

# Political Identity and Trust

Pablo Hernandez-Lagos

Dylan Minor

New York University, Abu Dhabi

Northwestern University

December, 2018

## Abstract

We explore how political identity affects trust. In particular, we examine the extent to which political identity and objective information shape perceptions about others' trustworthiness. Using an incentivized experimental survey over a sample of the general U.S. population, we vary information about partners' political identity to elicit trust behavior, beliefs about others' trustworthiness, and actual reciprocation. We find that beliefs depend on the political identity of the partner, but they are not always biased against out-groups. This suggests that the cross-party antagonism found in the literature does not necessarily translate into pessimism over what out-groups would do. We also find that people believe others are much less trustworthy than they actually prove to be. We then attempt to correct beliefs by disclosing historical trustworthiness. Subjects' beliefs shift only slightly, suggesting that incorrect stereotypes are difficult to correct.

**Keywords:** *Trust, Beliefs, Political Identity, Polarization*

Political polarization in the United States has increased and partisan antagonism is deeper and more extensive than at any point in the last two decades. Evidence suggests that political polarization based on partisan identity hinders cross-partisan trust, creating political and economic gridlocks (Abramowitz 2010). Such political identity stems from a strong emotional connection between the individual and the party (Mason 2014) and may lead to sentiments of dislike, anger, and even loathe towards political opponents (Fiorina and Abrams 2008; Goren, Federico and Kittilson 2009; Iyengar, Sood and Lelkes 2012; Mason 2014; Iyengar and Westwood 2014; Huddy, Mason and Aaroe 2015; McConnell, Margalit, Malhotra and Levendusky 2017).

In this paper, we explore the effect of political identity on trust. We focus on trust—the willingness to accept vulnerability based on *positive expectations* about the behavior of others—as it is fundamental for economic and political organization, and has been found to be stronger in societies with more inclusive institutions and with more representative democracies (see, e.g., Arrow 1974; Inglehart 1990; Putnam, Leonardi and Nanetti 1994; Knack and Keefer 1997; Alesina and La Ferrara 2002).<sup>1</sup>

Political identity affects citizens’ trust through two fundamental mechanisms: Sentiments that lead to discrimination based on identity (“taste for discrimination”) and stereotypes about the behavior of others. A taste for discrimination represents individuals’ willingness to give away money (i.e., profits, wages, or income) to cater to their prejudice (Becker 1957, 1993). This prejudice is part of one’s individual preferences and may reflect animus and sentiments of dislike, loathe, anger, or similar emotions towards certain groups. Stereotypes, on the other hand, are simplified perceptions about the behavior of a group of people. These perceptions affect individual behavior through beliefs rather than preferences (Arrow 1972; Phelps 1972).

Beliefs about the behavior of others do not entail direct enjoyment or gratification. Instead, beliefs serve to assess the benefit or damage produced by other people’s actions. The literature documents that individuals enjoy favoring in-groups, and sometimes even hurting

---

<sup>1</sup>An interdisciplinary definition of trust comes from Rousseau, Sitkin, Burt and Camerer (1998):

“Trust is a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intention or behaviors of others.”

As we describe below, we refer to *trust behavior* as the action of accepting vulnerability in a transaction. We label the expectation about cooperative behavior of others as *trust beliefs*. Finally, we label the action that honors trust as *trustworthiness* or *reciprocity*. We use these labels because our set-up allows us to disentangle these three aspects of trust.

out-groups, regardless of how in-groups or out-groups actually behave (Iyengar et al. 2012; Iyengar and Westwood 2015; McConnell et al. 2017). This paper focuses on how political identity affects beliefs about behavior. In particular, we explore whether political identity evokes negative perceptions about the trustworthiness of out-groups, whether perceptions about trustworthiness are correct and whether they can be changed.

Beliefs about the trustworthiness of others, which we label *trust beliefs*, are important for a number of reasons. Trust beliefs that are biased against out-groups can open deep and persistent cleavages in the society. For example, given that polarization based on political identity seems to be at least as strong as polarization based on race (Iyengar and Westwood 2015), stereotypes based on political identity may influence lending, employment offers, or school admissions, just as stereotypes based on race do (see, e.g. List 2004; Gift and Gift 2015). Moreover, stereotypes and preferences do not always coincide. A lender may dislike individuals who identify with the opposite party, yet be willing to extend credit to an opposite-party applicant if she expects the applicant will pay back the loan. In this case, the pecuniary gains expected from trustworthy behavior of the borrower override the intrinsic animosity towards out-groups.

Beliefs about the behavior of others are also important because they can undermine democratic institutions directly. Evidence suggests that people’s negative expectations about the behavior of incumbent groups may affect the incumbent’s actual performance (Hetherington 1998; Hetherington and Husser 2012). It has also been documented that incumbent groups may be unwilling to surrender power if they anticipate opportunistic behavior by opposition groups (Inglehart 1990, 1999). In the same way, opposition groups may hamper well-intended policies if they hold negative expectations about the behavior of incumbent groups (Alesina, Baqir and Easterly 1999). Moreover, biased trust beliefs are potentially difficult to correct because political identity exacerbates the tendency to form expectations that cater to pre-existing views (Campbell, Converse, Miller and Stokes 1960; Kunda 1990; Taber and Lodge 2006; Gerber and Huber 2010; McGrath 2017). As a result, individuals often disregard disconfirming information, which may lead them to persistently and incorrectly believe that out-groups are not trustworthy (e.g., Green, Palmquist and Schickler 2002; McCarty, Poole and Rosenthal 2006; Bénabou and Tirole 2011; Nyhan and Reifler 2010; Bordalo, Coffman, Gennaioli and Shleifer 2016; Lelkes and Westwood 2016).

We conduct experimental surveys in which we vary the political identity (in terms of

partisan identity) and measure a proxy for trust behavior, beliefs about trustworthiness, and actual trustworthiness across a sample of the general U.S. population. In each experimental survey, we run a standard version of a trust game. In this game, Player A (the sender) can either choose to allocate \$5 for him/herself and \$5 for Player B (the receiver) or defer the decision to Player B, in which case Player B can either choose to double the amount each one of them gets (\$10 respectively) or behave opportunistically and increase his/her own payoff (to \$14) and reduce Player A's payoff (to \$0). This procedure allows us also to elicit trust beliefs directly, by asking Player A to report his/her beliefs about the trustworthiness of Player B and rewarding Player A's accuracy. This game is labeled a "trust game" because Player A choosing to defer the decision to Player B is an indication of what we want to measure: An intention to accept vulnerability based upon positive beliefs about the behavior of others.

One important task when eliciting trust beliefs is to overcome the potentially exaggerated pessimism about out-groups' trustworthiness. Incentivizing truthful reporting of beliefs often leads to more sensible responses and less "party cheerleading" (Bullock, Gerber, Hill and Huber 2015). Using a trust game also helps us address common challenges of survey questions, such as ambiguity in language (Zak and Knack, 2001). Our game features a sender and a receiver with clearly stated monetary outcomes. Another common challenge with surveys is a low response rate, which can create sample selection issues. With a highly incentivized game, however, we reach a 97% response rate. We also opt for this research design because by measuring a proxy for trust behavior as well as actual trustworthiness, we can assess the extent to which trust beliefs drive trust behavior, and whether trust beliefs are statistically correct.

The experimental surveys are organized in two experiments. The first experiment (*main treatments*) varies the political identity of Player A and Player B, but does not provide any other information regarding the identity of the matched partners. The second experiment (*belief treatments*) features the same trust situation as in the main treatments, except now Player A also receives information about the actual trustworthiness corresponding to previous Players B. For example, if a Democrat Player A is matched to a Republican Player B in one of the belief treatments, we disclose to Player A the percentage of Republican Players B who reciprocated trust in the corresponding main treatment.

The data from the main treatments reveal variation in Player A's trust beliefs in each of

the possible identity matches. The data also show a shared belief, common across parties, that Democrats are more trustworthy. People’s behavior is consistent with this in general, as Democrats trust people identified with the other party less, whereas Republicans trust those identified with the other party slightly more. However, college-educated Democrat Players A believe that Democrats are not more trustworthy than Republicans. This suggests that the cross-party antagonism found in the literature does not always translate into pessimism over what out-groups would do. These results highlight the importance of beliefs about others’ behavior, as opposed to preferences.

Another result is that optimistic trust beliefs are associated with trust, but the relationship is stronger in in-group matches. We interpret these results as suggesting that political identity affects stereotypes in subtle ways and that such stereotypes do matter for trust.

Stereotypes, however, are not statistically correct. Regardless of political identity, participants reciprocate more often than expected. In addition, Republicans (who are thought to be less trustworthy by Democrats) reciprocate slightly more often than Democrats. Using the data from both experiments, we find that Republicans and Democrats react differently to new information. Democrats’ beliefs about the trustworthiness of others are generally incorrect (pessimistic) and favor in-groups, but new information makes them more accurate. Republicans’ beliefs are also pessimistic, but they show less in-group bias; nevertheless, new information does not change them.

Our paper builds on the literature that studies how affective polarization impacts political and economic behavior (Gerber and Huber 2009; Gerber et al. 2010; Iyengar et al. 2012; Iyengar and Westwood 2015). McConnell et al. (2017) find that individuals forgo money in order to favor (hurt) in-groups (out-groups) in cross-partisan interactions, which is consistent with a preference for discrimination. We complement this paper in that we focus on beliefs about the behavior of others rather than preferences, and also in that we consider an interaction that involves trust rather than (unilateral) giving. A related literature in political economy documents how information about political events and behavior of others changes individual behavior (see, e.g., Oliver and Wood 2014; Carey, Nihan, Valentino and Liu 2016; Perez-Truglia and Cruces 2017; Bursztyn, Egorov, and Fiorin 2017). We contribute to this literature by checking whether information has an effect on expectations about the behavior of others and whether such expectations explain behavior.

Expectations are at the core of trust behavior, yet measuring them is often challenging

(see, e.g., Bullock et al. 2015). By incentivizing truthful reporting, we can measure differences in expectations across political identity. In addition, by revealing actual behavior we can assess the extent to which objective information corrects misperceptions about the behavior of others. It has been documented that misperceptions about policy-relevant facts are hard to change and sometimes they are even strengthened by revealing disconfirming information (Taber and Lodge 2006; Nyhan and Reifler 2010; Bullock et al. 2015). To contribute to this literature, we study whether misperceptions about trustworthiness of others can be corrected.

Our paper also relates to the large literature on the drivers of trust using trust games. Trust games are a widely accepted method to elicit and measure important aspects of trust (Berg, Dickhaut and McCabe 1995; Falk and Zehnder 2013; Garbarino and Slonim 2009; Carlin and Love 2013). One recurrent finding in this literature is that mistrust toward out-groups is often driven by stereotypes (related to age, gender, income or ethnic groups) and not by a preference-driven taste for discrimination (see, e.g., Garbarino and Slonim 2009; Falk and Zehnder 2013; and Fershtman and Gneezy 2005).<sup>2</sup> To our knowledge, two papers study trust behavior in the context of political identity: Anderson, Mellor and Milyo (2005) and Carlin and Love (2013). Both use laboratory experiments and ask students from U.S. educational institutions to participate. Both papers find higher levels of trust among individuals who share the same partisan identity or political orientation. One key difference between these papers and our study is that we focus on beliefs about trustworthiness in addition to trust behavior. Another important difference is that we explore the determinants of trust beliefs by varying the political identity of the partner and the information about the trustworthiness of others. Further, our participants are drawn from the general U.S. population.

We see the contributions of this paper as follows. First, it explores the role of stereotypes, rather than preferences, on trust. Second, the paper assesses whether political identity evokes biased stereotypes. Third, it studies whether stereotypes can be changed by revealing objective information. Finally, the paper uses a sample of the general U.S. population and features a relatively well-paid survey experiment.

---

<sup>2</sup>Other psychological elements may affect trust as well such as guilt-aversion (Charness and Duwfenberg 2006), betrayal-aversion (Bohnet and Zeckhauser 2004) and risk-aversion (e.g., Eckel and Wilson 2004). Risk-aversion is the most widely studied phenomenon, but risk-aversion is found to be weakly related to trust, at best.

## Experimental Design

The main experiment (main treatments), described in this section, seeks to examine how trust behavior and beliefs change when we manipulate partisan identity of individuals. The second experiment (belief treatments) intends to assess the effect of information on beliefs. This was run after the main experiment and is described in the section on manipulating beliefs. Both experiments use the Kellogg School of Management E-lab system, which maintains a pool of 7,045 participants from across the United States. E-lab staff pre-screens individuals in this subject pool through a survey instrument from which partisan inclination and other demographic information are collected. Subjects in this pool are then provided an opportunity to periodically participate in research surveys sponsored by faculty. We use the information regarding partisan inclination to determine who receives the experimental survey (we provide the details below). The main treatments were run in December 2013 and the belief treatments in November 2014.

In the implementation of the design, the questions were asked in order, and once participants moved on to the next one, they were not able to go back to the previous question. Individuals did not receive any feedback regarding the results of their and their partner's decisions while they were answering the experimental survey. The procedure involved individual decision making, as in any other survey experiment; the only difference was that they were told that their decision on each question would be matched to the decision of a partner to compute payoffs. The exact procedure was as follows.

Each potential participant received an invitation to participate in the experimental survey.<sup>3</sup> Balancing parsimony and the need to identify both trust behavior and beliefs about partners' trustworthiness, the survey consisted of 8 questions: the first 4 were incentivized and the last 4 were not.

From the incentivized questions, the first one was a standard dictator game, in which subjects were told to allocate \$5 anonymously between themselves and another participant. We did not mention political affiliation to subjects so as to obtain a measure of participants' unconditional altruism and to avoid cueing on the subsequent trusting decision.

The second question corresponded to the sender role (Player A) in the trust game shown

---

<sup>3</sup>The survey was administered via Qualtrics. A copy of the survey instrument can be found in the online appendix.

in Figure 1.<sup>4</sup> Participants had a choice of trusting or not trusting a matched partner (Player B), who would then make a final allocation decision. If the sender decided not to trust, each participant received \$5. If Player A decided to trust Player B, the decision left payoffs as a function of Player B's choice. Our intervention started in this second question. We varied the identity of Player B by letting Player A know that the otherwise-anonymous Player B identified as either a Democrat or Republican. We also ran a baseline treatment in which there was no such mention of the political identity of the subject's partner.

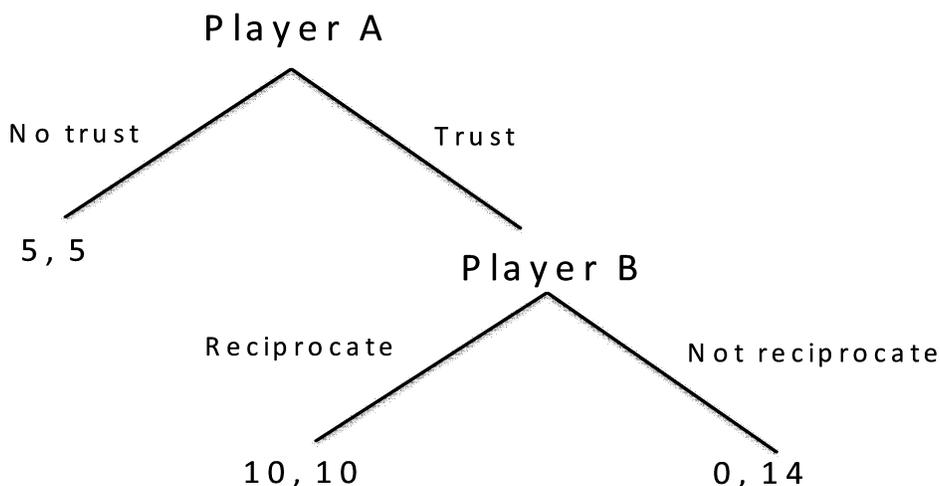


Figure 1: **Game tree.**

The third question asked the participant to act as Player B, making an allocation choice if entrusted with a decision by Player A. The participant had to decide whether to allocate \$10 to each player or to behave opportunistically and take \$14 for him or herself while leaving \$0 for Player A. In this question, we varied the identity of Player A to be Democrat or Republican. We also ran a baseline treatment in which Player A's political identity was not revealed.<sup>5</sup>

Crucial to our analysis is the fourth question, in which we elicited the participant's belief

<sup>4</sup>This game is similar to that in Charness and Dufwenberg (2006).

<sup>5</sup>We decided to make the reciprocation payoff also the utilitarian payoff:  $10 + 10 > 0 + 14$ . This may conflate reciprocation motives with utilitarian motives. We are willing to face this trade-off for two reasons. First, we want to give the reciprocation outcome an additional appeal because, before running the experiments, we expected little cross-party reciprocation given the results in the literature. The lack of cross-party variation could have entailed little variation in beliefs, which is something we wanted to avoid. Second, we are interested in differences in trust beliefs and behavior across treatments due to exogenous variation of the identity configuration of the pair, keeping payoff configuration constant.

about what proportion of those in the role of Player B (in the previous question) would prove to be trustworthy. A payment of \$3 was awarded if the participant predicted the percent of participants in the sample who would choose the (\$10,\$10) option when given the role of Player B, to be within 10 percentage points of the actual value.<sup>6</sup> As in the previous two questions, we varied the information regarding the identity of those in the role of Player B as either Democrat or Republican according to their response in the pre-screen survey. Consistent with the other questions, we also ran a no-identity baseline treatment.

Each participant received only one type of survey instrument. That is, we fixed the identity of the partner across questions. For example, when a participant was told in the second question she will be matched with a Democrat Player B, in the third question she was also told she will be matched with a Democrat Player A, and in the fourth question she had to state her beliefs about the proportion of Democrats in the role of Player B who would prove to be trustworthy. We followed the same pattern when the partner was a Republican and when his or her identity was not revealed. We used the so-called “strategy method” to elicit behavior: participants made decisions individually, and those decisions were matched across subjects after everyone responded to the survey to compute payoffs (for a comparison between the strategy method and the direct-response method, see Brandts and Charness, 2011). The outcomes of Player A’s and Player B’s decisions were paid in full. We believe this facilitated participants’ understanding of the procedure and maximized their participation, which for us was an initial concern as they were drawn from the general population.<sup>7</sup>

The next four non-incentivized questions presented the Cognitive Reflection Test from Frederick (2005) and asked for political orientation (from very liberal to very conservative), income range, and partisan identity (Republican, Democrat, Independent, or Other). Except for the Cognitive Reflection Test, these demographic questions used the exact same language as in E-lab’s pre-screen survey.<sup>8</sup>

We used the partisan inclination information from E-lab’s pre-screen survey to identify participants. E-lab sent the experimental survey to 250 Democrats and 250 Republicans in

---

<sup>6</sup>When studied in isolation, this belief elicitation procedure is incentive compatible even when subjects are risk averse. See the supplementary appendix for details.

<sup>7</sup>Paying in full may raise concerns about some participants under-reporting beliefs in order to hedge against Player B’s opportunistic behavior if entrusted. In the supplementary appendix, we show that hedging motives are practically non-existent, which is consistent with the results in Blanco et al. (2010, p. 415).

<sup>8</sup>We find that subjects’ answers to the last question are consistent with those given in the E-lab pre-screen survey, sometimes many months earlier. See the online appendix for details.

total. The baseline treatment survey (no partner identification) was sent to 100 out of the 250 Democrats, the survey type identifying the partner as a Democrat was sent to 100 of the remaining 150 Democrats, and the survey type identifying the partner as Republican was sent to the final 50 Democrats. The response rate was 100/100, 100/100, and 47/50, respectively. An analogous procedure was conducted with the 250 identified Republicans with response rates 95/100, 99/100, and 44/50, respectively. In total, we had 485 participants (97% response). Table 1 shows the treatments.

Table 1: **Main treatments.**

<b>Partner's ideology</b>	<b>Participant's ideology</b>	
	Democrat (D)	Republican (R)
Not revealed (NR)	<b>D-NR (N=100)</b>	<b>R-NR (N=95)</b>
Democrat (D)	<b>D-D (N=100)</b>	<b>R-D (N=44)</b>
Republican (R)	<b>D-R (N=47)</b>	<b>R-R (N=99)</b>

After the surveys were completed, we randomly matched subjects to compute payments. The survey payments averaged \$20 per participant, and it took on average less than three minutes to complete the survey. Payments were made via electronic Amazon gift cards within a week of completing the experiment. Subjects were allowed to complete only one experimental survey.

Our design is not without limitations. Using a game to elicit trust behavior and beliefs allows us to have clear measures, but it also generates the usual external validity and experimenter effects concerns. We try to minimize the role of external validity issues by using participants drawn from the general U.S. population (as opposed to student participants) in our experiment. This, however, still leaves open the question of whether a trust game captures general trust in the population. The answer to this question depends on the reader's particular notion of trust. Nevertheless, the trust game does indeed capture what we are trying to measure: Players A "intention to accept vulnerability based upon positive expectations of the intentions or behavior of others" (see e.g., Butler, Giuliano and Guiso 2015; Rousseau et al. 1998). Another concern is that participants may have felt compelled to answer the questions in a particular way to satisfy the experimenters. Since we are interested in differences in behavior rather than absolute magnitudes, experimenter effects are not an issue as long as Republicans and Democrats do not respond differently to the same

experimenter, on average. In order to minimize such differential behavior, our design features relatively high payoffs to incentivize profit maximizing behavior, keeps payoff constant across treatments, and does not prime subjects by asking for their own political identity prior to their decisions.

## Hypotheses

We use a simple framework to describe explicit channels through which political identity can shape beliefs and behavior. We assume Player A’s trust depends on her beliefs about Player B’s trustworthiness. These beliefs reflect a perception about the behavior of others, initially acquired through cultural transmission and subsequently updated through experience (Guiso, Sapienza and Zingales, 2008; Butler et al., 2015). Revealing the political identity of Player B provides Player A with a cue to infer the behavior of Player B. In other words, we assume that Player A’s beliefs  $p = p(I, \mathbf{x}) \in [0, 1]$  are a function of the information about Player B’s political identity  $I = s(ame), o(ther)$ , and a vector of background characteristics related to belief formation,  $\mathbf{x} \in \mathbb{R}^N$ .

Beliefs are often described as priors updated by information coming from biased sources (e.g., similar people, see Ortoleva and Snowberg, 2015). Individuals choose sources that reinforce their identity, which promotes their holding of more optimistic beliefs about in-group members (see, e.g., Taber and Lodge, 2006; Mason 2014; Lelkes and Westwood 2016). We should expect therefore that  $p(s, \mathbf{x}) > p(o, \mathbf{x})$  for all  $\mathbf{x}$ . In words:

**Hypothesis 1.** *Participants believe individuals with the same political identity are more likely to reciprocate trust than individuals with different political identity.*

Trust is an intention to accept vulnerability based upon positive expectations about the behaviors of others. To make explicit the role of beliefs (expectations), we need to account for the role of a taste for discrimination (animus) towards out-groups. We assume that Player A benefits from both players’ monetary gains (e.g., Charness and Rabin 2002) and such benefit depends on whether political identities coincide. That is, Player A’s utility is given by

$$u_A = u_A(\pi_A, \pi_B, I), \tag{1}$$

where  $I = s, o$  denotes the identity of Player B, and  $\pi_A, \pi_B \in \mathbb{R}$  represent monetary payoffs.  $u_A \in \mathbb{R}$  is strictly monotonic in  $\pi_A$  and  $\pi_B$  and captures a taste for discrimination towards out-group members if an individual is better-off giving away money to an in-group ( $I = s$ ) rather than an out-group ( $I = o$ ) member (Chen and Li, 2009). Since we focus on how identity and information affect beliefs, we work with the simplifying (and perhaps strong) assumption that  $u_A$  does not depend on  $\mathbf{x}$ .

According to Equation 1, the utility of not trusting is equal to  $u_A(5, 5, I)$ . The expected utility of trusting is  $p(I, \mathbf{x})u_A(10, 10, I) + (1 - p(I, \mathbf{x}))u_A(0, 14, I)$ .<sup>9</sup> Defining  $\Delta(I) = u_A(10, 10, I) - u_A(0, 14, I)$  and  $M(I) = u_A(0, 14, I) - u_A(5, 5, I)$ ; and assuming that there are other random elements that affect the benefits of trust and not-trust,  $\varepsilon_T$  and  $\varepsilon_{NT}$  respectively, Player A trusts if and only if

$$p(I, \mathbf{x})\Delta(I) + M(I) \geq \varepsilon_{NT} - \varepsilon_T.$$

Denoting  $F$  the cumulative distribution function of  $\varepsilon_{NT} - \varepsilon_T$ , the probability that we observe trust is given by

$$\Pr\{A \text{ trusts} | I, \mathbf{x}\} = F(p(I, \mathbf{x})\Delta(I) + M(I)). \quad (2)$$

Equation 2 reveals three channels through which political identity shapes trust in this simple framework.<sup>10</sup> On top of the effect of identity on beliefs, identity also determines how sensitive trust is to beliefs through  $\Delta(I)$ , as well as the baseline level of trust through  $M(I)$ . A positive relation between trust and beliefs has been found in laboratory experiments when identities are not revealed (see e.g., Charness and Dufwenberg 2006; Buchan, Croson and Solnick, 2008). This is also our second hypothesis:

**Hypothesis 2.** *Player A's trust is positively associated with beliefs regardless of Player B's political identity.*

---

<sup>9</sup>This approach is consistent with Williamson (1993).

<sup>10</sup>Identity may affect trust through second-order beliefs as well, which are not included in the decision to trust in this model. The literature exploring second-order beliefs on trust mainly focuses on the receiver's decision (see, e.g., Charness and Dewfenberg, 2006). In this case, however, Butler et al, (2015) show that receiver's own beliefs about what is morally correct have more explanatory power than his second-order beliefs. We do not include second-order beliefs in our design because we focus on sender's behavior, and because we seek to assess the direct effect of political identity on first-order beliefs while keeping the experiment simple enough for a general U.S. subject pool.

The positive relation between trust behavior and trust beliefs in Hypothesis 2 means that Player A prefers Player B to cooperate regardless of the latter’s identity, i.e.,  $\Delta(s), \Delta(o) > 0$  (recall that  $\Delta(I) = u_A(10, 10, I) - u_A(0, 14, I)$ ). If Player A enjoys Player B’s cooperation (compared to defection) more when it comes from an in-group rather than an out-group Player B, then we should also expect that  $\Delta(s) > \Delta(o)$ . In other words, Player A’s trust decision is more sensitive to beliefs when she shares Player B’s political identity. This is our next hypothesis:

**Hypothesis 3.** *Player A’s decision to trust is more sensitive to her beliefs about Player B’s trustworthiness when both have the same political identity.*

Without further assumptions, this simple individual decision-making framework is silent about whether beliefs are statistically correct—an important question that motivates our design. We treat this question as an empirical one, which can be answered by comparing Player A’s beliefs to Player B’s actual trustworthiness. Drawing on social psychology research, however, we should at least expect more cooperation from Player B toward in-groups justified by the moral foundations that prescribe reciprocation and loyalty for Democrats and Republicans, respectively (see, e.g., Graham, Haidt and Nosek, 2009; Greene, 2014).

**Hypothesis 4.** *The frequency with which Players B reciprocate trust is higher when Players A share their political identity than when Players A do not share their political identity.*

A key goal of this paper, reflected in Hypothesis 1, is to test whether others’ political identity shapes own beliefs. But political identity only provides coarse information, if any, about the potential cooperation of Player B. Our final task is therefore to determine whether beliefs respond to objective information and whether this response depends on players’ political identities. Beliefs are largely characterized by self-deception and conformity to group values rather than by the incorporation of objective, sometimes disconfirming, information (see e.g., Bénabou and Tirole 2011; Friesen, Campbell and Kay 2014). In our setting this means that political polarization may motivate individuals to under-respond to news that disconfirms their beliefs (Taber and Lodge, 2006). For example, a Democrat Player A may fail to update her beliefs about a Republican Player B’s trustworthiness when she receives information that disconfirms her pessimistic priors. The same holds when Player A is Republican and Player B is Democrat. Whether individuals respond to information is arguably an important question because one would hope that positive and objective information leads

to efficiency gains in transactions that involve trust, especially in polarized societies. Our final hypotheses can be stated as follows:

**Hypothesis 5a.** *Democrat (Republican) Player A updates her beliefs about Democrat (Republican) Player B's trustworthiness in response to information about high frequency of cooperation by Democrats (Republicans) in previous experiments.*

**Hypothesis 5b.** *Democrat (Republican) Player A does not update her beliefs about Republican (Democrat) Player B's trustworthiness in response to information about high frequency of cooperation by Republicans (Democrats) in previous experiments.*

We test these hypotheses in the next two sections.

## Results

Consistent with existent survey data, Republicans and Democrats have different demographic characteristics in our sample. Table 2 shows the means of the most relevant ones in our study. Regarding demographics, we observe a greater proportion of individuals who are male, white, born between 1976-1985, conservative (and very conservative), or married with annual earnings between US\$ 75k and US\$ 150k among Republicans than among Democrats. We observe no significant differences across partisanship in the proportion of black individuals, those who have at least college education, those who are currently not working, those who earn between US\$ 150k and US\$ 250k a year, or those ones born before 1975. Nor do we observe significant differences in scores on the Cognitive Reflection Test across partisan identity. Pew Research Center (2015) presents demographic breakdowns of party identification by gender, race, marital status, education, employment, family income and cohort (among other categories) which we use to compare with our sample. Overall, the Pew sample demographics are consistent with ours: their sample also features more individuals who consider themselves or lean towards Republicans who are male, white or married with annual family earnings between US\$ 75k and US\$ 150k. One difference with respect to our sample, however, is that Pew documents that blacks are more often Democrats and those with at least college education are also more often Democrats. We do not find significant differences in those categories in our sample. Pew also documents slightly higher numbers of Republicans among older cohorts, whereas we do not observe significant differences. It is important to note that although Republicans and Democrats are different in many aspects,

they keep almost the same amount in the dictator game (US\$ 3.72 and US\$ 3.67 out of US\$ 5, respectively) and the difference is not statistically significant (t-test p-value = 0.67).<sup>11</sup>

Table 2: **Summary statistics.**

	Republican N=238	Democrat N=247			t-test
	mean	mean	diff.	se	p-value
Amount kept, dictator [0,5]	3.72	3.67	0.04	0.10	0.67
Female {0,1}	0.47	0.66	-0.19	0.04	0.00
White {0,1}	0.78	0.69	0.09	0.04	0.03
Black {0,1}	0.06	0.09	-0.03	0.02	0.16
Married {0,1}	0.62	0.43	0.19	0.04	0.00
At least college {0,1}	0.75	0.68	0.06	0.04	0.12
Not Working {0,1}	0.20	0.25	-0.05	0.04	0.16
Income [75k,150k] {0,1}	0.42	0.24	0.18	0.04	0.00
Income (150k,350k) {0,1}	0.12	0.15	-0.03	0.03	0.37
Income more 350k {0,1}	0.01	0.05	-0.04	0.02	0.02
Born in [1931-1950] {0,1}	0.03	0.05	-0.02	0.02	0.20
Born in [1951-1965] {0,1}	0.20	0.17	0.03	0.04	0.44
Born in [1966-1975] {0,1}	0.24	0.25	-0.01	0.04	0.85
Born in [1976-1985] {0,1}	0.33	0.25	0.08	0.04	0.05
Born in [1986-1995] {0,1}	0.20	0.28	-0.08	0.04	0.05
Average CRT {0,1,2,3}	1.53	1.48	0.05	0.11	0.66
Conservative {0,1}	0.33	0.03	0.30	0.03	0.00
Very Conservative {0,1}	0.11	0.01	0.10	0.02	0.00

The table shows the means for the most important demographics in the samples of Democrats and Republicans and a t-test of the difference in means between Democrats and Republicans for each characteristic. The interval [0, 5] in the first row, first column shows the range of that variable, and the set {0, 1} indicates that the corresponding variable takes the values of 0 or 1 in the remaining rows. Cognitive Reflection Test (CRT) scores can take values {0, 1, 2, 3}.

## Beliefs, trust and identity

Table 3 shows trust behavior and beliefs about partners' trustworthiness across the main treatments. Trust and beliefs are relatively high and similar across political identities of Player A. Democrats trust 56% of the time and believe Player B will reciprocate 62% of

<sup>11</sup>The average giving rate (26%) is similar to other experiments. For example, Bohnet and Frey (1999) document 25% average giving; and Andreoni and Miller (2002) 23%.

the time. The same figures for Republicans are 59% and 63%, respectively (see column “Overall” in Table 3). We fail to reject that trust behavior overall is equal across identity of Players A (chi-squared p-value = 0.5). Beliefs are also not statistically different across partisan identities of Player A (two-sided rank-sum test p-value = 0.89).<sup>12</sup>

Table 3: **Trust behavior and trust beliefs.**

Democrat’s partner identity is	Not revealed	Democrat	Republican	Overall
Fraction of Trust	0.57	0.63	0.4	0.56
# of Players A who Trust/Total	57/100	63/100	19/47	139/247
Mean beliefs about trustworthiness	0.62	0.67	0.52	0.62
s.d.	(0.22)	(0.21)	(0.28)	(0.24)
Republican’s partner identity is	Not revealed	Democrat	Republican	Overall
Trust	0.58	0.66	0.58	0.59
# of Players A who Trust/Total	55/95	29/44	57/99	141/238
Mean beliefs about trustworthiness	0.64	0.63	0.62	0.63
s.d.	(0.21)	(0.21)	(0.23)	(0.22)

The table shows the fraction of Democrat and Republican Players A who trust, as well as their mean beliefs about Player B’s trustworthiness. The columns “Not revealed”, “Democrat” and “Republican” refer to the treatments in which Player B (receiver) comes from each one of those categories. The last column, “Overall,” shows the trust rates and mean beliefs for each sub-sample of Democrat and Republican Player A (sender).

Table 3 summarizes Player A’s behavior according to the political identity of Player B. When focusing on the main pairwise comparisons, Democrat Players A trust Democrats more often than they do Republicans: 63% of the time, compared with 40% of the time (chi-squared p-value = 0.01). Regarding beliefs, a Democrat Player A believes a Democrat Player B is more trustworthy on average. The mean expected frequency of reciprocal behavior is 67% when Player B is a Democrat compared with 52% when Player B is Republican. We reject equality of distribution of trust beliefs for Democrat Players A across identity of Player B as beliefs are more optimistic when Player B is Democrat than when Player B is Republican (one-sided rank-sum test p-value < 0.01). Republican Players A, on the other

<sup>12</sup>Trust rates when matched to an anonymous are 57% when pooling Democrat and Republican Players A. This is consistent with previous student-subject experiments: Charness and Dufwenberg (2006) reports 56%; and Eckel and Wilson (2002) report 52.8% (Game 2).

hand, do trust a Democrat Player B more often (66% of the time) than a Republican Player B (58% of the time). This difference, however, and all the other differences from pairwise comparisons (including trust beliefs) are not statistically significant at conventional levels for Republican Players A. Checking for a significant difference in trust beliefs provides a first test of Hypothesis 1, which is borne out by the data only for Democrats.

We estimate an empirical model of beliefs to assess whether the effect of identity is robust to controlling for individual characteristics. For each type of Player A, we separately estimate variants of the empirical model

$$p_i = \gamma_0 + \gamma I_s^i + D_i + \xi_i, \quad (3)$$

where the subscript  $i$  indexes Players A and  $p_i$  represents the beliefs of  $i$  about Player B's trustworthiness.  $I_s^i$  is the treatment variable and denotes whether the political identity of Player B is the same ( $s$ ) as that of Player A (the baseline corresponds to different political identity of Player B).  $D_i$  takes on three specifications, each representing a different regression.

In the first specification,  $D_i = \Theta' X_i$ ;  $X_i$  is a vector of individual characteristics, which we assume serve as proxies for prior beliefs shaped by elements different from the political identity of Player B (such as Player A's income or education), and  $\Theta$  is a vector of parameters.<sup>13</sup>

In the second model,  $D_i$  in Equation 3 equals  $\theta X_i^k + \gamma_X (I_s^i \times X_i^k)$ , the sum of the  $k$ -th characteristic multiplied by the parameter  $\theta$ , and the interaction between the same-identity dummy and the  $k$ -th characteristic multiplied by the parameter  $\gamma_X$ . This specification allows us to test whether individual characteristics moderate the identity treatment effect by assessing the significance of  $\gamma_X$ . In addition, we consider  $X_i^k$  to be dichotomous (0 or 1) and we focus on education, income and cognitive reflection (CRT). We include in the online appendix interactions with female, age and amount kept in the dictator game, as well. For each characteristic  $k$  (education, income and CRT), we code  $X_i^k = 1$  if education is college or beyond, income is above \$75k a year or CRT score is greater or equal to 2, and  $X_i^k = 0$  otherwise, respectively. Under this codification,  $\gamma$  denotes the effect of (same) identity in the  $X_i^k = 0$  group and  $\gamma + \gamma_X$  denotes the effect of identity in the  $X_i^k = 1$  group. Assessing the significance of  $\gamma$  and  $\gamma + \gamma_X$  allows us to determine which sub-population, if any, drives

---

<sup>13</sup>We include the variables in Table 2 as  $X_i$ .

in-group biases.

The last specification of Equation 3 pools Players A across parties and considers  $D_i$  to be an indicator variable equal to one if Player B is Democrat. This last model aims at testing whether there exists a shared optimism across parties about Democrats' trustworthiness (see Table 3). Finally,  $\xi_i$  is the error term.

Table 4 columns (1)-(4) present the results for Democrat Players A, and columns (5)-(8) the results for Republican Players A. The first row, column (1), shows that the coefficient of the political identity indicator is positive and significant for Democrat Players A, when including demographic controls. Democrat Players A believe a Democrat Player B is on average 12% more likely to reciprocate trust than a Republican Player B. In contrast, for Republican Players A, first row column (5), political identity of Player B does not significantly affect beliefs, as the coefficient of the identity indicator is no different from zero.<sup>14</sup>

Let us turn to the analysis of moderators and look first into the effect of identity on Democrat Players A's beliefs (columns (2) to (4)). We observe that college education is the only variable that moderates the effect of identity.<sup>15</sup> In addition, the positive and significant estimate of  $\gamma$  (column (2) first row) and the non-significant estimate of  $\gamma + \gamma_X$  (column (2) bottom row, F test p-value = 0.21), suggest that most of the treatment effect (in-group bias) is driven by Democrat Players A without college education. Individuals with at least college education do not perceive fellow Democrats to be more trustworthy than Republicans. Income and cognitive reflection do not moderate the positive effect of same-identity on trust beliefs. When looking at either the estimate of  $\gamma$  or the estimate of  $\gamma + \gamma_X$  (first and bottom row in column (3) for income and bottom row column (4) for CRT), we see that at least one of them is positive and significant at conventional levels in each specification.

For Republican Players A, however, no individual characteristic seems to moderate (or exacerbate) the null effect of political identity on beliefs, see the estimates in each column from (6) to (8) in Table 4. Thus, no sub-group presents a significant treatment effect at conventional levels either.

When we pool across partisan identity of Player A, being matched to an in-group affects

---

<sup>14</sup>These results remain qualitatively unaltered when we restrict the sample to "strongly Democrats" (Liberal or Very Liberal Democrats:  $\hat{\gamma} = 0.17$  s.e. = 0.05) and to "strongly Republicans" (Conservative or Very Conservative Republicans:  $\hat{\gamma} = 0.03$  s.e. = 0.05).

<sup>15</sup>Interactions between the same-identity indicator and age, gender and altruism are reported in the online appendix, under Interactions. However, they do not significantly moderate or exacerbate the effect of identity on beliefs for Democrat and for Republican Players A.

Table 4: **Effect of political identity on trust beliefs.**

Dependent variable:	Beliefs								
Player A (Sender):	Democrat				Republican				Pooled
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Is (Identity: Same)	0.12 (0.04)	0.30 (0.07)	0.20 (0.07)	0.12 (0.06)	-0.02 (0.04)	0.03 (0.08)	0.06 (0.08)	-0.03 (0.06)	0.08 (0.03)
College		0.10 (0.07)				-0.04 (0.08)			
Is.x.College		-0.24 (0.09)				-0.06 (0.09)			
Income (>\$75k/y)			-0.08 (0.07)				-0.01 (0.08)		
Is.x.Income			-0.08 (0.08)				-0.09 (0.09)		
CRT (2 or 3)				-0.03 (0.07)				0.06 (0.07)	
Is.x.CRT				0.06 (0.08)				0.05 (0.08)	
Player B Dem.									0.07 (0.03)
_cons	0.45 (0.12)	0.45 (0.06)	0.57 (0.06)	0.54 (0.05)	0.49 (0.12)	0.66 (0.07)	0.63 (0.07)	0.59 (0.05)	0.54 (0.03)
N	147	147	147	147	143	143	143	143	290
$R^2$	0.340	0.150	0.163	0.088	0.162	0.030	0.031	0.049	0.042
Ho: $\gamma+\gamma_X=0$ pval	-	0.21	0.02	0.00	-	0.65	0.51	0.64	-

The table shows a linear model in which the dependent variable is Player A's trust beliefs. The treatment variable consists of an indicator variable on whether Player B (receiver) is from the same partisan identity,  $I_s$ . Each specification includes a covariate (individual characteristic) and the interaction between the identity indicator and the covariate. Columns (1)-(4) (resp., (5)-(8)) show the results for Democrat (resp., Republican) Player A. Columns (2)-(4) and (6)-(8) present demographic interactions. The bottom row shows the p-value of a F-statistic that tests whether the identity treatment effect for individuals who have a high level covariate is different from zero. Column (9) shows the results from including a Player B Democrat indicator, when we run the regression pooling across identities of Players A. Standard errors in parentheses.

beliefs even when we control for the identity of Player B being Democrat, as shown in Table 4 column (9). The coefficient on the Player B Democrat indicator is also significantly different from zero at 5% level, which corroborates the existence of a perception, common across parties, that Democrats are more trustworthy.

The results in Table 3, 4 and the previous discussion are consistent with Hypothesis 1 for Democrats. For Republicans, Hypothesis 1 does not hold. Democrats hold pessimistic beliefs about out-groups' trustworthiness and trust out-groups less often. In contrast, Republicans are not more pessimistic about out-groups and do not discriminate against out-groups more often either. This asymmetric discrimination challenges the common finding of unconditional animosity toward out-groups and highlights the importance of studying political identity in context. These results also expose beliefs as a key channel through which political identity can affect behavior.

We next ask whether beliefs are associated with trust. We estimate the decision to trust according to Equation 2 for each different pair of identities separately. Using the notation above, we estimate six specifications of

$$Trust_i = \mu_I + \delta_I p_i + \varepsilon_i. \quad (4)$$

The first four account for each type of Player A matched to each type of Player B (i.e.,  $I = s, o$ ). Each of the remaining two specifications pools Democrat and Republican Players A matched to either in-groups or out-groups. We opt for this parsimonious model in Equation 4 because it directly represents Equation 2. A test for Hypothesis 2 is whether the estimates of  $\delta_s$  and  $\delta_o$  are positive. Similarly, a test for Hypothesis 3 is whether the estimate of  $\delta_s$  is greater than the estimate of  $\delta_o$  for each type of Player A.

Table 5 presents the results. Columns (1) and (2) show the coefficients for Democrat Players A and columns (3) and (4) for Republican Players A. Column (1) shows the coefficient on beliefs and the constant when Player B is also Democrat. Beliefs correlate with trust positively and significantly. When Player B is Republican (column 2), beliefs are not significantly correlated with trust. A similar relation toward in-group members exists for Republican Players A. When matched to a Republican Player B (column 4), beliefs and trust correlate positively and significantly, but the point estimate is smaller than the estimate between Democrats. Trust and beliefs do not correlate when Player B is a Democrat (column

3).<sup>16</sup>

Table 5: **Relationship between beliefs and trust.**

	Dem. Player A (Sender)		Rep. Player A (Sender)		Pooled and CRT = 2 or 3	
Player B (Receiver):	Democrat	Republican	Democrat	Republican	Same	Other
Dependent variable:			Trust			
	(1)	(2)	(3)	(4)	(5)	(6)
Beliefs (p)	0.85 (0.22)	0.3 (0.26)	0.08 (0.35)	0.49 (0.21)	0.66 (0.21)	0.50 (0.28)
_cons	0.06 (0.15)	0.25 (0.15)	0.61 (0.23)	0.27 (0.14)	0.24 (0.15)	0.30 (0.18)
N	100	47	44	99	96	51
$R^2$	0.14	0.03	0.00	0.05	0.10	0.06

The dependent variable is whether Player A (sender) trusts Player B (receiver). The explanatory variables are Player A’s beliefs about Player B’s trustworthiness and the constant. All the results in this table come from the linear probability model. Standard errors in parentheses.

The perhaps unintuitive result that beliefs do not correlate with trust when Player A is matched with an out-group can be explained by taking into account cognitive reflection and a potential lack of power. Column (5) shows the results from pooling in-group participants (Democrats matched to Democrats and Republicans to Republicans) and considers participants who score high (2 or 3) in the CRT. The point estimate is positive and significant, and lies between the estimates in columns (1) and (4). Column (6) shows the results from pooling out-groups (Democrat Players A matched to a Republican Player B and Republican Players A matched to a Democrat Player B) and individuals with high CRT score (2 or 3). The coefficient on beliefs is 0.5 and marginally significant.<sup>17</sup> We interpret these results as lending support to Hypothesis 2, but the association between beliefs about the trustworthiness of others and actual trust is stronger when players’ political identities coincide.

<sup>16</sup>We also run the same specification over the sample of Democrat Players A who were not informed about the political identity of Player B. Beliefs have a positive, though marginally significant relationship with trust (point estimate 0.37, s.e. 0.22) in that case. Similarly, when we run the regression over Republican Players A who were not informed about the political identity of Player B, the point estimate is not significantly different from zero.

<sup>17</sup>If one is willing to take one step further and include observations from the belief treatment to increase power (and adding an indicator variable that controls for this treatment), then both coefficients, in-group and out-group matches, are positive and statistically significant at conventional levels (in-group matches sample  $N = 167$ :  $\hat{\delta}_s = 0.98$  s.e. = 0.17; out-group matches sample  $N = 109$ :  $\hat{\delta}_o = 0.73$  s.e. = 0.23).

The estimates in Table 5 also suggest that the correlation between Player A’s trust and beliefs is higher when Player B has the same political identity. In order to provide a statistical comparison of coefficients across regressions, we stack specifications (1), (2), (3) and (4) together in a seemingly unrelated regression model (which assumes the error terms  $\varepsilon_i$  are correlated across individuals and across treatments) and test for equality of the coefficients on beliefs. This exercise yields a significant difference between the coefficients on beliefs across types of Player B for Democrat Players A (one-sided standard normal test p-value = 0.03). This difference is not significant, however, when Player A is Republican (one-sided standard normal test p-value = 0.14). When we run a seemingly unrelated regression model on the specifications resulting in columns (5) and (6), however, there is no significant difference between the coefficients on beliefs about in-groups and out-groups trustworthiness (one-sided standard normal test p-value = 0.31). In sum, beliefs matter for trust, but the evidence that trust is more sensitive to trust beliefs in in-group matches is weak and mainly driven by Democrats, lending little support to Hypothesis 3.

If one is willing to take Equation 2 at face value, rejecting Hypothesis 3 can be interpreted as the *net utility* for Player A being similar if an in-group or an out-group Player B cooperates (i.e.,  $\Delta(s) = \Delta(o)$  in Equation 2). From Player A’s perspective, if the net utility of trusting an in-group Player B is similar to the net utility of trusting an out-group Player B, then part of the variation in trust due to Player B’s identity must come from its effect on beliefs, not on preferences. This is not to say that a preference for discrimination does not play any role, because preferences may affect trust through  $M(I)$  in Equation 2. Although our experiment is not designed to identify a preference channel through which identity can affect trust, one could do two things using our data to explore this question: compare the constant coefficient (a proxy for  $M$  in Equation 2) across identity of Player B in Table 5 or check whether there are differences in trust behavior for extremely pessimistic participants ( $p$  is close or equal to zero) across identity of Player B. The outcomes of both of these exercises, however, suggest that Player B’s identity does not play a significant role explaining trust behavior in our data, once beliefs are controlled for. The details are in the online appendix.

## Beliefs and actual trustworthiness

We now proceed to test Hypothesis 4. Table 6 shows that beliefs about Player B’s trustworthiness are not statistically correct—they are more pessimistic. Democrat Players A believe on average that 62% of Players B will reciprocate, but 80% end up doing so. However, not all are excessively pessimistic: 53 out of 247 Democrat Players A state that at least 80% of Players B will reciprocate. Similarly, Republican Players A believe on average that 63% of Players B will cooperate. Sixty out of 238 Republican Players A state that at least 80% of Players B will reciprocate.

Table 6: **Beliefs and trustworthiness.**

Democrat Player A			
Player B identity is	Democrat	Republican	Overall
Mean Player A’s beliefs about Player B’s trustworthiness	0.67	0.52	0.62
s.d.	(0.21)	(0.28)	(0.25)
Player B’s actual trustworthiness	0.85	0.93	0.8
# of participants / Total	85/100	41/44	389/485
Republican Player A			
Player B identity is	Democrat	Republican	Overall
Mean Player A’s beliefs about Player B’s trustworthiness	0.63	0.62	0.63
s.d.	(0.21)	(0.23)	(0.23)
Player B’s actual trustworthiness	0.81	0.87	0.8
# of participants / Total	38/47	86/99	389/485

The table shows average beliefs about Player B’s trustworthiness and actual rate of Player B’s reciprocation. The “Democrat” (“Republican”) column shows the beliefs Player A holds about a Democrat (Republican) Player B’s trustworthiness and the actual fraction of Players B who choose to reciprocate. The “Overall” column shows average beliefs for all Democrats and Republican Players A and the rate of reciprocation over the whole sample.

These differences are more pronounced when we separate them by the political identity of Player B. Republican Players B reciprocate trust to a Democrat Player A 93% of the time, which is notably higher than the mean belief a Democrat Player A holds about a Republican Player B: 52%. Only 8% (4 out of 47) of Democrat Players A are right in believing that

at least 90% of Republican Players B would reciprocate trust. Similarly, when Player B is Democrat and Player A is Republican, mean beliefs are 63%, but the actual reciprocation rate is 81%. Twenty-five percent (11 out of 44) of Republican Players A believe at least 80% of Democrat Players B would reciprocate trust.

In the previous section, we see that Player B's political identity has a significant effect on beliefs (see Table 4) only for Democrat Players A. In this case, the actual reciprocation rate by Democrats is 85% (85 out of 100) compared with 93% (41 out of 44) by Republican Players B. This difference, however, is not statistically significant (chi-squared p-value = 0.17) and if anything, it points in the other direction: Republicans are more trustworthy than Democrats, when matched with a Democrat Player A.

Republicans do not show a statistically different perception about Player B's trustworthiness across political identity (mean beliefs are 63% when matched with a Democrat Player B versus 62% when matched with a Republican Player B). Moreover, Republican Players B reciprocate more often than Democrat Players B when matched with a Republican Player A (81% by Democrats versus 87% by Republicans). Even though these two reciprocation frequencies are not statistically different (chi-squared p-value = 0.34), they suggest that Republicans are more trustworthy than Democrats when matched to a Republican Player A, just as they seem to be when matched to a Democrat Player A. Taken together, these results do not support Hypothesis 4.

Table 6 also reveals no significant differences in reciprocation toward in-group and out-group Players A when holding fixed the identity of Player B (chi-squared p-values are 0.53 and 0.27 when Player B is Democrat and Republican, respectively). In other words, there is no discrimination in the reciprocation decision. This raises the question: How can discrimination in the trust decision coexist with no discrimination in the reciprocation decision, at least for Democrats? In our simple game both decisions are similar in that they could depend on a preference for benefiting or hurting the partner. However, the trust decision differs from the reciprocation decision in one fundamental aspect. Trust depends on beliefs about behavior of others, while reciprocation does not. The monetary outcome Player B receives after deciding whether to reciprocate or not depends solely on the decision of Player B, which reflects his or her preferences only. Thus, the lack of discrimination in the reciprocation decision suggests that an identity-driven preference for discrimination is not playing a significant role. This observation further emphasizes the importance of beliefs, as opposed

to preferences, in trust decisions.

## **Manipulating beliefs**

We next explore whether beliefs can be manipulated. We ran four additional experimental surveys in which we revealed the frequency of reciprocation by Players B from the main treatments. We added one sentence with this information. See the online appendix section 2.1. No other change was introduced to the format or content of the survey relative to the survey in the main treatments. The four experimental surveys in this beliefs treatment were sent to a total of 200 participants: 100 Democrats and 100 Republicans. Out of the 100 Democrats, 50 received the survey that revealed the identity of the matching partner to be Democrat and the proportion of Democrat Players B who reciprocated trust in the main treatments to be 85%. The remaining 50 Democrats received the survey that stated the partner was a Republican and Republican Players B reciprocated 93% of the time in past surveys. Similarly, each of the 50 (out of 100) Republicans who received the experimental survey was matched to a Democrat and was informed Democrats reciprocated trust 81% of the time when they were matched with a Republican Player A. Each of the 50 remaining Republicans was matched to a Republican and was informed that Republicans reciprocated trust 87% of the time when matched with another Republican.

The response rate was 46/50 for a Democrat Player A matched with a Democrat Player B, 48/50 for a Democrat Player A matched with a Republican Player B, 47/50 for a Republican Player A matched with a Democrat Player B, and 50/50 for a Republican Player A matched with a Republican Player B.

## **Effect of information on trust beliefs**

Does revealing actual reciprocation change beliefs? Figure 2 shows the distribution of beliefs reported by Democrats in the main treatments (in which no information about actual reciprocation rates is revealed) and the beliefs treatment in which participants are informed about reciprocity. Panel 1 shows the distribution of beliefs reported by Democrat Players A facing a Democrat Player B. The vertical line represents the revealed reciprocation rate in the previously run main treatments. Many participants, 46% (21 out of 46), report that

Democrat Players B reciprocate trust at the same rate as that previously revealed (85%). In addition, trust beliefs in the beliefs treatment are statistically larger than the corresponding beliefs in the main treatments (one-sided Wilcoxon rank-sum test p-value = 0.01). Panel 2 shows the distribution of beliefs reported by Democrat Players A facing a Republican Player B. Although we marginally reject equality of distributions in favor of the alternative that beliefs are larger in the belief treatment (one-sided Wilcoxon rank-sum test p-value = 0.09), Democrat Players A seldom report trust beliefs about Republican Players B that coincide with the proportion revealed (93%). In other words, it seems Democrats become more optimistic about Republicans' cooperation rate, but their beliefs are still pessimistic.

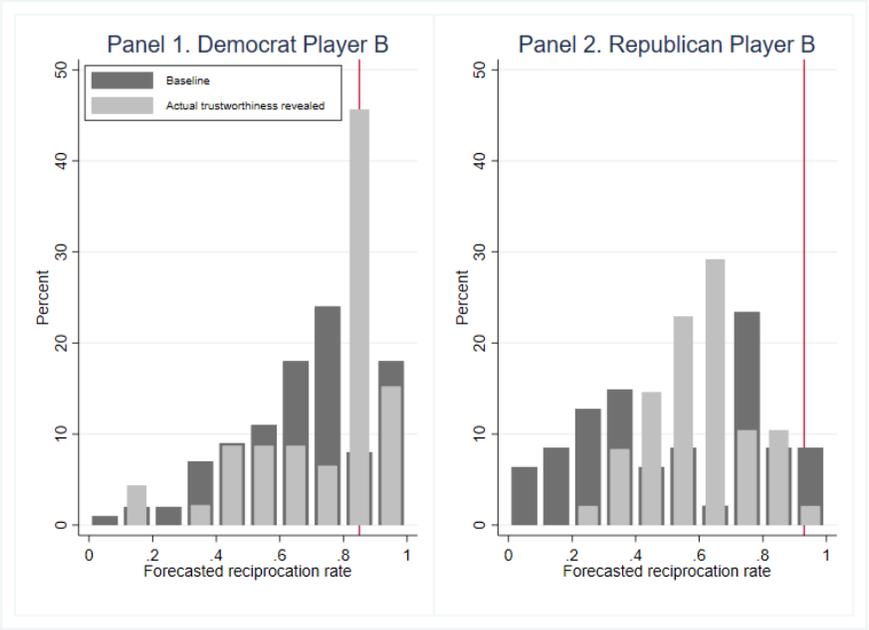


Figure 2: **Effect of information on Democrat Players A's beliefs.**

Figure 3 shows the distribution of beliefs reported by Republicans for the baseline condition in which no information about actual reciprocation is revealed and the treatment condition in which participants are informed about the actual rates. Panel 1 presents the distribution of beliefs reported by Republican Players A facing a Democrat Player B. The vertical line represents the revealed reciprocation rate in the corresponding main treatment. At conventional levels, we fail to reject equality against larger beliefs after revelation (one-sided Wilcoxon rank-sum test p-value = 0.93). Panel 2 shows the distribution of beliefs reported by a Republican Players A facing a Republican Player B. We also fail to reject the

null of equality of distributions (one-sided Wilcoxon rank-sum test  $p$ -value = 0.77).

The null effect of positive information on Republican Players A's trust beliefs is interesting and invites further exploration. The natural first step is to look for individual characteristics that moderate the effect of (positive) information on beliefs. We report the results from this exercise in the online appendix section 6, for both Democrat and Republican Players A. The data suggest that for Democrat Players A matched to Democrat Players B, college education, income and age moderate the effect of information on beliefs, and higher levels of these variables are associated with larger treatment effects. For Democrat Players A matched to Republican Players B, income and being female moderate the effect of information on beliefs. Both high income and female Democrats seem to drive the positive effect of information in Figure 2 panel 2. For Republican Players A, these covariates do not seem to moderate (or exacerbate) the null effect of information illustrated in Figure 3. Nonetheless, the results presented in Table 5 in the online appendix, do suggest that low education, low income or selfish Republican Players A seem to react slightly negatively to positive information about Republican Players B's trustworthiness. These results indicate that people process information based on an interplay between the type of information (e.g., whether it is about in-groups or out-groups) and their backgrounds (e.g., individual characteristics). This is broadly consistent with previous work on motivated reasoning (e.g., Taber and Lodge, 2006; Mason, 2014; Lelkes and Westwood, 2016) and cultural transmission (Guiso et al. 2008; Butler et al., 2015; Ortoleva and Snowberg, 2015). However, our results stress the importance of taking into account the exact situation, a trust game in our case, type of information and individual heterogeneity when drawing conclusions about beliefs and, perhaps more generally, political stereotypes.

In sum, revealing high previous cooperation rates has a positive effect on the beliefs of Democrat Players A about Democrat Players B's trustworthiness, which is consistent with Hypothesis 5a. We also find some evidence that Democrat Players A slightly update beliefs about Republican Players B, which is not consistent with Hypothesis 5b. When Player A is Republican, on the other hand, revealing high cooperation rates has no positive effect on beliefs about Republican Players B trustworthiness, which is not consistent with Hypothesis 5a. The null reaction to information when Player B is Democrat, however, is consistent with Hypothesis 5b. One conjecture from this exercise is that manipulating incorrect beliefs by revealing actual behavior is not straightforward. Beliefs about reciprocity across partisanship

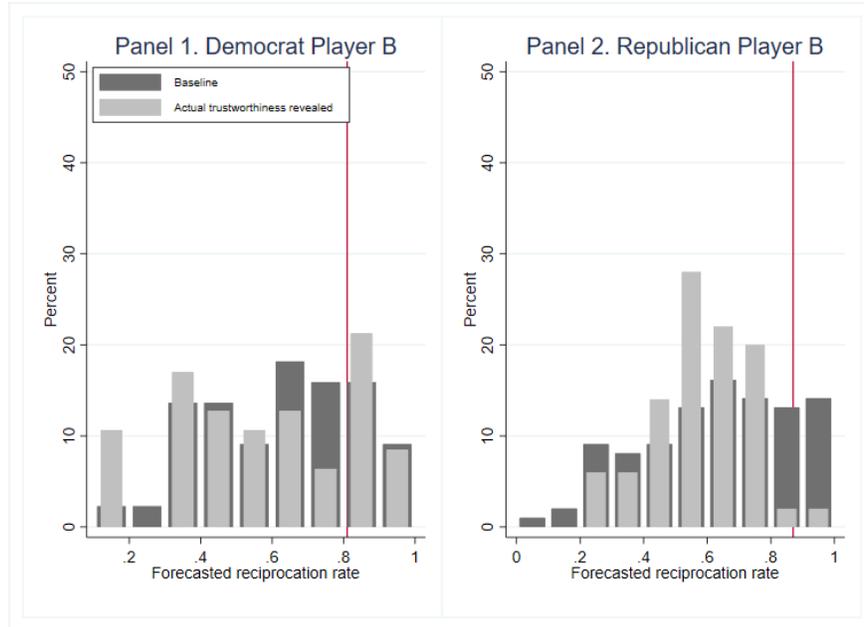


Figure 3: **Effect of information on Republican Players A's beliefs.**

may not be malleable.<sup>18</sup>

## Conclusion

Political polarization is an important phenomenon. Scholars and commentators usually conclude that it may be hindering trust not only in political spheres but also in the daily lives of U.S. citizens. A key component of trust is the expectation that others will deliver on their promises or duties. Research across social sciences has shown that a taste for discrimination, dislike or even hatred usually exist across political identities. We study whether people also have different beliefs about how trustworthy are people of different partisan identities.

To explore what could drive different beliefs, we study the relationship between political identity and trust among a sample of the U.S. population. In particular, we ask paired subjects to play a simple trust game with either the same or competing political identity partner. We find that there are partisan identity-based differences on trust and, importantly,

<sup>18</sup>Although the focus of this paper is on trust beliefs, we explore the effect of information on trust behavior in the online appendix. The illustrative results suggest that revealing high reciprocation in previous treatments has an overall positive effect on trust only for Democrat Players A (see Table 6 in the online appendix). In addition, such information is positively correlated with trust for optimistic Players A, and negatively correlated for pessimistic Players A, suggesting motivated reasoning.

on trust beliefs. Both, Republicans and Democrats, believe that Democrats are more trustworthy, but Democrats make a more notorious distinction. Moreover, we find that beliefs do matter for trust, especially for individuals who show high levels of reflective thinking. We also find that although there is no difference in trustworthiness as a function of political identity, individuals do hold beliefs that are much more pessimistic than actual behavior. This holds regardless of the partisan identity of both the trustee and trustor.

It is encouraging that some of the distrust is driven by incorrect beliefs, and not just by a taste for discrimination. However, these incorrect beliefs are apparently not easy to change. Although we note that Democrats update beliefs based on objective information about both Democrats and Republicans' trustworthiness, revealing such information does not change the beliefs of Republicans, and this result is robust to controlling for many demographic factors. Further analyzing the robustness of the differential response to information by Democrats and Republicans is a rather interesting avenue for future research.

## References

- [1] Abramowitz, Alan. 2010. *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy*. New Haven, CT: Yale University Press.
- [2] Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions" *Quarterly Journal of Economics* 114(4): 1243-1284.
- [3] Alesina, Alberto, and Eliana La Ferrara. 2002. "Who Trusts Others?." *Journal of Public Economics* 85: 207-34.
- [4] Anderson, Lisa R., Jennifer M. Mellor, and Jeffrey Milyo. 2005. "Do Liberals Play Nice? The Effects of Party and Political Ideology in Public Goods and Trust games." Vol. 13, pp. 107-131. J. Morgan Ed. Elsevier. Emerald Group Publishing Limited.
- [5] Andreoni, James, and John Miller. 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 70 (2), 737-753.
- [6] Arrow, Kenneth. 1972. "Some Mathematical Models of Race Discrimination in the Labor Market", in (A.H. Pascal, ed.), *Racial Discrimination in Economic Life*, pp. 187-204, Lexington, MA: D.C. Heath.

- [7] Arrow, Kenneth. 1974. *The Limits of Organization*. NY: Norton.
- [8] Becker, Gary. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press, 1957.
- [9] Becker, Gary. 1993. "The Evidence against Banks Doesn't Prove Bias," *Business Week*, April 19.
- [10] Bénabou, Roland, and Jean Tirole. 2011. "Identity, Morals, and Taboos: Beliefs as Assets." *The Quarterly Journal of Economics*, 126(2), 805-855.
- [11] Berg, Joyce, Dickhaut, John, and McCabe, Kevin. 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior*, vol. 10, no 1, p. 122-142.
- [12] Blanco, Mariana, Dirk Engelmann, Alexander K. Koch, and Hans-Theo Normann. 2014. "Preferences and Beliefs in a Sequential Social Silemma: A Within-subjects Analysis." *Games and Economic Behavior*, 87, 122-135.
- [13] Bohnet, Iris, and Richard Zeckhauser. 2004. "Trust, Risk and Betrayal." *Journal of Economic Behavior & Organization*, 55(4), 467-484.
- [14] Bohnet, Iris and Bruno S. Frey. 1999. "Social Distance and Other-regarding Behavior in Dictator Games: Comment." *American Economic Review*, 89(1), pp.335-339.
- [15] Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *The Quarterly Journal of Economics*, 131(4), 1753-1794.
- [16] Brandts, Jordi, and Gary Charness. 2011. "The Strategy Versus the Direct-response Method: A First Survey of Experimental Comparisons." *Experimental Economics*, 14(3), 375-398.
- [17] Buchan, Nancy R., Rachel Croson, and Sara Solnick. 2008. "Trust and Gender: An Examination of Behavior and Beliefs in the Investment Game." *Journal of Economic Behavior & Organization*, 68(3), 466-476.
- [18] Bullock, John G., Alan S. Gerber, Seth J. Hill and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs about Politics," *Quarterly Journal of Political Science*, vol 10(4), pages 519-578.

- [19] Bursztyn, L., Georgy Egorov, and Stefano Fiorin. 2017. “From Extreme to Mainstream: How Social norms Unravel.” (No. w23415). National Bureau of Economic Research.
- [20] Butler, Jeffrey V., Paola Giuliano, and Luigi Guiso. 2015. “Trust, Values and False Consensus,” *International Economic Review*, August, 56 (3), 889-915.
- [21] Campbell, Angus, Philip Converse, Warren Miller, and Donald Stokes. 1960. *The American Voter*. University of Chicago Press.
- [22] Carey, John M., Nyhan, Brendan, Valentino, Benjamin, and Liu, Mingnan. 2016. “An Inflated View of the Facts? How Preferences and Predispositions Shape Conspiracy Beliefs about the Deflategate Scandal.” *Research and Politics*, 3(3).
- [23] Carlin, Ryan E., and Gregory J. Love. 2013. “The Politics of Interpersonal Trust and Reciprocity: An Experimental Approach.” *Political Behavior*, 35(1), 43-63.
- [24] Charness, Gary, and Martin Dufwenberg. 2006. “Promises and Partnership.” *Econometrica*, 74(6), 1579-1601.
- [25] Charness, Gary and Matthew Rabin. 2002. “Understanding Social Preferences with Simple Tests.” *Quarterly Journal of Economics*, 817-869.
- [26] Chen, Y., and Sherry X. Li. 2009. “Group Identity and Social Preferences.” *The American Economic Review*, 99(1), 431-457.
- [27] Eckel, Catherine C., and Rick K. Wilson. 2004. “Is Trust a Risky Decision?” *Journal of Economic Behavior & Organization*, 55(4), 447-465.
- [28] Falk, Armin, and Christian Zehnder. 2013. “A City-wide Experiment on Trust Discrimination.” *Journal of Public Economics*, 100, 15-27.
- [29] Fershtman, Chaim, and Uri Gneezy. 2001. “Discrimination in a Segmented Society: An Experimental Approach.” *Quarterly Journal of Economics*, 351-377.
- [30] Fiorina, Morris P., and Samuel J. Abrams. 2008. “Political Polarization in the American Public.” *Annual Review of Political Science* 11: 563-88.

- [31] Frederick, Shane. 2005. "Cognitive Reflection and Decision Making." *Journal of Economic Perspectives*, 25-42.
- [32] Friesen, Justin P., Troy H. Campbell, and Aaron C. Kay. 2014. "The Psychological Advantage of Unfalsifiability: The Appeal of Untestable Religious and Political Ideologies." *Journal of Personality and Social Psychology*, Nov 17.
- [33] Garbarino, Ellen, and Robert Slonim. 2009. "The Robustness of Trust and Reciprocity Across a Heterogeneous U.S. Population." *Journal of Economic Behavior & Organization*, 69(3), 226-240.
- [34] Gerber, Alan, and Gregory Huber. 2009. "Partisanship and Economic Behavior: Do Partisan Differences in Economic Fore-casts Predict Real Economic Behavior?" *American Political Science Review* 103(3): 407–26.
- [35] Gerber, Alan, and Gregory Huber. 2010. "Partisanship, Political Control, and Economic Assessments." *American Journal of Political Science* 54, no. 1 (2010): 153-173.
- [36] Gerber, Alan, Gregory Huber, and Ebonya Washington. 2010. "Party Affiliation, Partisanship, and Political Beliefs: A Field Experiment." *American Political Science Review* 104(4):720–44.
- [37] Gift, Karen, and Thomas Gift. 2015. "Does Politics Influence Hiring? Evidence from a Randomized Experiment." *Political Behavior* 37, no. 3: 653-675.
- [38] Goren, Paul, Christopher M. Federico, and Miki Caul Kittilson. 2009. "Source Cues, Partisan Identities, and Political Value Expression." *American Journal of Political Science*, 53(4), 805-820.
- [39] Graham, Jesse, Jonathan Haidt, and Brian A. Nosek. 2009. "Liberals and Conservatives Rely on Different Sets of Moral Foundations." *Journal of Personality and Social Psychology*, 96(5), 1029.
- [40] Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. "Partisan Hearts and Minds: Political Parties and the Social Identities of Voters" Yale University Press.

- [41] Greene, Joshua. 2014. *Moral tribes: Emotion, Reason, and the Gap Between Us and Them*. Penguin.
- [42] Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2008. "Social Capital as Good Culture," *Journal of the European Economic Association*, 6(2-3), 295-320.
- [43] Hetherington, Marc J. 1998. "The Political Relevance of Political Trust." *American political science review* 92, no. 4: 791-808.
- [44] Hetherington, Marc J., and Jason A. Husser. 2012. "How Trust Matters: The Changing Political Relevance of Political Trust." *American Journal of Political Science* 56, no. 2: 312-325.
- [45] Huddy, Leonie, Lilliana Mason, and Lene Aaroe. 2015. "Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity." *American Political Science Review* 109(1): 1-17.
- [46] Inglehart, Ronald. 1990. *Culture Shift in Advanced Industrial Society*. Princeton University Press.
- [47] Inglehart, Ronald. 1999. "Trust, Well-being and Democracy." *Democracy and Trust*, 88.
- [48] Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. "Affect, not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly*, 76(3), 405-431.
- [49] Iyengar, Shanto, and Sean J. Westwood. 2015. "Fear and Loathing Across Party Lines: New Evidence on Group Polarization." *American Journal of Political Science*, 690-707.
- [50] Knack, Stephen, and Philip Keefer. 1997. "Does Social Capital Have an Economic Pay-off? A Cross-country Investigation." *The Quarterly Journal of Economics*, 1251-1288.
- [51] Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108:480-4.
- [52] Lelkes, Yphtach, and Sean J. Westwood. 2017. "The Limits of Partisan Prejudice." *Journal of Politics* 79(2):485-501.

- [53] List, John. 2004. "The Nature and Extent of Discrimination in the Marketplace: Evidence From the Field", *Quarterly Journal of Economics*, vol. 119(1), pp. 49–89.
- [54] Mason, Lilliana. 2014. "I Disrespectfully Agree": The Differential Effects of Partisan Sorting on Social and Issue Polarization. *American Journal of Political Science*.
- [55] McGrath, Mary C. 2017. "Economic Behavior and the Partisan Perceptual Screen," *Quarterly Journal of Political Science*: Vol. 11: No. 4, pp 363-383. <http://dx.doi.org/10.1561/100.00015100>
- [56] McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: MIT Press.
- [57] McConnell, Christopher, Margalit, Yotam, Malhotra, Neil and Levendusky, Matthew. 2017. "The Economic Consequences of Partisanship in a Polarized Era." *American Journal of Political Science*. doi:10.1111/ajps.12330
- [58] Nyhan, Brendan, and Jason Reifler. 2010. "When Corrections Fail: The Persistence of Political Misperceptions." *Political Behavior*, 32(2), 303-330.
- [59] Oliver, J. Eric, and Thomas J. Wood. 2014. "Conspiracy Theories and the Paranoid Style (s) of mass opinion." *American Journal of Political Science* 58, no. 4: 952-966.
- [60] Ortoleva, Pietro, and Erik Snowberg. 2015. "Overconfidence in Political Behavior." *The American Economic Review*, 105(2), 504-535.
- [61] Perez-Truglia, R., and Cruces, G. 2017. "Partisan Interactions: Evidence from a Field Experiment in the United States." *Journal of Political Economy*, 125(4), 1208-1243.
- [62] Pew Research Center. 2015. A Deep Dive Into Party Affiliation, <http://www.people-press.org/files/2015/04/4-7-2015-Party-ID-release.pdf>
- [63] Phelps, Edmund 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review*, 62(4), 659-661.
- [64] Putnam, Robert D., Robert Leonardi, and Raffaella Y. Nanetti. 1994. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University press.

- [65] Rousseau, Denise M., Sim B. Sitkin, Ronald S. Burt, and Colin Camerer. 1998. "Not So Different After All: A Cross-discipline View of Trust." *Academy of Management Review*, 23(3), 393-404.
- [66] Taber, Charles S., and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science*, 50(3), 755-769.
- [67] Williamson, Oliver. 1993. "Calculativeness, Trust, and Economic Organization." *Journal of Law and Economics*, 453-486.
- [68] Zak, Paul J., and Stephen Knack. 2001. "Trust and Growth." *The Economic Journal*, 111(470), 295-321.