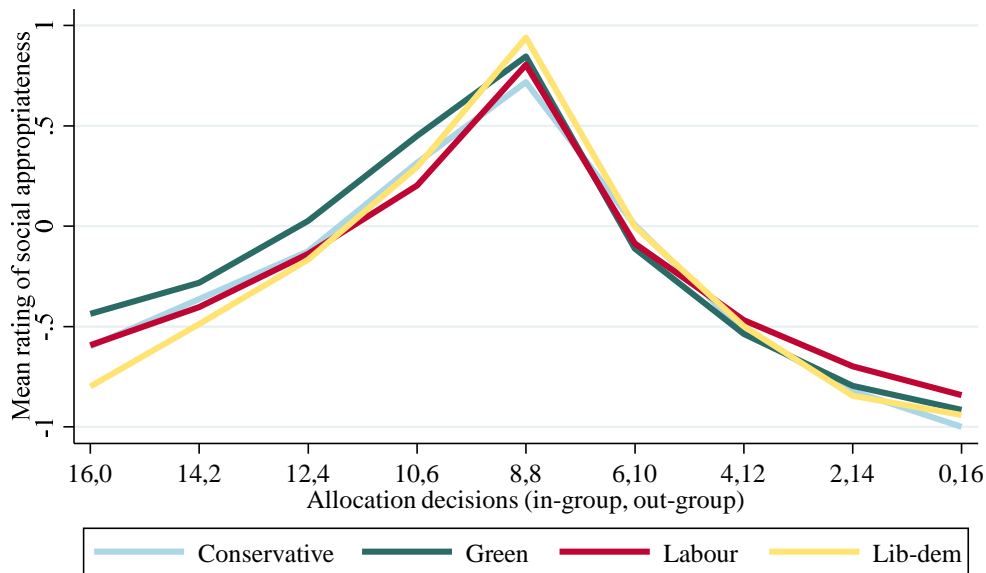
Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

Figures E1.1 and E1.2 show results from the norm-elicitation task. Figure E1.1 displays average ratings of social appropriateness based on the political affiliation of the allocator, while Figure E1.2 aggregates rating based on the identity of the out group member. That implies that, for example, those results labelled as *Conservative* include the ratings of Green, Labour and Liberal Democrat participants when assessing decisions of a participant from their group involving a *Conservative* out-group member. We see no clear patterns that deviate much from the aggregated results. We highlight briefly two results. *Green* participants consider favouring the in-group as slightly less inappropriate. On the other hand, discriminating against a *Conservative* participant is also considered slightly less inappropriate.
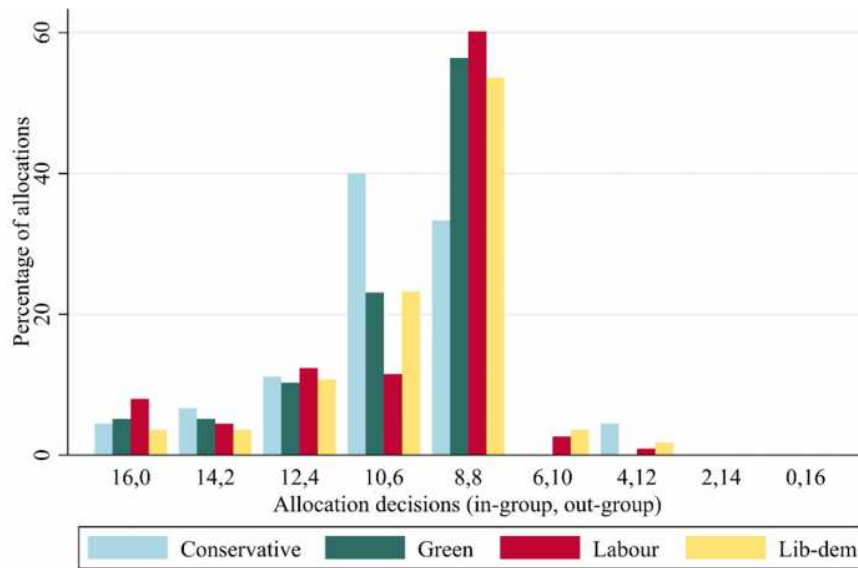
Figure E1.2: Mean evaluation of social appropriateness of allocation decisions in the *Politics* treatment, by political identity of the **out-group participant**
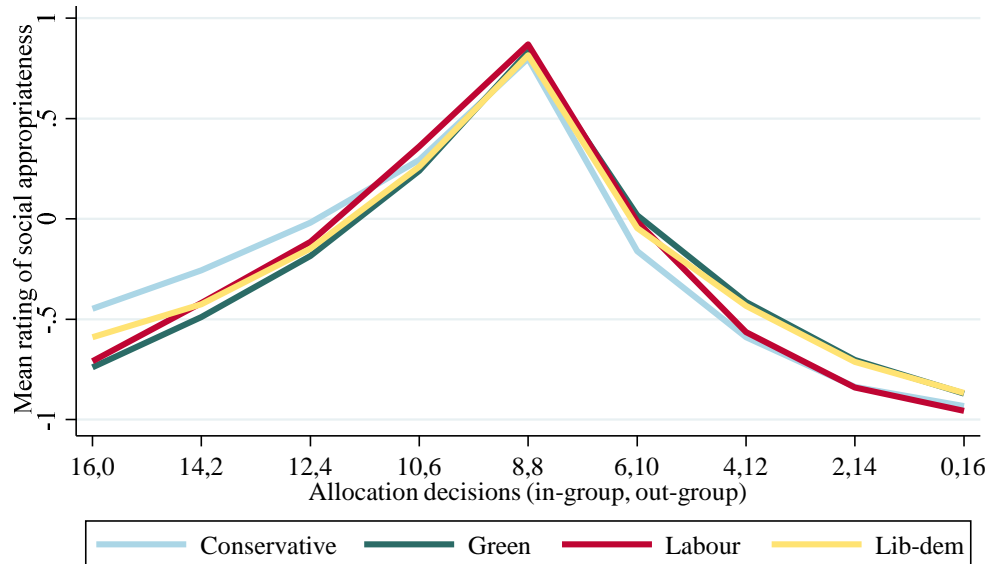
Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness). For example, those bars labelled as *Conservative* include the aggregated ratings by Green, Labour and Liberal Democrat participants for allocations of money between one member of their group and a Conservative member.

Figures E1.3 and E1.4 show the distribution of allocations across political affiliations while Table E1.1 includes the regression analysis of those decisions. The main noticeable pattern is that the equal split is not the modal choice for *Conservative* allocators. 40% of their allocations entail a small premium for the in-group, while 33% result in splitting the money equally and 22% in giving a higher premium to their *Conservative* counterpart. The equal (8,8) split represents between 53% and 60% of the allocations by participants of other political parties. Yet, as the in-group favouritism by *Conservative* allocators is on average small, differences in the in-group premium resulting from allocations by supporters of different parties are not statistically significant. Moreover, this behavioural pattern is not clearly reflected in a differentiated recognition of the social appropriateness of political discrimination by those identifying as Conservative.

Figure E1.3: Allocation decisions in the *politics* treatment, by party affiliation of the allocator

Note: Frequencies of allocation decisions in the *Politics* treatment. Coloured bars represent the political affiliation of the **allocator**.

Figure E1.4: Allocation decisions in the *politics* treatment, by party affiliation of the out-group member
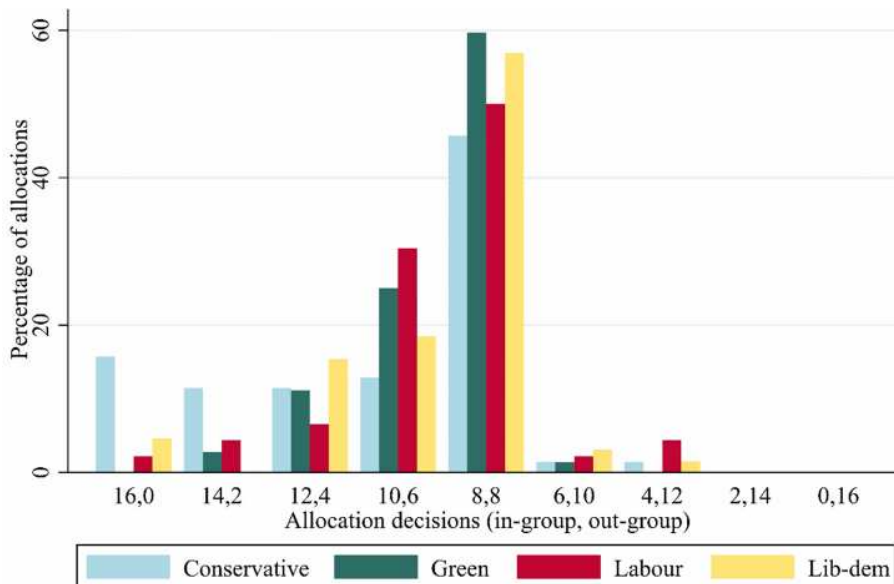
Note: Frequencies of allocation decisions in the *Politics* treatment. Coloured bars represent the political affiliation of the **out-group** member. For example, those bars labelled as *Conservative* include the aggregated decisions of Green, Labour and Liberal Democrat participants when allocating money between one member of their group and a Conservative member.

Table E1.1: Regression analysis of allocation decisions in the *Politics* treatment

| | Difference in amount allocated to in-group and out-group participant | | | |
	(1)	(2)	(3)	(4)
Allocator:				
Green	-0.465	-0.883		
	(1.501)	(1.527)		
Labour	-0.549	-0.203		
	(1.208)	(1.287)		
Liberal-democrats	-1.171	-1.07		
	(1.37)	(1.426)		
Out-group:				
Green			-3.226***	-3.212***
			(0.58)	(0.579)
Labour			-3.394***	-3.473***
			(0.681)	(0.682)
Liberal-democrats			-3.063***	-3.086***
			(0.6)	(0.6)
Controls:	✗	✓	✗	✓
Constant	3.644***	8.61**	5.385***	10.255**
	(1.023)	(4.155)	(0.557)	(3.989)
Observations	253	253	253	253
R^2	0.005	0.058	0.067	0.122

Note: Results for random effects models. Omitted category both for allocator and out-group identity is *Conservative*. Control variables include: gender (binary variable), age, year of university degree the participant is in, if they are originally from a rural or urban background, and a categorical variable of annual household income. Robust standard errors are in parentheses. Stars indicate significance level: *** p<.01, ** p<.05, * p<.1.

Looking at the decisions based on the out-group participant considerer, the higher deviations from the equal split happen when the out-group member is a *Conservative*, to the point of statistical significance. 51.4% of participants give more money to the member with whom they share political affiliation when the alternative is a *Conservative*, while that number is around 40% for the other parties. In fact, 15% of participants allocate all the money to their in-group member when the alternative is a *Conservative* supporter, in contrast to 2 to 4% of participants when the alternative is another out-group affiliation.

Table E1.2: Distribution of social appropriateness ratings in the *Politics* treatment, by political identity of the **allocator**

	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
Conservatives									
1	17.78	13.33	11.11	15.56	71.11	2.22	0	0	0
.33	0	11.11	24.44	66.67	20	53.33	4.44	0	0
-.33	6.67	33.33	48.89	17.78	4.44	37.78	62.22	26.67	0
-1	75.56	42.22	15.56	0	4.44	6.67	33.33	73.33	100
Mean rating	**-0.60**	**-0.36**	**-0.13**	**0.32**	**0.72**	**0.01**	**-0.52**	**-0.82**	**-1.00**
Green									
1	15.38	17.95	23.08	23.08	76.92	0	0	0	0
.33	10.26	17.95	23.08	71.79	23.08	43.59	10.26	5.13	5.13
-.33	17.95	17.95	38.46	5.13	0	46.15	47.72	20.51	2.56
-1	56.41	46.15	15.38	0	0	10.26	41.03	74.36	92.31
Mean rating	**-0.44**	**-0.28**	**0.03**	**0.45**	**0.85**	**-0.11**	**-0.53**	**-0.79**	**-0.91**
Labour									
1	14.16	9.73	7.08	19.47	80.53	5.31	1.77	0.88	0.88
.33	4.42	12.39	36.28	45.13	12.39	38.05	13.27	8.85	7.08
-.33	9.73	35.4	35.4	31.86	4.42	45.13	47.79	24.78	7.08
-1	71.68	42.48	21.24	3.54	2.65	11.5	37.17	65.49	84.96
Mean rating	**-0.59**	**-0.40**	**-0.14**	**0.20**	**0.81**	**-0.09**	**-0.47**	**-0.70**	**-0.84**
Lib-dem									
1	5.36	7.14	7.14	23.21	91.07	3.57	0	0	0
.33	5.36	7.14	33.93	53.57	8.93	48.21	8.93	0	0
-.33	3.57	41.07	35.71	17.86	0	42.86	57.17	23.21	8.93
-1	85.71	44.64	23.21	5.36	0	5.36	33.93	76.79	91.07
Mean rating	**-0.80**	**-0.49**	**-0.17**	**0.30**	**0.94**	**0.00**	**-0.50**	**-0.84**	**-0.94**

Note: Frequencies of perceived social appropriateness ratings of the nine possible decisions in the allocator game. Decisions go from allocating all the money to the in-group participant (16,0) to allocating all the money to the put-group participant (0,16). The modal evaluation for each outcome is highlighted. Values of 1, 0.33, −0.33 and −1 correspond to assessments of "very socially appropriate", "somewhat socially appropriate", "somewhat socially inappropriate" and "very socially inappropriate", respectively.

Table E1.3: Distribution of social appropriateness ratings in the *Politics* treatment, by political identity of the **out-group member**

	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
Conservatives									
1	18.57	15.71	15.71	24.29	78.57	1.43	0	0	0
.33	10	15.71	35.71	50	15.71	37.14	8.57	2.86	1.43
-.33	7.14	32.86	28.57	21.43	2.86	47.14	44.29	18.57	7.14
-1	64.29	35.71	20	4.29	2.86	14.29	47.14	78.57	91.43
Mean rating	**-0.45**	**-0.26**	**-0.02**	**0.29**	**0.80**	**-0.16**	**-0.59**	**-0.84**	**-0.93**
Green									
1	9.72	8.33	8.33	19.44	80.56	6.94	1.39	0	0
.33	1.39	6.94	27.78	50	15.28	47.22	12.5	6.94	5.56
-.33	6.94	37.5	41.67	27.78	2.78	37.5	58.33	30.56	8.33
-1	81.94	47.22	22.22	2.78	1.39	8.33	27.78	62.5	86.11
Mean rating	**-0.74**	**-0.49**	**-0.18**	**0.24**	**0.83**	**0.02**	**-0.42**	**-0.70**	**-0.87**
Labour									
1	8.7	8.7	10.87	21.74	84.78	2.17	0	0	0
.33	4.35	13.04	28.26	63.04	10.87	47.83	6.52	2.17	2.17
-.33	8.7	34.78	43.43	13.04	4.35	45.65	52.17	19.57	2.17
-1	78.26	43.48	17.39	2.17	0	4.35	41.3	78.26	95.65
Mean rating	**-0.71**	**-0.42**	**-0.12**	**0.36**	**0.87**	**-0.01**	**-0.56**	**-0.84**	**-0.96**
Liberal Democrats									
1	13.85	10.77	6.15	15.38	80	3.08	1.54	1.54	1.54
.33	3.08	12.31	33.85	60	15.38	44.62	12.31	6.15	6.15
-.33	13.85	29.23	41.54	23.08	1.54	44.62	55.38	26.15	3.08
-1	69.23	47.69	18.46	1.54	3.08	7.69	30.77	66.15	89.23
Mean rating	**-0.59**	**-0.43**	**-0.15**	**0.26**	**0.81**	**-0.05**	**-0.43**	**-0.71**	**-0.87**

Note: Frequencies of perceived social appropriateness ratings of the nine possible decisions in the allocator game. Decisions go from allocating all the money to the in-group participant (16,0) to allocating all the money to the put-group participant (0,16). The modal evaluation for each outcome is highlighted. Values of 1, 0.33, −0.33 and −1 correspond to assessments of "very socially appropriate", "somewhat socially appropriate", "somewhat socially inappropriate" and "very socially inappropriate", respectively.

E.2. Castellón, SPAIN

We now move to examine the results of the second set of experiments in Castellón, Spain, within the Politics treatment, looking at differences between interactions across different political affiliation groups. Again, it should be noted that sample sizes are small and vary across subgroups. Hence, any conclusions should be regarded with care.

Figure E2.1: Mean evaluation of social appropriateness of allocation decisions in the *Politics* treatment, by political identity of the **allocator**
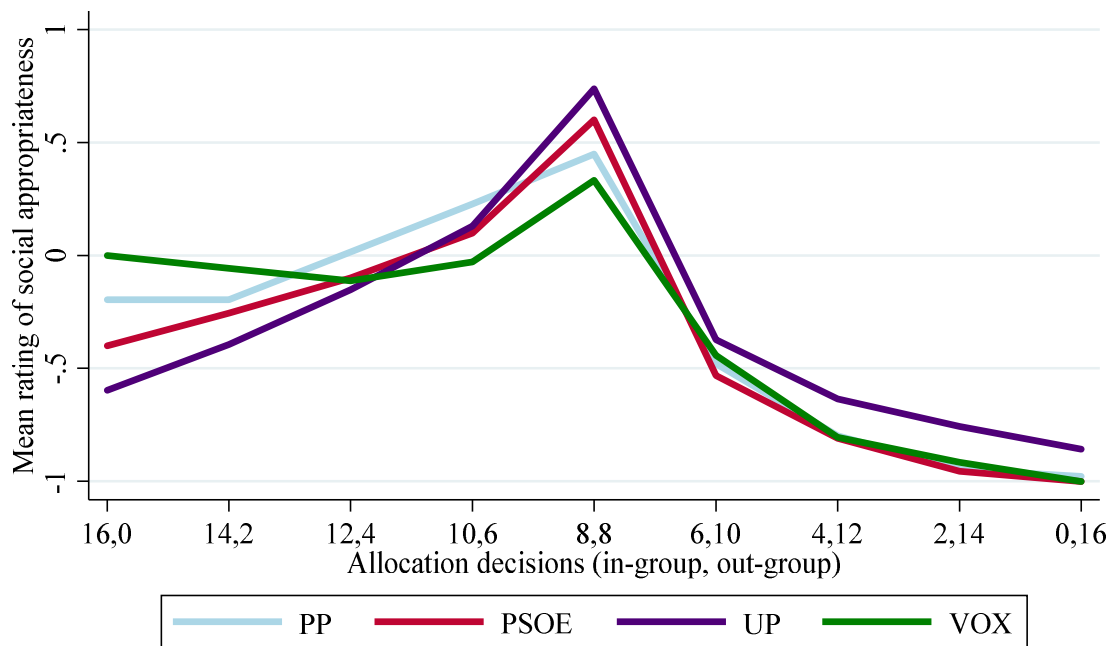
Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

Figures E2.1 and E2.2 show results from the norm-elicitation task. Figure E2.1 displays average ratings of social appropriateness based on the political affiliation of the allocator, while Figure E2.2 aggregates rating based on the identity of the out group member. That implies that, for example, those results labelled as *PP* include the ratings of PSOE, UP and VOX participants when assessing decisions of a participant from their group and a *PP* member. When looking at ratings based on the identity of the allocator, we observe a clear ranking in the social appropriateness of the equal split that correspond to the ideological positions of the parties: those more to the left in the ideology scale rate the equal split as more appropriate. This ranking perfectly inverses for the

ratings of the (16,0) split. No such strong pattern outstands when looking at social appropriateness assessments aggregated by out-group participant.
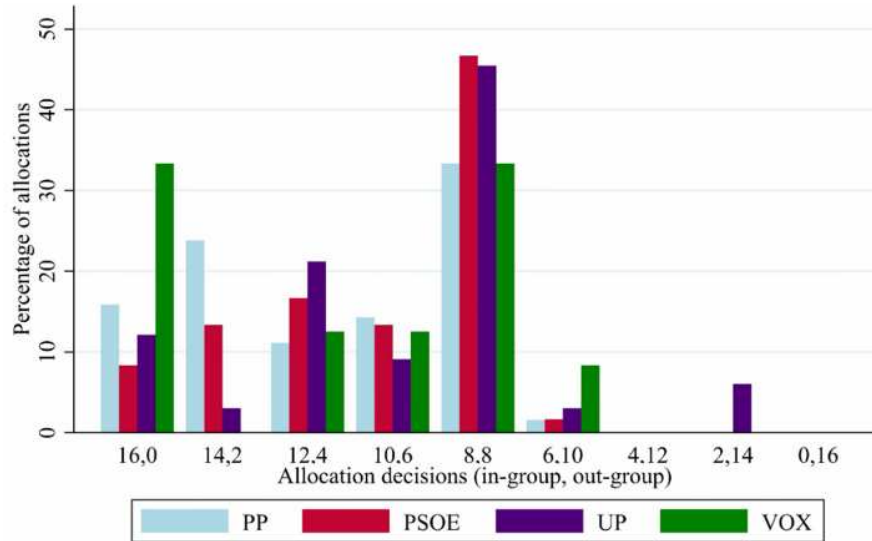
Figure E1.2: Mean evaluation of social appropriateness of allocation decisions in the *Politics* treatment, by political identity of the **out-group participant**
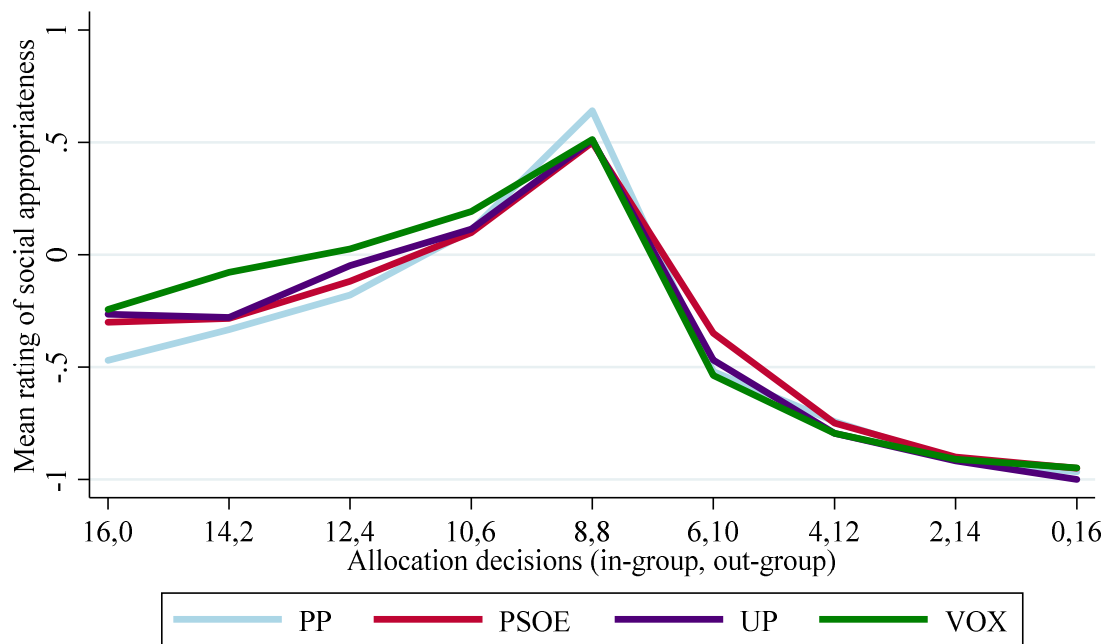
Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness). For example, those bars labelled as *PP* include the aggregated ratings by PSOE, UP and VOX participants for allocations of money between one member of their group and a PP member.

Figures E2.3 and E2.4 show the distribution of allocations across political affiliations while Table E2.1 includes the regression analysis of those decisions. The main noticeable pattern is that, in correspondence with results from the norm-elicitation task, those participants who support of a left-wing party (PSOE and UP) opt for the equal split with a higher frequency –around 11 points of difference– than those supporters of the right -wing parties (PP and VOX). On the other hand, when we look at these results aggregated by out-group member, we observe that those that have to allocate between one of their own and a supporter of a political party closer to the centre (PP and PSOE), choose the equal split in higher proportion than if the out-group choice is a supporter of the more extreme parties (UP and VOX).

Figure E1.3: Allocation decisions in the *politics* treatment, by party affiliation of the allocator

Note: Frequencies of allocation decisions in the *Politics* treatment. Coloured bars represent the political affiliation of the **allocator**.

Figure E1.4: Allocation decisions in the *politics* treatment, by party affiliation of the out-group member
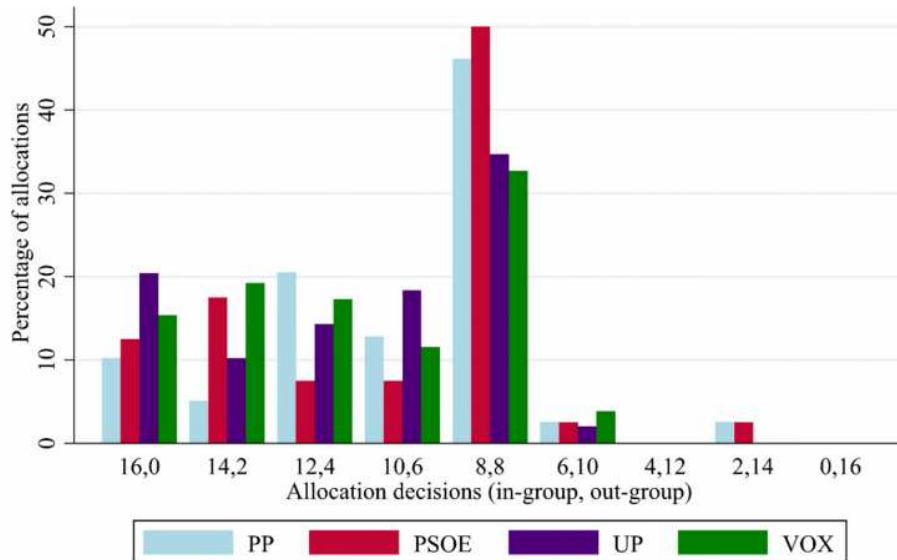
Note: Frequencies of allocation decisions in the *Politics* treatment. Coloured bars represent the political affiliation of the **out-group** member. For example, those bars labelled as *PP* include the aggregated decisions of PSOE, UP and VOX participants when allocating money between one member of their group and a PP member.

Table E2.1: Regression analysis of allocation decisions in the *Politics* treatment

| | Difference in amount allocated to in-group and out-group participant | | | |
	(1)	(2)	(3)	(4)
Allocator:				
PSOE	-2.06	-2.202		
	(1.529)	(1.67)		
UP	-3.278*	-3.209		
	(1.821)	(2.054)		
VOX	-0.294	-1.495		
	(2.033)	(2.196)		
Out-group:				
PSOE			-0.107	-0.172
			(1.205)	(1.207)
UP			1.658	1.58
			(1.128)	(1.13)
VOX			2.222**	2.258**
			(1.108)	(1.108)
Controls:	✗	✓	✗	✓
Constant	6.794***	11.149**	4.397***	8.262*
	(1.068)	(5.071)	(0.984)	(4.826)
Observations	180	180	180	180
R^2	0.039	0.076	0.025	0.074

Note: Results for random effects models. Omitted category both for allocator and out-group identity is *PP*. Control variables include: gender (binary variable), age, year of university degree the participant is in, if they are originally from a rural or urban background, and a categorical variable of annual household income. Robust standard errors are in parentheses. Stars indicate significance level: *** p<.01, ** p<.05, * p<.1.

These patterns observed empirically from aggregated results are only partially statistically significant. However, we emphasise the small size of the groups. In fact, the strongest results are obtained for the bigger groups. This sub-group analysis should be regarded as tentative.

Table E1.2: Distribution of social appropriateness ratings in the *Politics* treatment, by political identity of the **allocator**

	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
PP									
1	28.57	15.87	19.05	22.22	47.62	1.59	0	0	0
.33	9.52	28.57	38.1	46.03	30.16	9.52	1.59	1.59	0
-.33	15.87	15.87	19.05	25.4	14.29	53.97	26.98	4.76	3.17
-1	46.03	39.68	23.81	6.35	7.94	34.92	71.43	93.65	96.83
Mean rating	**-0.20**	**-0.20**	**0.02**	**0.23**	**0.45**	**-0.48**	**-0.80**	**-0.95**	**-0.98**
PSOE									
1	23.33	20	21.67	20	61.67	0	0	0	0
.33	5	16.67	21.67	41.67	18.33	13.33	0	0	0
-.33	10	18.33	26.67	21.67	18.33	43.33	28.33	6.67	0
-1	61.67	45	30	16.67	1.67	43.33	71.67	93.33	100
Mean rating	**-0.40**	**-0.26**	**-0.10**	**0.10**	**0.60**	**-0.53**	**-0.81**	**-0.96**	**-1.00**
UP									
1	9.09	6.06	9.09	12.12	75.76	0	0	0	0
.33	15.15	24.24	36.36	54.55	15.15	30.3	12.12	3.03	3.03
-.33	3.03	24.24	27.27	24.24	3.03	33.33	30.3	30.3	15.15
-1	72.73	45.45	27.27	9.09	6.06	36.36	57.58	66.67	81.82
Mean rating	**-0.60**	**-0.39**	**-0.15**	**0.13**	**0.74**	**-0.37**	**-0.64**	**-0.76**	**-0.86**
VOX									
1	37.5	20.83	16.67	8.33	37.5	4.17	0	0	0
.33	16.67	37.5	37.5	50	33.33	8.33	4.17	0	0
-.33	4.17	4.17	8.33	20.83	20.83	54.17	20.83	12.5	0
-1	41.67	37.5	37.5	20.83	8.33	33.33	75	87.5	100
Mean rating	**0.00**	**-0.06**	**-0.11**	**-0.03**	**0.33**	**-0.44**	**-0.80**	**-0.92**	**-1.00**

Note: Frequencies of perceived social appropriateness ratings of the nine possible decisions in the allocator game. Decisions go from allocating all the money to the in-group participant (16,0) to allocating all the money to the put-group participant (0,16). The modal evaluation for each outcome is highlighted. Values of 1, 0.33, −0.33 and −1 correspond to assessments of "very socially appropriate", "somewhat socially appropriate", "somewhat socially inappropriate" and "very socially inappropriate", respectively.

Table E1.3: Distribution of social appropriateness ratings in the *Politics* treatment, by political identity of the **out-group member**

	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
PP									
1	17.95	12.82	15.38	15.38	66.67	0	0	0	0
.33	7.69	20.51	25.64	48.72	15.38	12.82	2.56	0	0
-.33	10.26	20.51	25.64	23.08	15.38	46.15	33.33	12.82	5.13
-1	64.1	46.15	33.33	12.82	2.56	41.03	64.1	87.18	94.87
Mean rating	**-0.47**	**-0.33**	**-0.18**	**0.11**	**0.64**	**-0.52**	**-0.74**	**-0.91**	**-0.97**
PSOE									
1	25	15	7.5	7.5	52.5	0	0	0	0
.33	12.5	20	45	60	27.5	20	5	0	0
-.33	5	22.5	20	22.5	12.5	57.5	27.5	15	7.5
-1	57.5	42.5	27.5	10	7.5	22.5	67.5	85	92.5
Mean rating	**-0.30**	**-0.28**	**-0.12**	**0.10**	**0.50**	**-0.35**	**-0.75**	**-0.90**	**-0.95**
UP									
1	26.53	12.24	20.41	18.37	53.06	2.04	0	0	0
.33	8.16	26.53	28.57	46.94	28.57	16.33	2.04	0	0
-.33	14.29	18.37	24.49	18.37	10.2	40.82	26.53	12.24	0
-1	51.02	42.86	26.53	16.33	8.16	40.82	71.43	87.76	100
Mean rating	**-0.27**	**-0.28**	**-0.05**	**0.11**	**0.51**	**-0.47**	**-0.80**	**-0.92**	**-1.00**
VOX									
1	26.92	23.08	25	26.92	53.85	1.92	0	0	0
.33	11.54	30.77	30.77	34.62	23.08	9.62	3.85	3.85	1.92
-.33	9.62	7.69	17.31	28.85	19.23	44.23	23.08	5.77	3.85
-1	51.92	38.46	26.92	9.62	3.85	44.23	73.08	90.38	94.23
Mean rating	**-0.24**	**-0.08**	**0.03**	**0.19**	**0.51**	**-0.54**	**-0.79**	**-0.91**	**-0.95**

Note: Frequencies of perceived social appropriateness ratings of the nine possible decisions in the allocator game. Decisions go from allocating all the money to the in-group participant (16,0) to allocating all the money to the put-group participant (0,16). The modal evaluation for each outcome is highlighted. Values of 1, 0.33, −0.33 and −1 correspond to assessments of "very socially appropriate", "somewhat socially appropriate", "somewhat socially inappropriate" and "very socially inappropriate", respectively.

Appendix F. Analysis by religious identity in Religion treatment

F.1. Nottingham, UK

This section includes the detailed results within the Religion treatment. Once again, sample sizes are small and not completely balanced between different religious affiliations.

Figure F1.1: Mean evaluation of social appropriateness of allocation decisions in the *Religion* treatment, by religious identity of the **allocator**
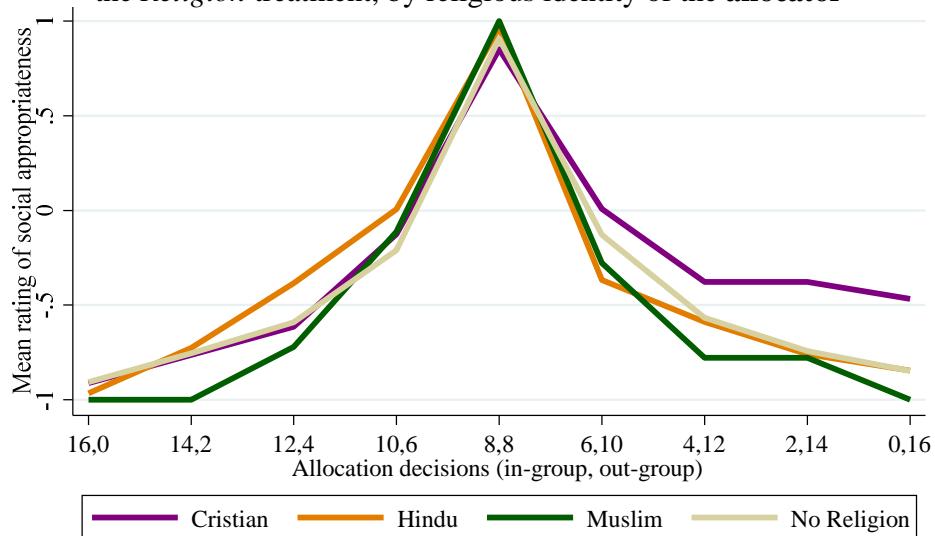

Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

Figure F1.2: Mean evaluation of social appropriateness of allocation decisions in the *Religion* treatment, by religious identity of the **out-group participant**
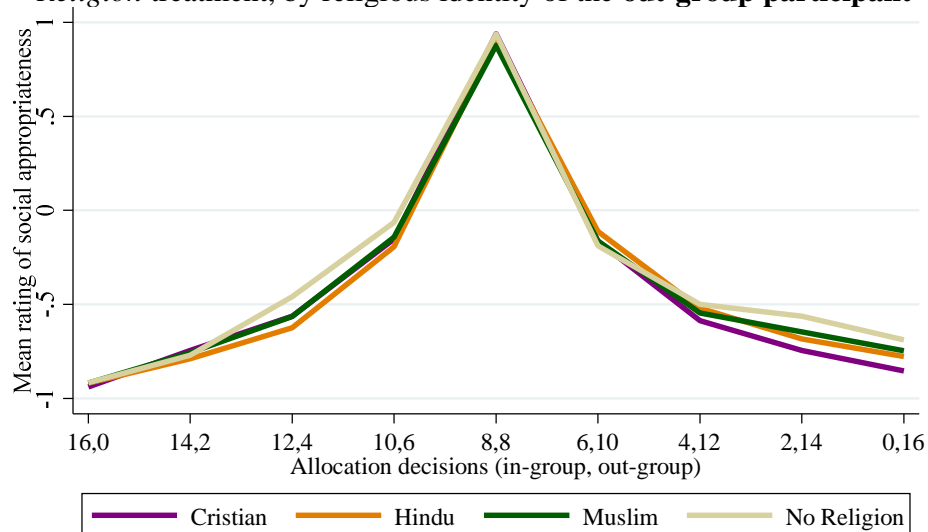
Figure F1.3: Allocation decisions in the *Religion* treatment, by political identity of the **allocator**
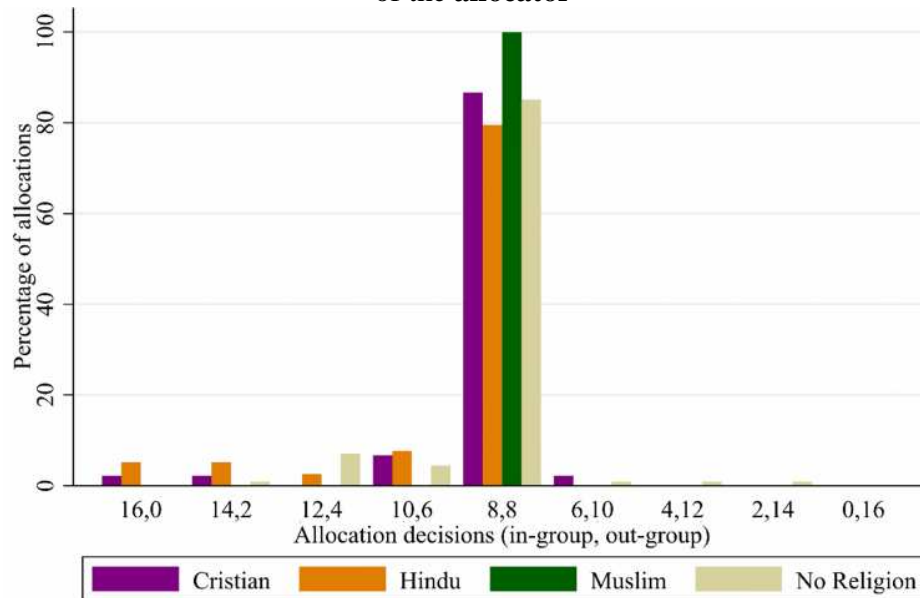
Note: Frequencies of allocation decisions in the *Religion* treatment. Coloured bars represent the religious affiliation of the **allocator.**

Figure F1.4: Allocation decisions in the *Religion* treatment, by political identity of the **out-group member**
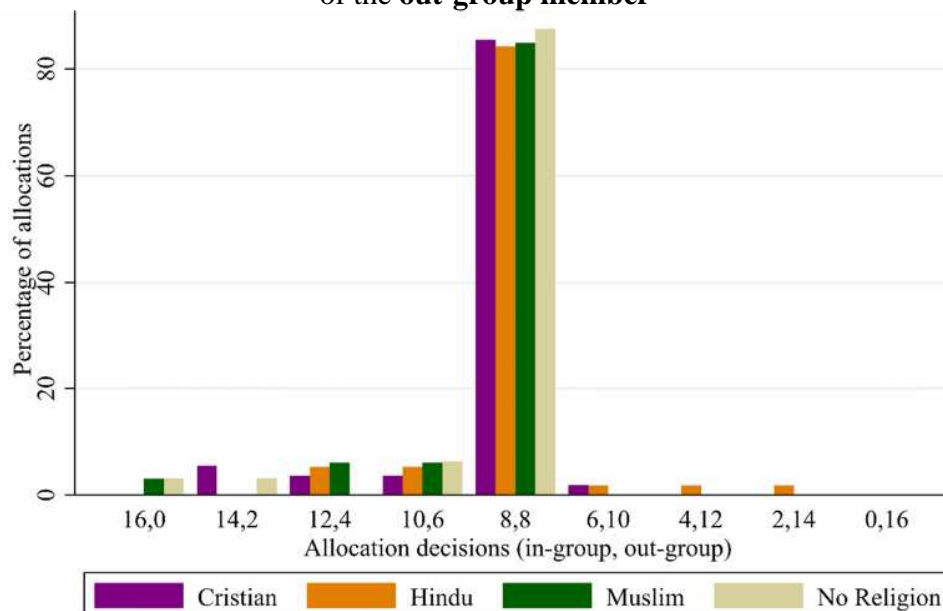
Note: Frequencies of allocation decisions in the *Religion* treatment. Coloured bars represent the religious affiliation of the **out-group** member. For example, those bars labelled as *Cristian*,

include the aggregated decisions of *Hindu*, *Muslim* and *Non-Religious* participants when allocating money between one member of their group and a *Cristian* participant.

Table F1.1: Regression analysis of allocation decisions in the *Religion* treatment

	Difference in amount allocated to in-group and out-group participant			
	(1)	(2)	(3)	(4)
Allocator:				
Cristian	1.149	1.747*		
	(0.931)	(1.013)		
Hindu	-0.8	-1.247		
	(1.383)	(1.388)		
Muslim	-0.168	0.027		
	(0.75)	(0.762)		
Out-group:				
Cristian			-0.672	-0.697
			(0.497)	(0.501)
Hindu			0.169	0.17
			(0.476)	(0.478)
Muslim			-0.05	-0.074
			(0.61)	(0.614)
Controls:	✗	✓	✗	✓
Constant	0.8	1.491	1.013**	3.31
	(.635)	(3.323)	(0.424)	(3.092)
Observations	210	210	210	210
R^2	0.028	0.090	0.016	0.061

Note: Results for random effects models. Omitted category both for allocator and out-group identity is *Non-Religous*. Control variables include: gender (binary variable), age, year of university degree the participant is in, if they are originally from a rural or urban background, and a categorical variable of annual household income. Robust standard errors are in parentheses. Stars indicate significance level: *** $p<.01$, ** $p<.05$, * $p<.1$.

F.2. Castellón, SPAIN

This section includes the detailed results within the Religion treatment in Spain. The *Muslim* group in this treatment was formed only by one person, therefore results are merely illustrative, but no further conclusions should be raised.

Figure F2.1: Mean evaluation of social appropriateness of allocation decisions in the *Religion* treatment, by religious identity of the **allocator**
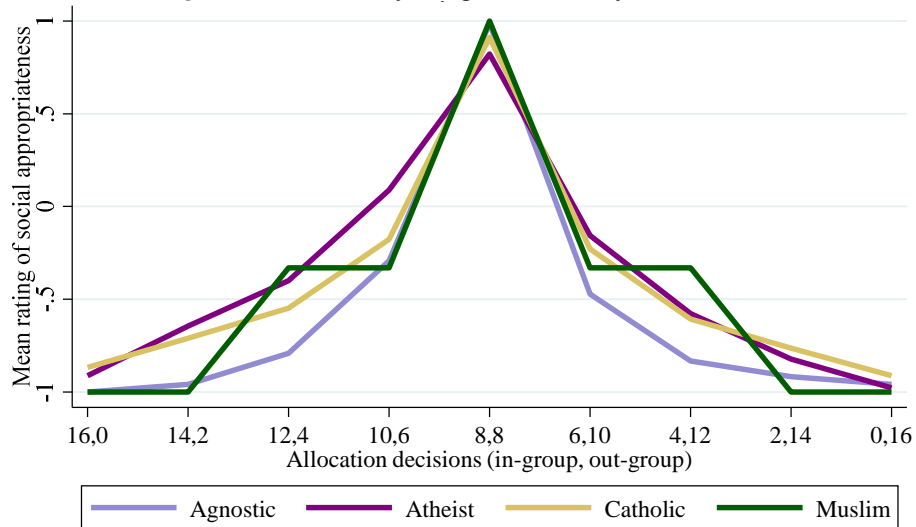

Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

Figure F2.2: Mean evaluation of social appropriateness of allocation decisions in the *Religion* treatment, by religious identity of the **out-group participant**

Note: Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

50
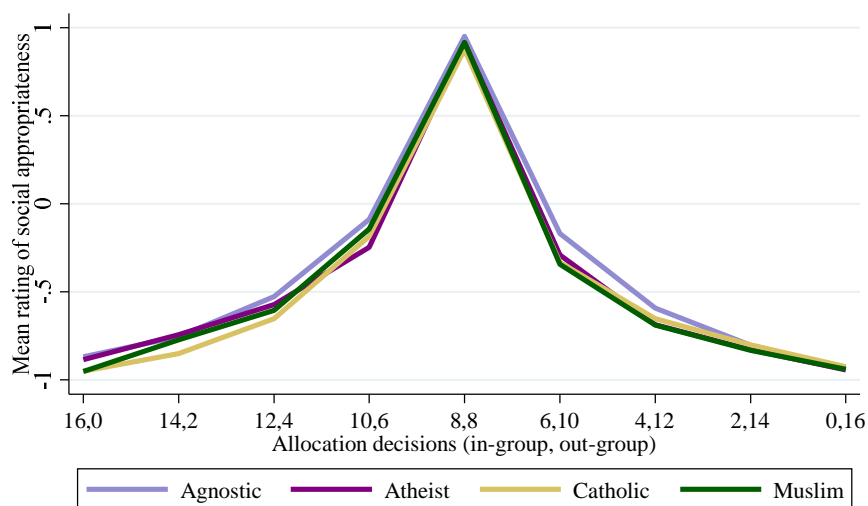
Figure F2.3: Allocation decisions in the *Religion* treatment, by political identity of the **allocator**
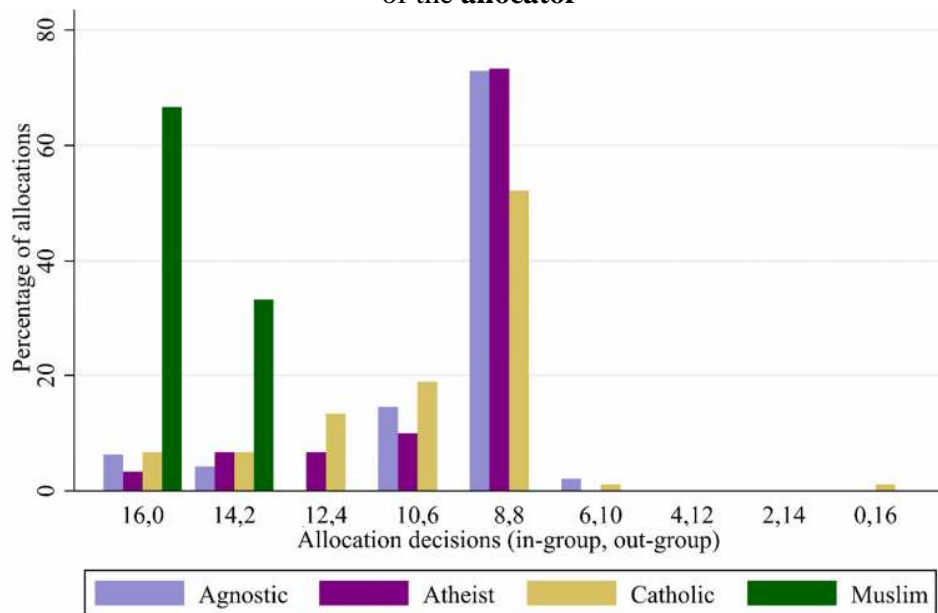
Note: Frequencies of allocation decisions in the *Religion* treatment. Coloured bars represent the religious affiliation of the **allocator.**

Figure F2.4: Allocation decisions in the *Religion* treatment, by political identity of the **out-group member**
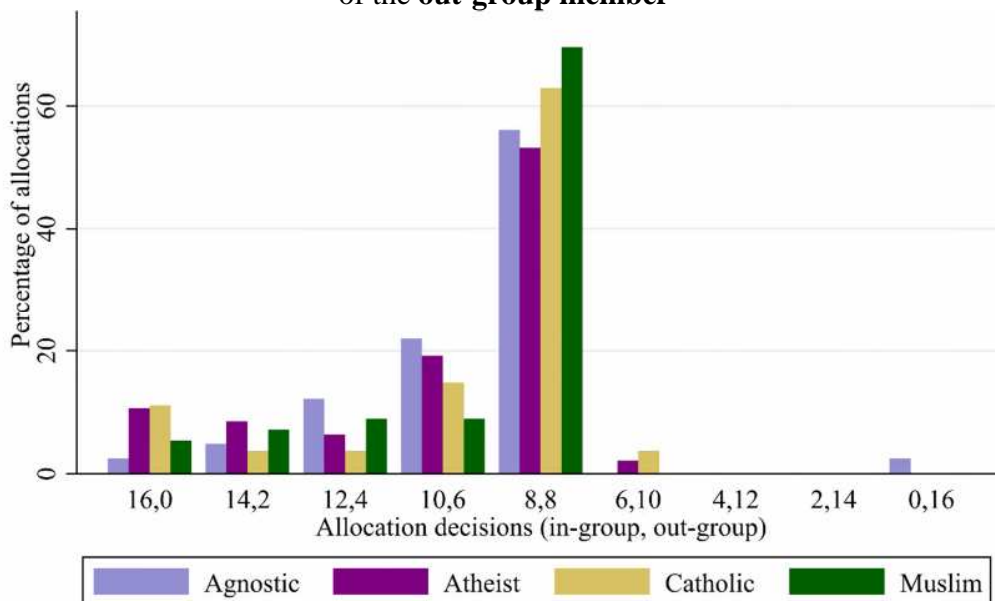
Note: Frequencies of allocation decisions in the *Religion* treatment. Coloured bars represent the religious affiliation of the **out-group** member. For example, those bars labelled as *Agnostic*, include the aggregated decisions of *Atheist*, *Catholic* and *Muslim* participants when allocating money between one member of their group and an *Agnostic* participant.

Table F2.1: Regression analysis of allocation decisions in the *Religion* treatment

| | Difference in amount allocated to in-group and out-group participant | | | |
	(1)	(2)	(3)	(4)
Allocator:				
Atheist	0.267	0.198		
	(1.579)	(1.792)		
Catholic	1.467	1.433		
	(1.212)	(1.332)		
Muslim	12.667***	12.843***		
	(4.037)	(4.419)		
Out-group:				
Atheist			1.761**	1.807**
			(0.882)	(0.882)
Catholic			1.27	1.373
			(1.074)	(1.076)
Muslim			0.845	0.903
			(0.838)	(0.836)
Controls:	✗	✓	✗	✓
Constant	2**	1.298	2.08***	4.552
	(0.979)	(3.822)	(.778)	(3.468)
Observations	171	171	171	171
R^2	0.105	0.119	0.010	0.034

Note: Results for random effects models. Omitted category both for allocator and out-group identity is *Agnostic*. Control variables include: gender (binary variable), age, year of university degree the participant is in, if they are originally from a rural or urban background, and a categorical variable of annual household income. Robust standard errors are in parentheses. Stars indicate significance level: *** p<.01, ** p<.05, * p<.1.

Appendix G. Robustness analysis

In Nottingham, some procedural errors were made in conducting the experiments. In the first *Artificial* session, a typo at one point on the instructions referred to "religion", where it should have said "ball colour grouping". At this point, the experimenter immediately verbally corrected the instructions to clarify what they should have said. Because these errors could in principle bias our results, we perform a robustness analysis to check whether our main results are driven specifically by the flawed sessions. The analysis does not suggest this is the case. Note that we also have incomplete allocation decisions from five subjects as a result of glitches in the software which prevented them from entering responses. We also noticed that one subject managed to participate twice, first in session 4 and also in session 7. Data from their second participation was excluded from the analysis.

Figure G1: Mean evaluation of social appropriateness of allocation decisions excluding session 1 of *Artificial* treatment (UK)
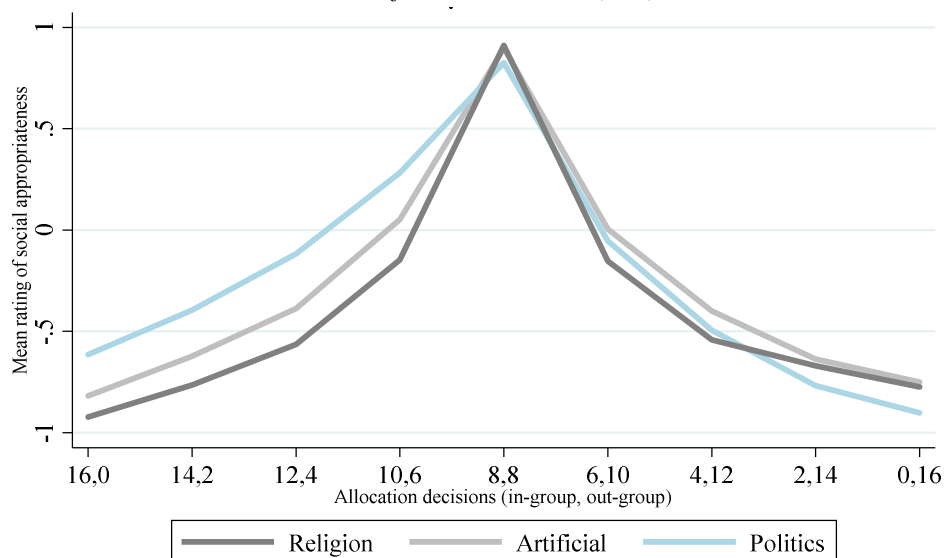
Notes:. Social appropriateness ratings range from -1 (if complete consensus on social inappropriateness) to 1 (in the case of absolute consensus on social appropriateness).

Table G1: Ratings of social appropriateness of decisions for the *Artificial* treatment excluding session 1 and updated p-values

	16-0	14-2	12-4	10-6	8-8	6-10	12-4	14-2	0-16
Mean rating of social appropriateness									
Artificial treatment	-0.82	-0.62	-0.39	0.05	0.90	0.00	-0.40	-0.64	-0.75
P-values									
Politics vs. Artificial	0.001	0.000	0.000	0.000	0.049	0.249	0.056	0.008	0.001
Artificial vs. Religion	0.012	0.004	0.000	0.000	0.766	0.006	0.012	0.540	0.710

Notes: Mean ratings are computed assigning values of 1, 0.33, −0.33 and −1 to assessments of "very socially appropriate", "somewhat socially appropriate", "somewhat socially inappropriate" and "very socially inappropriate", respectively. P-values from Fisher-Pitman permutation tests for two independent samples.

As it can be seen, none of our main findings change when excluding session 1 of the *Artificial* treatment. The distribution of the allocation decisions and ratings of social appropriateness remain equivalent, and the statistically significance is not altered. Therefore, we reject the idea that results could be driven even partially by that single flawed session.

Figure G2: Allocation decisions by treatment
excluding session 1 of *Artificial* treatment (UK)
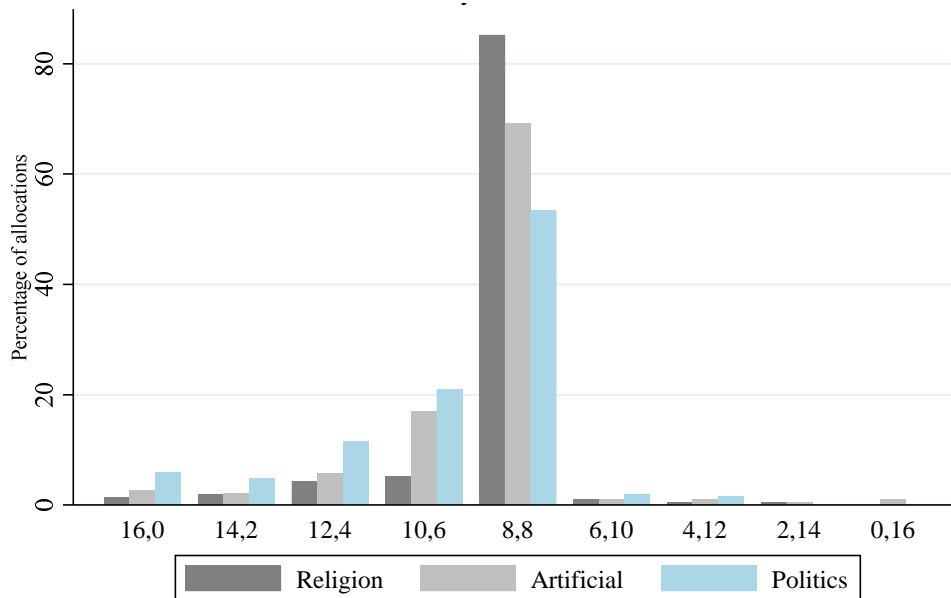
Notes: Percentages of allocation decisions, by treatment. Each participant made three allocation decisions, one for each different possible affiliation of the out-group participant.

In one *Religion* session in which the allocator game was conducted first, the full instructions from the previous *Religion* session (in which the norm-elicitation task was run first) were inadvertently left on subjects' desks and only spotted and removed after a few subjects had already taken their seats – therefore, these subjects may have briefly been exposed to the information about both tasks at the beginning of the experiment. Removing that session data from the analysis also replicates all our findings. Results are qualitative the same. This also corresponds with section H of the Appendix, in which we show the absence of order effects in our results (which would suggest exposing subjects to the second task ought not to affect their first-task responses).

Appendix H. Order-effects analysis

In this sections, we expand the discussion on possible order effect showing a regression analysis of said effects. Table G1 reproduces the allocation model we have used throughout the article and Appendix including an *Order* variable that captures whether the norm-elicitation task was presented first. We show that the order in which the tasks were carried out had no effect neither on allocation decisions nor ratings of social appropriateness.

Table H1: Results estimated considering order effects

	Allocation decisions	Social appropriateness ratings		
		(16,0)	(8,8)	(0,16)
Treatment:				
Politics	2.383***	0.138	-0.076	-0.234*
	(0.915)	(0.128)	(0.089)	(0.119)
Religion	-0.22	-0.076	-0.008	-0.009
	(0.786)	(0.11)	(0.077)	(0.103)
Order	-1.002	-0.126	0.14	-0.189
	(0.904)	(0.126)	(0.088)	(0.118)
*Order*Artificial*	0.955	0.08	-0.107	-0.027
	(1.207)	(0.169)	(0.118)	(0.157)
*Order*Politics*	-0.214	0.175	-0.081	0.236
	(1.255)	(0.175)	(0.122)	(0.164)
Constant	3.032	-0.602*	0.44*	-0.372
	(2.424)	(0.339)	(0.236)	(0.316)
Observations	712	712	712	712
R^2	0.0475	0.059	0.0556	0.059

Notes: The dependant variable for the allocation decisions is the *Difference in amount allocated to in-group and out-group participant*, the same variable used in all allocation models throughout the paper and appendix. For the norm-elicitation task, the dependant variables are the ratings of social appropriateness given for the two extreme allocations and the equal split. Results for random effects models. Control variables are included: gender (binary variable), age, year of university degree the participant is in, if they are originally from a rural or urban background, and a categorical variable of annual household income. Robust standard errors are in parentheses. Stars indicate significance level: *** p<.01, ** p<.05, * p<.1.

Appendix I. Supplementary materials

I1. Questions of pre-survey for Nottingham experiment

[The same questionnaire in Spanish was used in Castellón, except that the names of political parties were changed, and there was no Brexit question]

[For all multiple choice questions. subjects must select exactly one answer]

Q1: First. please enter your student ID number and university email address in the boxes below. Note that this information will NEVER be released as published data. or revealed to other participants. Please enter your student ID number here:

Q2: Please enter your university email address here:

Q3: If the email account associated with your Paypal account is different from your university email address. please enter it here. (this information will only be used for payment purposes)

Q4: How would you define yourself in religious terms?
- Catholic
- Muslim
- Jewish
- Protestant
- Hindu
- Sikh
- Buddhist
- No religion
- Other
- Prefer not to say

Q5: How often do you participate in worship or other religious events (apart from social ceremonies such as weddings and funerals)?
- Almost never
- A few times a year
- A few times a month
- Almost every week
- Prefer not to say

Q6: Which party did you vote for in the last UK General Election in December 2019?
- Conservative

- Labour
- Liberal Democrats
- Scottish National Party
- Green Party
- Brexit Party
- Plaid Cymru
- Other
- I did not vote
- Prefer not to say

Q7: Is there a particular political party you feel closer to than any other?
- Yes
- No Prefer not to say

Q8 [DISPLAYED ONLY IF ANSWER TO Q7 IS YES]: Following on from the previous question. please name this party that you feel closer to than any other.
- Conservative
- Labour
- Liberal Democrats
- Scottish National Party
- Green Party
- Brexit Party
- Plaid Cymru
- Other
- I did not vote
- Prefer not to say

Q9 [DISPLAYED ONLY IF ANSWER TO Q7 IS YES]: And how close do you feel to the party you have selected?
- Very Close
- Somewhat Close
- Not very close
- Prefer not to say

Q10: If the next general election were held tomorrow. which party would you vote for?
- Conservative
- Labour
- Liberal Democrats
- Scottish National Party

- Green Party
- Brexit Party
- Plaid Cymru
- Other
- I would not vote[19]
- Prefer not to say

Q11: To what extent are you interested in politics? Please. place yourself in a scale between 0 and 10 where 0 means you are not interested at all and 10 you are very interested.
(If you prefer not to answer. you may skip this question)

```
0    1    2    3    4    5    6    7    8    9    10
```

Q12: When people talk about politics they normally use the words 'left' and 'right'. Could you place yourself in a scale from 0 to 10 where 0 means left. 10 right and 5 the centre?
(If you prefer not to answer. you may skip this question)

```
0    1    2    3    4    5    6    7    8    9    10
```

Q13: In hindsight. do you think Britain was right or wrong to vote to leave the European Union? In other words. do you think Brexit was right or wrong?
- Brexit was right
- Brexit was wrong
- I don't know / prefer not to say

---

[19] In the first wave in which we sent out this survey, this option was incorrectly presented as "I did not vote". We corrected this subsequently.

I2. Instructions for Nottingham experiment

Instructions for subjects in the Politics treatment, playing allocator game first

Instructions

Welcome to this experiment. This is an experiment about decision-making. During the experiment, we request that you remain quiet and do not attempt to communicate with other participants. Mobile phones must not be used during the experiment. Participants not following these requests may be asked to leave without receiving payment. If you have any questions, please raise your hand and the experimenter will come to you. For your participation, you will be paid a show-up fee of £4. You may also receive some additional money based on your choices and the choices of others in the tasks described below. All payments will be made by Paypal within the next 24 hours.

There will be two tasks for all participants to perform. At the end of the experiment, the experimenter will toss a fair coin. If it lands on heads, all participants will receive payment for the first task only; if it lands on tails, all participants will receive payment for the second task only. As you will not know until the end of the experiment which task you will receive payment for, *please make your decisions in each task carefully*. You will not receive feedback on the outcome of any task until after the end of the experiment, and your decisions in the first task will have no effect on the nature or outcome of the second task. You will not receive any instructions for or information about the second task until you have completed the first task. After the second task, there will also be a questionnaire. The anonymity of your responses to all parts of all tasks and questions is guaranteed.

Please now answer two questions on your screen, to ensure you understand the process of the experiment.

In this experiment, participants support four different political parties**: Conservative, Green, Labour and Liberal Democrat**[20]. Your **political** grouping is based upon **your response to a question in a survey you completed last month, in which you were asked who you would vote for in the next General Election if it were held tomorrow**. Hereafter, we will refer to **the party you stated you would vote for** as your **political** grouping. Your **political** grouping will be relevant to your interactions with other participants in this experiment.

---

[20] Those words and expressions in bold were changed across treatments, and represent our experimental manipulation.

Task One

In this experiment, one third of you will be randomly assigned by the computer into a role entitled 'Individual A'. The decisions made by Individual As during the task will determine the payments from the task received by the other two thirds of participants. Exactly who the Individual As are will not be revealed until after the experiment. In the meantime, we ask all participants to make decisions **as if** they are an Individual A.

Please make each decision carefully, as it may be used to determine participants' payments.

Assume for the rest of this section that you are an Individual A. Your task will be to decide how to divide £16 between two other participants in the experiment, one who has the same **political** grouping as you, and another who has a different **political** grouping from you. You may divide the money any way you like so long as the amount allocated to each person is a multiple of two. You may not allocate any of the money to yourself. However, you will also receive a payment. This payment might be £6, £8 or £10. This will be randomly decided at the end of the experiment by the computer, which is equally likely to select any of these amounts.

The participant from a different **political** grouping with whom you are matched in this task might possibly belong to any one of the other three groupings. Since you will not know which of the other three groupings this participant belongs to, we ask you to make your decision three times – once for each possible grouping this participant might belong to. We will only implement the decision corresponding to the grouping that this participant actually belongs to.

To clarify: imagine that the four groupings in this experiment were **Party 1, Party 2, Party 3 and Party 4**; and that you support **Party** 1. You would be asked to make three decisions:

Decision 1: How to divide £16 between one participant supporting **Party** 1 and one participant supporting **Party** 2

Decision 2: How to divide £16 between one participant supporting **Party** 1 and one participant supporting **Party** 3

Decision 3: How to divide £16 between one participant supporting **Party** 1 and one participant supporting **Party** 4

Then, if it turns out that the other-party participant you are matched with in this task supports **Party** 3, your Decision 2 would be used to determine participants' payoffs, while your Decisions 1 and 3 would not be implemented.

Please now answer three questions on your screen, to ensure you understand this part of the experiment.

<u>**Task Two**</u>

In the second part of this experiment, you will receive a description of a situation. This description corresponds to a situation in which one person, "Individual A," must decide how to act. You will be given a description of various possible actions Individual A can choose to take.

After you receive the description of the situation, you will be asked to evaluate each of the various possible actions Individual A can choose to take. You must indicate, for each of the possible actions, whether taking that action would be "socially appropriate" or "socially inappropriate". By socially appropriate, we mean behaviour that you think most participants of your **political** grouping in this experiment would agree is the "correct" thing to do. Another way to think about what we mean is that if Individual A were to select a socially inappropriate action, then another participant of your **political** grouping might be angry at Individual A.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behaviour.

To give you an idea of how the experiment will proceed, we will go through an example situation and show you how you will indicate your responses.

Example Situation

Individual A is at a local coffee shop near campus. While there, Individual A notices that someone has left a wallet at one of the tables. Individual A must decide what to do. Individual A can choose four possible actions: take the wallet, ask others nearby if the wallet belongs to them, leave the wallet where it is, or give the wallet to the shop manager.

The table below presents the list of the possible actions Individual A can choose. For each of the actions, you would be asked to indicate whether you believe choosing that action is very socially inappropriate, somewhat socially inappropriate, somewhat socially appropriate, or very socially appropriate. To indicate your response, you would click on the corresponding button.

	Take the wallet	Ask others nearby if the wallet belongs to them	Leave the wallet where it is	Give the wallet to the shop manager
	The table below presents all actions Individual A can possibly take. Please tick one box for each action corresponding to how socially appropriate you think the action is.			
Very socially inappropriate	○	○	○	○
Somewhat socially inappropriate	○	○	○	○
Somewhat socially appropriate	○	○	○	○
Very socially appropriate	○	○	○	○

Submit

If this was the situation for this study, you would consider each of the possible actions above and, for that action, indicate the extent to which you believe taking that action would be "socially appropriate" or "socially inappropriate". Recall that by socially appropriate we mean behaviour that most participants of your **political** grouping agree is the "correct" thing to do.

For example, suppose you thought that taking the wallet was very socially inappropriate, asking others nearby if the wallet belongs to them was somewhat socially appropriate, leaving the wallet where it is was somewhat socially inappropriate, and giving the wallet to the shop manager was very socially appropriate. Then you would indicate your responses as follows:

If you have any questions about this example situation or about how to indicate your responses, please raise your hand now.

You will next be given the description of a situation where Individual A, a participant in an experiment, has to choose between various possible actions. After you read the description, you must consider the possible actions and indicate on your computer screen how socially appropriate these are in a table similar to the one shown above for the example situation.

After this, the computer will randomly select one participant of your **political** grouping. The computer will then randomly select one action Individual A can choose. Your evaluation of this action will be compared with that of the randomly selected participant of your **political** grouping. If your evaluation is the same as theirs, you will receive £8 for this task; otherwise you will receive zero.

For instance, imagine the example situation above was the actual situation and the possible action "Leave the wallet where it is" was selected by the computer. If your evaluation had been "somewhat socially inappropriate" then your task earnings would be £8 if the person you are matched with also evaluated the action as "somewhat socially inappropriate" and zero otherwise.

The situation

The situation you are asked to evaluate is like the one you participated in in the previous task. Here is a summary.

Individual A is taking part in an experiment in this lab. The room contains participants who support four different **political parties**: **Conservative, Green, Labour and Liberal Democrat.** Hereafter, the party each participant supports is referred to as their political grouping. The anonymity of Individual A's decisions in the experiment is guaranteed.

Individual A's task will be to decide how to divide £16 between two other participants in the experiment, one who has the same **political grouping** as Individual A, and another who has a different **political grouping** from Individual A. Individual A may divide the money any way they like so long as the amount allocated to each person is a multiple of two. Individual A may not allocate any of the money to themself. However, Individual A will also receive a payment. This payment might be £6, £8 or £10. This will be randomly decided at the end of the experiment by the computer, which is equally likely to select any of these amounts.

The participant from a different **political** grouping with whom Individual A is matched in this task might possibly belong to any one of the other three groupings. Since Individual A will not know which of the other three groupings this participant belongs to, Individual A is asked to make their decision three times – once for each possible grouping this participant might belong to. Only the decision corresponding to the grouping that this participant actually belongs to will be implemented.