# Political Identity and Trust

Pablo Hernandez
Dylan Minor

**Working Paper 16-012**

# Political Identity and Trust

Pablo Hernandez
New York University AD

Dylan Minor
Harvard Business School

**Working Paper 16-012**

# Political Identity and Trust[*]

Pablo Hernandez

New York University AD

Dylan Minor

Harvard University

This version: July 2015

### Abstract

We explore how political identity affects trust. Using an incentivized experimental survey conducted on a representative sample of the U.S. population, we vary information about partners' partisan identity to elicit trust behavior, beliefs about trustworthiness, and actual reciprocation. By eliciting beliefs, we are able to assess whether differences in trust rates are due to stereotyping or a "taste for discrimination." By measuring actual trustworthiness, we are able to determine whether beliefs are statistically correct. We find that trust is pervasive and depends on the partisan identity of the trustee. Differential trust rates are explained by incorrect stereotypes about the other's lack of trustworthiness rather than by a "taste for discrimination." Given the importance of beliefs, we run additional treatments in which we disclose previous reciprocation rates before participants decide whether to trust. We find that beliefs are slightly more optimistic compared with the previous treatments, suggesting that incorrect stereotypes are hard to change.

**Keywords:** *Trust, Beliefs, Social Preferences, Political Ideology*

1

# 1 Introduction

Political polarization of the American public has increased, and partisan antagonism is "deeper and more extensive than at any point in the last two decades" (Pew Research Center, 2014). The consequences of this cross-partisan antipathy manifest themselves in a myriad of ways, both in politics and in everyday life. Evidence suggests that political polarization may be hindering cross-partisan trust, creating political and economic gridlocks (Carlin and Love 2013; Abramowitz 2007, 2010; Abramovitz and Saunders, 2005; Brewer and Pierce 2005). Political polarization based on partisan identity stems from the current strong emotional connection between the individual and the party (Mason 2014). Partisan bias and mistrust across partisanship have been explained therefore by sentiments of dislike, anger, and even loathe towards political opponents (Iyengar et al. 2012; Haidt and Adams 2014; Mason 2014; Iyengar and Westwood 2014).

In this paper, we explore the effect of partisan identity in trust behavior. We focus on trust, as it is fundamental for economic organization (see, e.g., Arrow 1974; Knack and Keefer 1997).[1] We are interested in analyzing the mechanism underlying trust. Ferschtman and Gneezy (2001) argue that there are two forces driving it: sentiments that lead to discrimination based on identity ("taste for discrimination") or stereotypes about others' behavior. 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 individual preferences and may reflect 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 about others' behavior rather than preferences (Phelps 1972). Such stereotypes, however, may or may not be statistically correct.[2] The main goal of this paper

---

[1]An interdisciplinary definition of trust comes from Rousseau et al. (1998):

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

[2]Williamson (1993) provides a similar view of trust. In his view there are two dimensions to trust: calculative and noncalculative. The former refers to trusting decisions based on calculations

is to determine whether monetary trust is driven by preferences or by stereotypes.

We conducted a highly incentivized experimental survey in which we varied the political identity (in terms of partisan identity) and measured a proxy for trust behavior, beliefs about trustworthiness, and actual trustworthiness across a sample of the general U.S. population. The following are the main questions we answer in this paper. 1) Do trust rates vary with political identity of a partner? 2) If so, do these differences in trust depend on beliefs about a partner's trustworthiness or on (social) preferences favoring in-group members relative to out-of-group members? 3) Do beliefs respond to the partisan identity of the partners in the interaction? 4) Are beliefs statistically correct?

In order to address these questions, we base our analysis on a model of identity and social preferences introduced by Chen and Li (2009), which builds upon Charness and Rabin (2002) and incorporates Akerlof and Kranton's (2000) "prescribed behavior" according to identity.

In our experiment, we run a simplified version of the standard trust game in Berg, Dickhaut, and McCabe (1995) similar to the one used in Charness and Dufwenberg (2006). This simplified version is a two-player game in which Player A (the sender) chooses ($5,$5) for himself and the other player, respectively, or defers the decision between ($10,$10) and ($0,$14) to Player B (the receiver). In this game, efficiency would require Player A to defer the decision to Player B. The subgame perfect equilibrium, however, implies Player B choosing ($0,$14). Thus, the potential gains from cooperation are not exploited. This game is labeled "trust game" because Player A choosing to defer the decision to Player B is an indication of her trust in Player B. We used this simplified version because it allowed us to elicit beliefs about trustworthiness directly. In the main treatments, we manipulated the information about the partisan identity of Player A and Player B, and we did not provide any other information regarding the identity of the matched partners. (We also run a baseline treatment in which partisan identity is not revealed.) The outcomes of

---

of expected monetary costs and benefits, while the latter refers to decisions based on sentiments and affection. Williamson's calculative trust encompasses stereotypes as beliefs that are fundamental in calculating the expected benefits and costs of trust. His noncalculative trust, on the other hand, refers to nonstandard other-regarding concerns.

Player A's and player B's decisions were paid in full. Hence, to the best of our knowledge, this was the first highly incentivized experimental survey administered to a sample of the general U.S. population incorporating partisan identity.

We find that the Democrat's and Republican's trust rates are around 60%. In answering question 1, we do find that trust rates depend on the partisan identity of Player B. Democrats and Republicans trust other Democrats more often, on average. However, only Democrat Player A types have such different trust levels as to be statistically significant. The answers to questions 2 and 3 are the following: This difference in trust rates is explained by the optimistic beliefs about trustworthiness that Democrats hold in favor of Democrats. Other-regarding preferences, which Fershtman and Gneezy (2001) suggest represent a taste for discrimination, do not seem to be related to trust rates across partisan identity, at least for our game.

Are these beliefs about a partner's trustworthiness statistically correct? Social psychology research (see, e.g., Chambers and Melnyk, 2006) shows that individuals of different partisan identity hold perceptions of large disagreements with opponents in core values such as abortion, while in reality opinions and actions are more similar than perceived (Mason 2014). We find that, overall, our evidence is consistent with this observation. Regardless of partisan identity, participants engaged in higher rates of reciprocation (i.e., Player B chooses ($10,$10) almost 80% of the time if entrusted with making the decision) than expected (mean of roughly 60%). In addition, Republicans (who are thought to be less trustworthy by both Democrats and Republicans) reciprocate slightly more often than Democrats. Hence, we answer question 4 in the negative: beliefs are not statistically correct.

Given that beliefs drive trust and the fact they were overall more pessimistic than actual reciprocation rates, we ran secondary treatments ("belief treatments") in which we revealed reciprocation rates from the main treatments, before participants decided to trust. These secondary treatments featured the exact same game as in the main treatments where partisan identity was manipulated, except now Player A also received information about the reciprocation rate corresponding to previous Players B. For example, if Player A was a Democrat and Player B a Republican in one of the belief treatments, we disclosed to Player A the reciprocation rate of Republican

Players B in the corresponding main treatment. We found that, regardless of the identity of Player B, revealing actual reciprocation rates made Democrats slightly more optimistic about Player B's trustworthiness, but it did not have a significant effect on Republicans' beliefs. Trust behavior, however, remained largely unchanged.

Our paper relates to the large literature on the drivers of trust in general and trust across partisan identity in particular. While providing a thorough account of this literature goes beyond the scope of this paper, it can be useful to sketch a brief note, primarily focused on the drivers of trust. Early experimental studies have provided evidence that a non-negligible proportion of senders do actually trust, and receivers usually reciprocate (Camerer and Weigelt 1988; Berg et al. 1995; Bolle 1998; Dufwenberg and Gneezy 2000; Ortmann, Fitzgerald, and Boeing 2000). These robust findings have spurred a series of papers analyzing the mechanisms underlying trust. Expectations about return and other-regarding preferences have been found to explain trust behavior to a large extent (Cox 2004; Ashraf, Bohnet, and Piankov 2006; Fehr 2009; Sapienza, Toldra-Simats, and Zingales 2013), and they have been shown to depend on the identity of the matching partners in the trust interaction.[3] Chen and Li (2009) find that individuals favor induced (according to taste and randomly) in-group members in trust games. Their evidence is consistent with individuals being more altruistic towards an in-group match. When it comes to identity based on pre-determined characteristics, Garbarino and Slonim (2009) show that expectations about partners' trustworthiness drives trust in experiments focused on the effect of gender and age on trust behavior. Similarly, Falk, and Zehnder (2013) explore trust decisions conditional on the district of the receiver. They also find that in-group trust is in part driven by positive beliefs about in-group trustworthiness.

---

[3]Other important studies have found that elements of psychological game theory (e.g., Geanakoplos, Pearce, and Stacchetti 1989 and Rabin 1993) also explain trust behavior. Charness and Duwfenberg (2006), for example, find that preferences over beliefs explain trust and trustworthy behavior. In the same vein, Bohnet and Zeckhauser (2004) introduce the concept of betrayal aversion to explain why individuals are less optimistic about return rates when matched to a human than when matched to a person. Under the assumption that some receivers reciprocate trust, Eckel and Wilson (2004) study whether there is a relationship between risk aversion and trust. The authors conclude there is no relationship unless additional control variables are included; but even in this case, the relationship is weak, at best. A similar conclusion is reached by Houser, Schunk, and Winter (2010).

They also find evidence that senders differentiate their investments depending on the receiver's district (favoring senders living in high-income districts). Along the same lines, but instead varying ethnic affiliation to Israeli Jewish society, Ferschtman and Gneezy (2001) find systematic mistrust towards men of Eastern origin to be driven by mistaken stereotypes and not by a "taste for discrimination."

To our knowledge, there are two papers that experimentally link trust and partisan identity in the U.S. Both papers use a design similar to Berg et al. (1995) and employ undergraduate students as participants. Carlin and Love (2013) find higher levels of trust among individuals from the same partisan identity. Anderson, Mellor, and Milyo (2005) find that self-described liberals trust more often than self-described conservatives. These papers focus on trust rates and on reciprocation rates. They also use small stakes in their experiments. While small stakes seem suitable to motivate the student subjects in their studies, they may induce self-selection in surveys directed towards the general U.S. public.

Our contribution to this literature is threefold. First, we explore the mechanism that drives trust (beliefs over other-regarding concerns) according to political identity. Second, these studies use college students while our population consists of individuals living in the U.S. ranging from 18 to 82 years old. Third, the stakes in the present experiment are high considering the duration of the experiment and the incentives we use (i.e., average payoffs were equivalent to $400 per hour).

## 2 Experimental Design

The main experiment ("main treatments") was designed to examine how trust behavior changes when we manipulate partners' partisan identity. We describe the main treatments in this section. The secondary experiment ("belief treatments") was run after the main experiment and it is described in section 6. Both of them used the Kellogg School of Management E-lab system, which maintains a pool of 7,045 participants from across the United States. E-lab staff prescreens individuals in this subject pool through a survey instrument from which partisan identity and other demographic information are collected. We used this information regarding

6

partisan identity to determine who received the experimental survey (the details are provided below). Subjects in this pool are then provided an opportunity to periodically participate in research surveys sponsored by faculty.

To maximize the response rate, our experimental survey was highly incentivized based on individual decisions. (Response rate: 97%, 485 out of 500 individuals.)

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; the only difference was that they were told that their decision on each question will be matched to the decision of a matched partner to compute payoffs. The exact procedure was as follows.

Each potential participant received an invitation to participate in the experimental survey.[4] 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, where subjects were told to allocate $5 anonymously between themselves and another participant.

The second question corresponded to the sender role in a trust game similar to that in Charness and Dufwenberg (2006) in Figure 1. Participants had a choice of trusting or not trusting a trustee (Player B), who would then make a final allocation decision. If the sender (Player A) 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 first intervention took place in this second question.[5] We varied the identity of Player B by letting Player A know that the otherwise-anonymous Player

---

[4]The survey was administered via Qualtrics. A copy of the survey instrument can be found in the appendix.

[5]We did not mention political affiliation to subjects until after the dictator game decision so as to obtain a measure of the Democrats' and Republicans' unconditional altruism and to avoid cueing on the subsequent trusting decision.
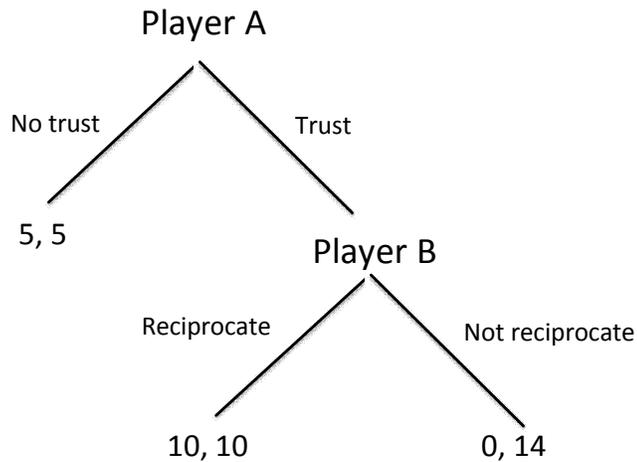
Figure 1: Game tree.

B identified him or herself as a Democrat or a Republican.[6] We also ran a baseline treatment where there was no such mention of the political identity of the subject's partner.

The third question asked the participant to make an allocation choice if entrusted with that decision by Player A. Following Charness and Dufwenberg (2006), the participant, now in the role of Player B, had to decide whether to allocate $10 to each player or to behave opportunistically and take $14 for him or herself and provide $0 for Player A. In this question, we also varied the identity of Player A to be Democrat or Republican.[7] We also ran a baseline treatment where Player A's political identity was not revealed.

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

---

[6]Question Q2 in the Appendix "Survey instrument, treatment conditions: Revealing a partner's political identity."

[7]Question Q3 in the Appendix "Survey instrument, treatment conditions: Revealing a partner's political identity."

predicted the sample's actual trustworthiness rate decile of Player B (i.e., those who would choose the ($10,$10) option when given the role of Player B). As in the previous two questions, we varied the information regarding the identity of those in the role of Player B by whether they identified themselves as Democrat or Republican in the prescreen survey.[8] 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 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. In this sense, 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.

It is important to note that the outcomes of Player A's and Player B's decisions were paid in full. Hence, to the best of our knowledge, this is the first widely administered and highly incentivized experimental survey incorporating partisan identity.

The next four nonincentivized 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 prescreen survey.[9]

---

[8]Question Q4 in the Appendix "Survey instrument, treatment conditions: Revealing a partner's political identitya"

[9]We use the information in the last four questions to control for intelligence and check for consistency with the E-lab prescreen survey. We find that subjects' answers to our questions are consistent with those given the E-lab, sometimes many months earlier. In particular, 221 out of 247 Democrats declared they were Democrat in the partisan question in our survey. Of the 26 remaining, 14 declared themselves Republican, 9 Independent, and 3 Other. In the case of Republicans, 182 out of 238 declared they were Republican in our survey. Of the 56 remaining, 39 declared themselves Democrat, 16 Independent and 1 Other. We find evidence, however, that declaring a different partisan identity may be driven by previous behavior and the partisan identity

9

| | Participant's ideology | |
|---|---|---|
| | Democrat (D) | Republican (R) |
| **Partner's ideology:** | | |
| Not revealed (NR) | **D-NR (N=100)** | **R-NR (N=95)** |
| Democrat (D) | **D-D (N=100)** | **R-D (N=44)** |
| Republican (R) | **D-R (N=47)** | **R-R (N=99)** |

Table 1: Treatment conditions. Total number of participants: 485.

We used the partisan identity information from E-lab's prescreen survey to identify participants. E-lab sent the experimental survey to 250 Democrats and 250 Republicans in total. The baseline treatment survey (no partner's 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 remaining 50 Democrats. The response rate was 100/100, 100/100, and 47/50, respectively. An analogous procedure was conducted with the 250 identified Republicans. The baseline was sent to 100, the type of survey identifying a partner as a Republican was sent to 100 of the 150 remaining, and the type identifying the partner as a Democrat was sent to the remaining 50. The response rates were 95/100, 99/100, and 44/50, respectively. In total, we had 485 participants. No subject was allowed to answer more than one incentivized survey. Table 1 shows the treatments.

After the surveys were completed, subjects were randomly matched so that payments could be calculated. The survey was highly incentivized, 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 approximately a week of completing the experiment. Subjects were only allowed to participate in one treatment (i.e., only complete one experimental survey).

We now turn to our theoretical framework.

of the partner. We discuss this possibility in the Appendix.

# 3 Theoretical framework and hypotheses

In this section we describe a simple model that incorporates beliefs and other-regarding preferences into the decision to trust. Recall that we denote Player A as the sender and Player B as the receiver. Player A's feelings towards Player B's monetary gains may be reflected in Player A's utility function. Following Andreoni and Miller (2002) and Fisman et al. (2005), we represent Player A's utility by

$$u_A(\pi_A, \pi_B) = \alpha(I)\pi_B + [1 - \alpha(I)]\pi_A \tag{1}$$

where, $\alpha(I)$ is the weight on other's payoff, $I = s(ame)$, $o(ther)$ denotes the identity of the receiver, and $\pi_A$ and $\pi_B$ represent monetary payoffs. Equation (1) captures Gneezy and Fershtman's (2001) "taste for discrimination" as it represents, through $\alpha(I)$, the extent to which individuals from a group are willing to give away money in order to benefit others. Hence, $\alpha(I) < 0$ represents Player A loathing or disliking Player B, in our setting.

The decision to trust is inherently strategic, as it also depends on the beliefs about a partner's trustworthiness. Player A may decide to trust Player B even if she loathes him provided she is optimistic enough that Player B will honor trust. These beliefs reflect the perception about the behavior of others, and they may have little to do with own feelings about people from the same or different group. As Fershtman and Gneezy (2001) point out, the perception about others' behavior may come from stereotypes that may or may not be accurate. Specifically, Player A's decision to trust responds to her expected net benefit (Williamson's (1993) "calculative trust"), which in turn depends on her sentiments towards others, $\alpha(I)$, and the beliefs about the other player's trustworthiness, $p$. Assuming risk neutrality, and assuming preferences are as in equation (1), the utility of not trusting is equal to 5 and the expected utility of trusting is $p10 + (1-p)\alpha(I)14$. Assuming also that there are other random elements that determine the decision to trust or not to trust, which we denote as $\varepsilon_T$ and $\varepsilon_{NT}$ respectively, Player A will trust if and only if

$$p10 + (1-p)\alpha(I)14 + \varepsilon_T \geq 5 + \varepsilon_{NT}.$$

Let $F$ be a cumulative distribution function of $\varepsilon_{NT} - \varepsilon_T$ then the probability that we observe trust is given by

$$\Pr\{A\ trusts|p, I\} = F(p10 + (1-p)\alpha(I)14 - 5). \tag{2}$$

Note from equation (2) that the probability Player A trusts Player B conditional on $p$ is weakly increasing in $\alpha(I)$. Current theories describe polarization in the American society as being rooted in loathe and dislike (see, e.g., Iyengar 2012). As Carlin and Love (2014) discuss, this antagonism may be rooted in perceiving co-partisans as friends and members of the other party as foes when political competition is fierce. All this can be represented as $a(s) > \alpha(o)$: an individual benefits more from the other's payoffs when they share the same political identity. If this is the case, then Player A is more likely to trust a co-partisan than an opponent.

**Claim 1** *Conditional on the beliefs about the receiver's (Player B) trustworthiness, sender (Player A) trust rates are higher when the receiver has the same political identity than when she has a different political identity.*

Trust behavior, however, can also emerge in the extreme case of negative other-regarding concerns between members of different groups. In equation (2), for example, even if we let $\alpha(o) \leq 0$, more optimistic beliefs will make trust more likely.[10] More generally, for any fixed $\alpha(I)$ below $5/7$, the probability Player A trusts Player B is increasing in $p$.

**Claim 2** *The sender (Player A) propensity to trust is increasing in her beliefs about the receiver (Player B) trustworthiness.*

According to Akerlof and Kranton (2005, p. 12) the views as to how people should behave depend upon the situation and, in particular, between whom a transaction takes place. Republicans, for example, appealing to in-group loyalty (one of the "moral foundations" in Graham, Haidt, and Nosek (2009) by which conservatives

---

[10] In fact, with our assumptions about preferences, the probability of trusting is increasing in $p$ as long as the weight in the other's payoffs is "not too big:" $\alpha(I) < 5/7$.

abide) may reciprocate trust more often to fellow Republican than to Democrat Players A. Along the same lines, Democrats may believe Republicans' reciprocate less often than fellow Democrats, as Democrats show greater endorsement to the fairness/reciprocity "moral foundation" in Graham et al. (2009). In general, scholars have argued that political polarization in terms of partisan identity has caused people to be biased (favoring their own group) in their assessment of relative group merits (see, e.g., Oten and Wentura 1999, Mason 2014). The broader point is that beliefs about trustworthiness depend on Player B's partisan identity and on whether Player B's identity matches that of Player A. In our model, we represent this as $p = p(I)$. Our third hypothesis is therefore $p(s) > p(o)$.

**Claim 3** *Participants believe individuals with the same partisan identity are more likely to be trustworthy than individuals with different partisan identity.*

These stereotypes, however, may or may not reflect actual behavior. Using equation (1), and comparing the utility of reciprocating, $u_B(10, 10) = 10 + \eta_R$, with the utility of not reciprocating, $u_B(14, 0) = [1 - \alpha(I)]14 + \eta_{NR}$, the probability that Player B reciprocates is given by

$$\Pr\{B \ reciprocates|I\} = G(\alpha(I)14 - 4),$$

where $G$ is the cumulative distribution function of $\eta_{NR} - \eta_R$ . In this case, Player B honors trust more often when Player A identifies herself with the same party, i.e. $a(s) > \alpha(o)$.

**Claim 4** *Receiver's (Player B) rates of reciprocation are higher with co-partisan senders (Player A) than with others.*

In other words, beliefs about trustworthiness should be statistically correct.

We test these hypotheses in the next two sections. We first describe a reduced form model and then a simple structural model, based on social preferences as in Charness and Rabin (2002) and identity as in Chen and Li (2009).

|  | | Republicans | | Democrats | | | | ttest p-value |
|---|---|---|---|---|---|---|---|
|  | N | mean | N | mean | diff. | se | |
| Amount kept - dictator (range [0,5]) | 238 | 3.72 | 247 | 3.67 | 0.04 | 0.10 | 0.67 |
| Female {0,1} | 238 | 0.47 | 247 | 0.66 | -0.19 | 0.04 | 0.00 |
| White {0,1} | 238 | 0.78 | 247 | 0.69 | 0.09 | 0.04 | 0.03 |
| Black {0,1} | 238 | 0.06 | 247 | 0.09 | -0.03 | 0.02 | 0.16 |
| Married {0,1} | 238 | 0.62 | 247 | 0.43 | 0.19 | 0.04 | 0.00 |
| At least college {0,1} | 238 | 0.75 | 247 | 0.68 | 0.06 | 0.04 | 0.12 |
| NotWorking {0,1} | 238 | 0.20 | 247 | 0.25 | -0.05 | 0.04 | 0.16 |
| Income [75k,150k] {0,1} | 238 | 0.42 | 247 | 0.24 | 0.18 | 0.04 | 0.00 |
| Income (150k,350k] {0,1} | 238 | 0.12 | 247 | 0.15 | -0.03 | 0.03 | 0.37 |
| Income More 350k {0,1} | 238 | 0.01 | 247 | 0.05 | -0.04 | 0.02 | 0.02 |
| Conservative {0,1} | 238 | 0.33 | 247 | 0.03 | 0.30 | 0.03 | 0.00 |
| Very Conservative {0,1} | 238 | 0.11 | 247 | 0.01 | 0.10 | 0.02 | 0.00 |

Table 2: This table shows the means for the most important demographics in the samples of Democrats and Republicans. It also shows 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.

# 4  Results

Consistent with previous studies (see, e.g., Pew Research Center 2012), Republicans and Democrats have different demographic characteristics in our sample. Table 2 shows the means of the most relevant ones in our study. It is important to note that although Republicans and Democrats are different in many respects, they keep almost the same amount in the dictator game (US$ 3.72 and US$ 3.67 out of US$ 5, respectively). Regarding demographics, in our sample we observe a greater proportion of individuals who are male, white, conservative (and very conservative), or married with annual earnings between US$ 75k and US$ 150k among Republicans compared with Democrats. We observe no significant differences across partisanship in the proportion of black individuals, those who have at least a college education, those who are currently not working, or those who earn between US$ 150k and US$ 250k a year. These differences are consistent with those found in Pew Research Center (2012).

Before we run the reduced form and structural models, we describe the overall

14

results directly from the data in Table 3. Trust and beliefs about partners' trustworthiness are relatively high and fairly similar across political identities of Player A. Democrats trust 56% of the time and believe Player B will reciprocate 62% of the time. The same figures for Republicans are 60% and 63%, respectively (see column "Overall" in Table 3). The differences in trust rates are not statistically significant (chi-squared p-value = 0.4). The differences in beliefs are also not statistically different across partisan identities of Player A (Kolmogorov-Smirnov test for equality of distribution p-value = 1.0).

When we analyze Player A's behavior according to the partisan identity of Player B in Table 3, Democrats trust Democrats more often than they do Republicans: 63% of the time, compared with 40% of the time (chi-squared p-value < 0.01). Relative to an anonymous Player B, Democrats trust other Democrats more (63% versus 57%, chi-squared p-value = 0.4) and trust a Republican Player B less (40% versus 57%, chi-squared p-value = 0.05), although only the latter difference is statistically significant at conventional levels. 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. The distributions of beliefs are statistically different at conventional levels (Kolmogorov-Smirnov test p-value < 0.01). When compared with an anonymous Player B, the mean belief about Player B's trustworthiness is 62%. The difference between the distributions of beliefs about a Democrat and an anonymous Player B is not statistically significant at conventional levels (Kolmogorov-Smirnov test p-value = 0.6), but this difference is significant when we compare a Republican and an anonymous Player B (Kolmogorov-Smirnov test p-value = 0.02).

Perhaps surprisingly, Republicans trust a Democrat Player B more often (66% of the time) than a Republican Player B (58% of the time), although the difference is not statistically significant at conventional levels (chi-squared p-value = 0.3). Nor are there any statistical differences in trust for Republicans across other pair-wise comparisons.

In sum, we find that partisan identity has an effect only for Democrat Players A: they believe that other Democrats are more trustworthy, and they act consistently

| 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) |

Table 3: This table shows the fraction of Democrats and Republicans Player A who trust and their mean beliefs about about Player B's trustworthiness. The columns "Not revealed," "Democrat," and "Republican" refer to the treatments in which Player B (receiver) is of each one of those categories. The last column, "Overall," shows the trust rates and mean beliefs for each subpopulaiton of Democrat and Republican Player A (sender).

with this belief by more often trusting fellow Democrats. Note that this analysis of the raw data allows us to test Claim 3, which is supported by the data only for Democrats. To test Claim 1 (conditional on beliefs, Player B's partisan identity determines trust through other-regarding preferences) we need to fix beliefs to determine whether trust rates vary across Player B's identities. To test Claim 2 (the propensity to trust depends on Player A's beliefs about Player B's trustworthiness), we estimate a probability model of trust as a function of Player A's beliefs.

In the next section we first fit a reduced form model to determine whether the propensity to trust responds to beliefs—providing a test for Claim 2. Then we test Claim 1 by estimating a structural model to explore whether it is beliefs or preferences that explain these trust rates.

## 4.1 Empirical models

Player A's trust toward Player B may reflect other-regarding preferences or beliefs about Player B's trustworthiness in both cases—although identity has a significant effect only on Democrats. Is trust explained only by beliefs about Player B's trustworthiness? Or is it the case that when beliefs are fixed, identity determines Player A's trust through sentiments represented by preferences favoring co-partisans? In the following two subsections we provide evidence that contradicts Claim 1 (i.e., we find that partisan identity does not matter once beliefs are controlled for) and lends support to Claim 2 (i.e., we find that trust rates are increasing in Player A's beliefs about Player B's trustworthiness).

### 4.1.1 Reduced-form model

To analyze the impact of identity on the probability of trust, we estimate the following baseline reduced-form empirical model:

$$Trust_i = \beta_0 + \beta_s I_s^i + \beta_o I_o^i + \gamma p_i + \Gamma' X_i + \varepsilon_i$$

The subscript $i$ indexes individuals. $I_s^i$ and $I_o^i$ denote whether the political identity of Player B is either the same as ($s$) or other than ($o$) that of Player A (the baseline corresponds to the cases in which the partisan identity of Player B is not revealed to Player A). $p_i$ represents the beliefs about Player B's trustworthiness, and $X_i$ denotes demographic controls. We estimate this model for the overall sample, only for Democrat Player A, and only for Republican Player A. For each of these, we report regressions using a linear probability model without controlling for $p_i$ and without controls, a linear probability model controlling for $p_i$ and without controls, and a linear probability model controlling for $p_i$ and using demographic controls.[11] The controls included are Amount Kept in Dictator Game, Gender, Cohort, Ethnicity, Marital Status, English Writing Skills, Language at Home, Time Living in the U.S.,

---

[11]We also run Probit and Logit models (results upon request), and the qualitative results remain unchanged.

Citizenship, Full-Time School, Educational Attainment, Employment Status, Income Range, Size of Household, Minors at Home, and Political Orientation (from very liberal to very conservative).

Table 4 columns (1)—(3) show the results for the overall sample, columns (4)—(6) the results for Democrat Player A, and columns (7)—(9) the results for Republican Player A. Overall, we find that without controlling for beliefs about partner's trustworthiness, Player B's identity significantly determines trust only when Player A is a Democrat (which is another way of seeing the results in Table 3 for Democrat Players A). This result, however, confounds the impact of identity through beliefs. In short, when controlling for beliefs, the coefficient of identity is no longer significant regardless of Player A's partisan identity.

Moreover, Player A's beliefs about Player B's trustworthiness are highly significant for all the specifications. Columns (2) and (3) in Table 4 show that a change in one standard deviation in beliefs (23%) makes it 10% more likely that Player A trusts Player B. The large share of this effect is explained by Democrat Players A. On the one hand, when looking at the coefficient of beliefs only for Democrat Players A, an increase in one standard deviation in beliefs yields a 12% increase in the likelihood of Player A trusting Player B, according to this model. The corresponding effect for Republican Players A is roughly 9%.

These results give no support to Claim 1, as Player B's partisan identity does not have an effect on Player A's trust rates when beliefs are controlled for. Claim 2, however, is borne out by the data: Player A's propensity to trust is increasing in her beliefs about Player B's trustworthiness. This effect is more pronounced for Democrat Players A.

In the next section, we exploit the preferences specifications used in Charness and Rabin (2002) and Chen and Li (2009) to provide a further test of Claim 1 and Claim 2 by estimating a simple structural model of trust.

18

|  | Overall (Player A) | | | Democrat Sender (Player A) | | | Republican Sender (Player A) | | |
|---|---|---|---|---|---|---|---|---|---|
|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|  | Trust | Trust | Trust | Trust | Trust | Trust | Trust | Trust | Trust |
| Same partisan identity (Is) | 0.03 | 0.02 | 0.01 | 0.06 | 0.04 | -0.01 | 0.00 | 0.00 | 0.04 |
|  | (0.05) | (0.05) | (0.05) | (0.07) | (0.07) | (0.07) | (0.07) | (0.07) | (0.08) |
| Different partisan identity (Io) | -0.05 | -0.02 | -0.01 | -0.17* | -0.11 | -0.08 | 0.08 | 0.08 | 0.09 |
|  | (0.06) | (0.06) | (0.06) | (0.09) | (0.09) | (0.09) | (0.09) | (0.09) | (0.10) |
| Beliefs: B's trustworthiness |  | 0.41*** | 0.42*** |  | 0.51*** | 0.49*** |  | 0.24* | 0.34** |
|  |  | (0.10) | (0.10) |  | (0.13) | (0.15) |  | (0.15) | (0.15) |
| constant | 0.57*** | 0.31*** | 0.76*** | 0.57*** | 0.25*** | 0.99*** | 0.58*** | 0.42*** | -0.10 |
|  | (0.04) | (0.07) | (0.26) | (0.05) | (0.10) | (0.32) | (0.05) | (0.11) | (0.53) |
| CONTROLS | NO | NO | YES | NO | NO | YES | NO | NO | YES |
| N | 485 | 485 | 485 | 247 | 247 | 247 | 238 | 238 | 238 |
| R-sq | 0.00 | 0.04 | 0.14 | 0.03 | 0.08 | 0.24 | 0.00 | 0.02 | 0.19 |

Table 4: This table shows a linear probability reduced-form model. The dependent variable is whether Player A (sender) trusts Player B (receiver). The explanatory variables shown represent whether Player B's partisan identity coincides with Player A's, whetherthey do not coincide, Player A's beliefs about Player B's trustworthiness, and the constant. In columns (3), (6), and (9) demographic controls were considered. The first three columns pool all the observations, the second three columns consider only Democrat Player A, and the last three consider only Republican Player A.

### 4.1.2 Structural model

In this section we are more specific about how we model sentiments by representing preferences using the Charness and Rabin (2002) baseline model and estimating the actual parameters from Chen and Li (2009). According to Charness and Rabin (2002), equation (1) can be written as

$$
\begin{aligned}
u_A(\pi_A, \pi_B) &= \alpha\pi_B + [1 - \alpha]\pi_A \\
&= (\lambda l + \eta h)\pi_B + [1 - (\lambda l + \eta h)]\pi_A,
\end{aligned}
$$

where $h = 1$ if $\pi_B > \pi_A$ and $h = 0$ otherwise; $l = 1$ if $\pi_B < \pi_A$ and $l = 0$ otherwise; and $\lambda$, $\eta$ are parameters to be estimated. Following Chen and Li (2009)

19

we incorporate group identity by setting

$$\alpha = \lambda(1 + I_s q + I_o \overline{q})l + \eta(1 + I_s w + I_o \overline{w})h, \tag{3}$$

where $I_s = 1$ if Player A and Player B share the same partisan identity and zero otherwise, and $I_o = 1$ if they hold different partisan identities and zero otherwise. The parameters $q$ and $w$ (respectively $\overline{q}$ and $\overline{w}$) represent the utility benefit Player A receives for trusting a Player B who shares the same (respectively different) partisan ideology.

Replacing equation (3) with equation (2) yields to

$$
\begin{aligned}
\Pr\{A\ Trust|p\} &= F(p10 + (1-p)\alpha14 - 5) \\
&= F(\delta_0 + \delta_1 p + \delta_2 I_s + \delta_3 I_o + \delta_4 I_s p + \delta_5 I_o p),
\end{aligned}
$$

where $\delta_0 = (14\eta - 5)$, $\delta_1 = (10 - 14\eta)$, $\delta_2 = 14\eta w$, $\delta_3 = 14\eta\overline{w}$, $\delta_4 = -14\eta w$, and $\delta_5 = -14\eta\overline{w}$.

Claim 1 states that if trusting decisions are rooted on dislike or loathe, i.e., partisan identity matters through preferences, then for given beliefs about Player B's trustworthiness, $p$, the probability Player A trusts a co-partisan Player B ($I_s = 1$) should be larger than the probability Player A trusts a Player B with different a partisan ideology ($I_o = 1$). Hence, using this model, Claim 1 can be stated as

$$F(\delta_0 + \delta_2 + (\delta_1 + \delta_4)p) - F(\delta_0 + \delta_3 + (\delta_1 + \delta_5)p) > 0. \tag{4}$$

We estimate these probabilities using a linear model without controls.[12] Figure 2 shows the difference in (4) is not statistically different from zero for all values of $p$. This result does not support Claim 1, and it corroborates our results from the reduced-form model in the previous subsection.

---

[12]This result remains unchanged if we incrporate demographic controls and if we use nonlinear probability models such as Probit or Logit.

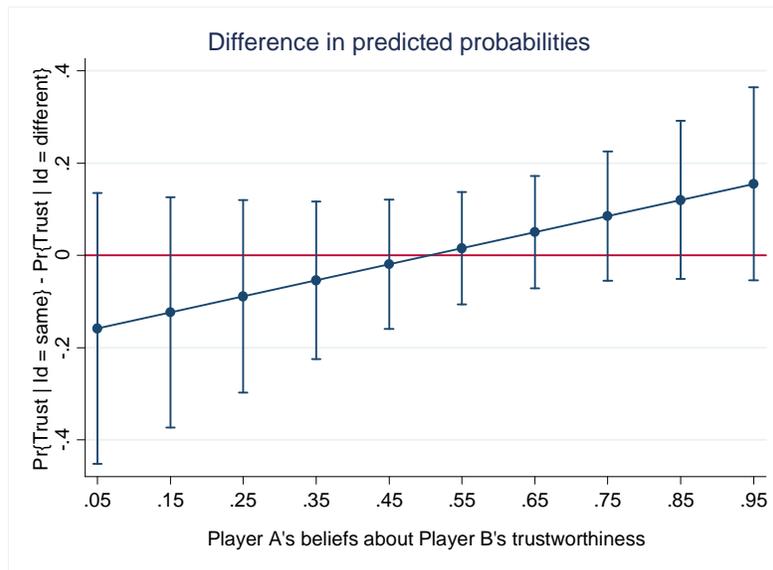Figure 2: This figure shows the difference in predicted probabilities (using a linear probability model) of Player A (sender) trusting Player B (receiver) when Player B shares the same partisan identity as Player A and when Player B holds a different partisan identity. The standard errors used to compute the confidence intervals represented by the vertical segments were calculated using the Delta method.
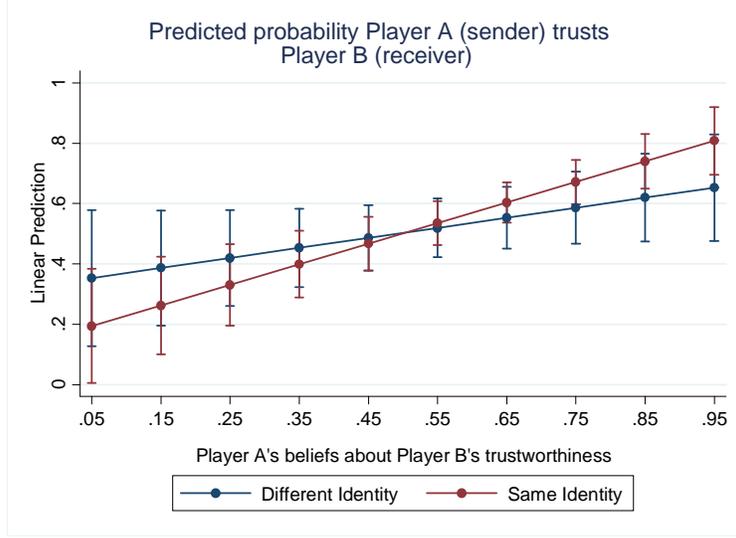
**Figure 3:** This figure shows the predicted probability Player A (sender) trusts a Player B (receiver) from the same partisan identity (green line) and from a different partisan identity (blue line) for different values of p. The standard errors used to compute the confidence intervals represented by the vertical segments were calculated using the Delta method.

In order to test Claim 2, we plot each one of the terms on the left-hand side of (4) in Figure 3 estimated for the full sample. Both are increasing in $p$. Moreover, the marginal effect of $p$ in the first term (when matched with a same-partisan-ideology Player B, $I_s = 1$), $F(\delta_0 + \delta_2 + (\delta_1 + \delta_4)p)$, is 68% (t-test p-value $< 0.01$) and the marginal effect of $p$ in the second term (when matched with a different-partisan-ideology Player B, $I_o = 1$), $F(\delta_0 + \delta_3 + (\delta_1 + \delta_5)p)$, is 33% (t-test p-value $= 0.09$). These results support Claim 2, as in the previous subsection.

In sum, the results from reduced-form and structural models suggest that trust behavior is driven by beliefs of trustworthiness rather than by affinity or dislike for a particular member's partisan identity.

Beliefs are driving trust in our setting. Do Player A's beliefs depend on the partisan identity of Player B? Before we move on to testing Claim 4 (whether beliefs are statistically correct), we answer this question in the next section.

# 5    Beliefs and Identity

In this section we ask whether partisan identity has an effect on out-of-group beliefs about a partner's trustworthiness, that is, whether in-group and out-of-group stereotypes determine trust.

Table 5 shows several variations of the baseline empirical model:

$$p_i = \gamma_0 + \gamma_s I_s^i + \gamma_o I_o^i + \Theta\prime X_i + \xi_i.$$

The subscript $i$ indexes individuals. $p_i$ represents the beliefs about Player B's trustworthiness, $I_s^i$ and $I_o^i$ denote whether the political identity of Player B is either the same as $(s)$ or other than $(o)$ that of Player A (the baseline corresponds to the case when individuals do not know the identity of the partner), and $X_i$ denotes demographic controls. We estimate this model for the overall sample, only for Democrat Player A (sender), and only for Republican Player A. For each of these subsamples, we report regressions using a linear probability model without demographic controls and a linear probability model using demographic controls. The controls included are Amount Kept in Dictator Game, Gender, Cohort, Ethnicity, Marital Status, English Writing Skills, Language at Home, Time Living in the U.S., Citizenship, Full-Time School, Educational Attainment, Employment Status, Income Range, Size of Household, Minors at Home, and Political Orientation (from very liberal to very conservative).

Table 5 columns (1) and (2) show the results for the overall sample, columns (3) and (4) the results for Democrat Player A, and columns (5) and (6) the results for Republican Player A. Our results are consistent with Claim 3. We find that partisan identity shapes beliefs. Overall, individuals are 6% less optimistic (t-test p-value $< 0.05$) about Player B's trustworthiness if his identity is not revealed. There is no difference in Player A's beliefs about Player B's trustworthiness between same-identity Player B and Player B whose identity is not revealed. As a result, we could say Player A's beliefs are 6% less optimistic when Player B's partisan identity is

23

|                                  | Overall |        | Democrat Sender |          | Republican Sender |        |
|----------------------------------|---------|--------|-----------------|----------|-------------------|--------|
|                                  | (1)     | (2)    | (3)             | (4)      | (5)               | (6)    |
|                                  | p       | p      | p               | p        | p                 | p      |
| Same partisan identity (Is)      | 0.01    | 0      | 0.04            | 0.03     | -0.02             | -0.04  |
|                                  | (0.02)  | (0.02) | (0.03)          | (0.03)   | (0.03)            | (0.04) |
| Different partisan identity (Io) | -0.06** | -0.06**| -0.11***        | -0.12*** | -0.01             | 0      |
|                                  | (0.03)  | (0.03) | (0.04)          | (0.04)   | (0.04)            | (0.05) |
| _cons                            | 0.63*** | 0.57***| 0.62***         | 0.64***  | 0.64***           | 0.58** |
|                                  | (0.02)  | (0.11) | (0.02)          | (0.14)   | (0.02)            | (0.24) |
| CONTROLS                         | NO      | YES    | NO              | YES      | NO                | YES    |
| N                                | 485     | 485    | 247             | 247      | 238               | 238    |
| R-sq                             | 0.015   | 0.144  | 0.054           | 0.269    | 0.002             | 0.134  |

Table 5: This table shows a linear reduced-form model in which the dependent variables are the beliefs a given Player A (sender) holds about Players B's trustworthiness. The explanatory variables consist of an indicator variable on whether Player B (receiver) is from the same, $I_s$, or differen,t $I_o$, partisan identity. Columns (1) and (2) show the results for the overall sample (with and without demographic controls, respectively), columns (3) and (4) the results for democrat Player A (with and without demographic controls, respectively), and columns (5) and (6) the results for Republican Player A (with and without demographic controls, respectively).

different. This result, however, is driven by Democrat Players A. Democrats are 11—12% (t-test p-value < 0.01) less likely to trust a Republican, when compared to a Player B whose partisan identity is unknown or whose partisan identity is also Democrat. For the Republicans in our sample, partisan identity does not seem to affect beliefs.

These results show that Democrats perceive Republicans as less trustworthy than Democrats, and that Republicans' perceptions about Player B's trustworthiness do not depend on Player B's partisan identity. Are these perceptions statistically correct? In the next section we explore this question in order to test Claim 4.

## 5.1  Beliefs and actual trustworthiness

We have provided tests for Claims 1, 2, and 3. In this section, we proceed to test Claim 4. In general, Table 6 shows that beliefs about a partner's trustworthiness are not statistically correct—they are more pessimistic. Overall, Democrat Players

| 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) | (0.24) |
| | | | |
| 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) | (0.22) |
| | | | |
| Player B's actual trustworthiness | 0.81 | 0.87 | 0.8 |
| # of participants / Total | 38/47 | 86/99 | 389/485 |

Table 6: This table shows the average beliefs about Player B's trustworthiness and the actual rate of Player B's reciprocation. The "Anonymous" column corresponds to the treatment in which neither player's indentiy is revealed. 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 chose to reciprocate, [10,10]. The "Overall" column shows the average beliefs for all Democrats and Republican Players A and the rate of reciprocation overall in the sample.

A believe on average that 62% of Players B will reciprocate, but 80% end up doing so. Not all, of course, 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.

These differences are more pronounced when we separate them by the partisan 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 were correct in their guesses: these 4 Democrats believed at least 90% of Republican Players B would reciprocate trust. The difference is less pronounced when Player B is Democrat and Player A is Republican: mean beliefs are 63% and the actual reciprocation rate is 81%. Twenty-five percent (11 out of 44) of Republican Players A believed at least 80% of Democrat Players B would reciprocate trust.

As we saw in the previous section, Player B's partisan identity has a statistically significant effect on beliefs (see Table 5 and the columns in Table 6) 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. Although this difference is not statistically significant (chi-squared p-value = 0.17), if anything, it points in the other direction: Republicans are more trustworthy than Democrats, when matched with a Democrat Player A (sender). In contrast, Republicans do not show a statistically different perception about Player B's trustworthiness across partisan identity (mean beliefs are 63% when matched with a Democrat Player B versus 62% when matched with a Republican Player B), and there is also no significant differences in terms of actual behavior: 81% versus 87%, chi-squared p-value =0.34. Although not statistically significant, Republican Players B tend to reciprocate more often than Democrat Players B when matched with a Republican Player A.

Overall, Player A's beliefs about Player B's trustworthiness are lower than actual reciprocation rates. For Democrat Players A, beliefs turn out to be incorrect, which does not support Claim 4. For Republican Players A, there is no statistical difference in Player B reciprocation rates between Democrats and Republicans, which is consistent with Republican Player A's beliefs. In total, these results do not support Claim 4.

In the next section we explore the role of identity in the trustworthy behavior.

### 5.1.1 Trustworthy behavior and preferences

The decision to reciprocate does not hinge on beliefs. A further test on whether preferences, on top of beliefs, play a role is to explore the effect of partisan identity on the decision to reciprocate. Table 7 reports results for the following linear probability model:

$$Trustworthiness_i = \gamma_0 + \gamma_s I_s^i + \gamma_o I_o^i + W'X_i + \omega_i.$$

The subscript $i$ indexes individuals. $I_s^i$ and $I_o^i$ denote whether the political identity of Player B is either the same ($s$) as or other than ($o$) that of Player A (the baseline corresponds to the case in which the partisan identity of Player B is not revealed to

| | Overall | | Democrat Player B | | Republican Player B | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Trustworthy | Trustworthy | Trustworthy | Trustworthy | Trustworthy | Trustworthy |
| Same partisan identity (Is) | 0.15*** | 0.07* | -0.01 | -0.02 | 0.31*** | 0.12** |
| | (0.04) | (0.04) | (0.05) | (0.05) | (0.06) | (0.06) |
| Different partisan identity (Io) | 0.16*** | 0.08* | -0.05 | -0.07 | 0.37*** | 0.16** |
| | (0.05) | (0.05) | (0.06) | (0.07) | (0.07) | (0.07) |
| _cons | 0.71*** | 1.20*** | 0.86*** | 1.29*** | 0.56*** | 1.17*** |
| | (0.03) | (0.18) | (0.04) | (0.22) | (0.04) | (0.37) |
| CONTROLS | NO | YES | NO | YES | NO | YES |
| N | 485 | 485 | 247 | 247 | 238 | 238 |
| R-sq | 0.034 | 0.305 | 0.003 | 0.231 | 0.145 | 0.481 |

Table 7: This table shows a reduced-form linear probability model of reciprocation (trustworthiness) as a function of whether Player A's (sender) partisan idenity coincides with individual $i$ which in this case is Player B (receiver).

Player A). $X_i$ denotes the usual demographic controls. Overall, a Wald test fails to reject the hypothesis that the coefficients for the dummy variables $(I_s)$ and $(I_o)$ are different from one another. That is, the decision to reciprocate does not significantly depend on the partisan identity of the sender when this identity is known.

It is worth noting that only Republican Players B tend to reciprocate trust significantly more often when Player A's partisan identity is revealed than when it is not. This effect diminishes when we incorporate demographic controls into the estimation. This suggests that some of the preferences for reciprocation may be driven by subpopulations that may condition their behavior when interacting with individuals from known and unknown partisan identities. Thus, in the next section we explore the differences in trust and trustworthiness for each demographic characteristic when we vary the identity of the matched partner.

# 6    Manipulating beliefs

Given the importance of beliefs in trust behavior, we further explore whether pessimistic and incorrect beliefs can be manipulated. We run four additional experimen-

tal surveys in which we reveal the rate of reciprocation by Players B from the main treatments. No other change is introduced to the format or content of the survey relative to the survey in the main treatments.[13] 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 91% 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 79% 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 88% 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.

---

[13]The only variation in Q3 with respect to the main treatments is the information about reciprocation rates. The question reads as follows: "You will receive a payment based on your decision in the following scenario: You will be matched to another survey participant who considered him or herself to be a [DEMOCRAT/REPUBLICAN]. You need to decide between the following two options: 1) You and the other participant each receive $5 OR You let the other participant choose. He/she will decide between one of two options: i) You receive $0 and he/she receives $14 or ii) Each of you receives $10.

In previous surveys, when the other Democrat participant was designated to choose, he/she chose alternative ii) "Each of you receives $10" X% of the time. Please enter your decision:

■       I choose option 1) (1)
■       I choose option 2) (2)"

## 6.1 Effect of revealing previous reciprocation rates on beliefs about trustworthiness

Does revealing actual reciprocation rates change beliefs? Figure 4 shows the distribution of beliefs declared by Democrats for the baseline condition in which no information about actual reciprocation rates is revealed and the treatment condition in which participants are informed about this. Panel 1 shows the distribution of beliefs by Democrat Players A facing a Democrat Player B. The red line represents the revealed reciprocation rate from the corresponding treatment in the previously run main treatments. Many participants, 46% (21 out of 46), forecasted that Democrat Players B reciprocate trust at the same rate as that previously revealed (85%). In addition, the declared beliefs in this beliefs treatment are statistically larger than the corresponding beliefs in the main treatments (Kolmogorov-Smirnov test p-value < 0.01). Panel 2 shows the distribution of beliefs by Democrat Players A facing a Republican Player B. Although we reject equality of distributions (Kolomogorov-Smirnov test p-value < 0.01), Democrat Players A seldom declare the rate of cooperation by Republican Players B to coincide with the one revealed (91%). In other words, it seems Democrats become more optimistic about Republicans' cooperation rate, but their beliefs are still pessimistic.

Figure 5 shows the distribution of beliefs declared by Republicans for the baseline condition in which no information about actual reciprocation rates is revealed and the treatment condition in which participants are informed about the actual rates. Panel 1 shows the distribution of beliefs by Republican Player A facing a Democrat Player B. The red line represents the revealed reciprocation rate from the corresponding treatment in the previously run main treatments. At conventional levels, we fail to reject equality of distributions about beliefs before and after revelation (combined Kolmogorov-Smirnov test p-value = 0.96). Panel 2 shows the distribution of beliefs by a Republican Player A facing a Republican Player B. Data reveals that beliefs are more skewed in the belief treatment and, if anything, Players A become more pessimistic about a Republican Player B's reciprocation rate (Kolmogorov-Smirnov
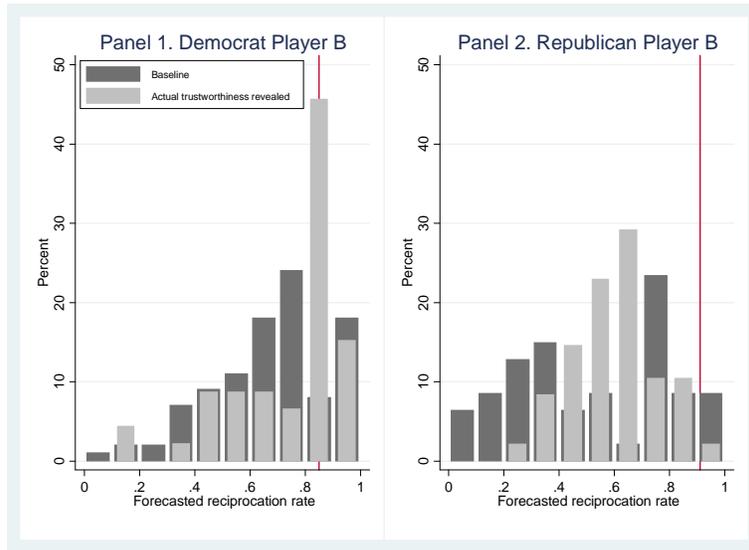
Figure 4: This figure shows the empirical distribution of beliefs declared in the main treatments (dark gray) and the beliefs treatments (light gray) when Player A is Democrat and Player B is Democrat (Panel 1) and Republican (Panel 2). In both figures, we reject the hypothesis that the baseline (main treatments) contains more optimistic beliefs than the beliefs treatment (Kolomogorv-Smirnov test p-value < 0.01 in each case). The red line represents the revealed reciprocation rate from the corresponding treatment in the previously run main treatments.
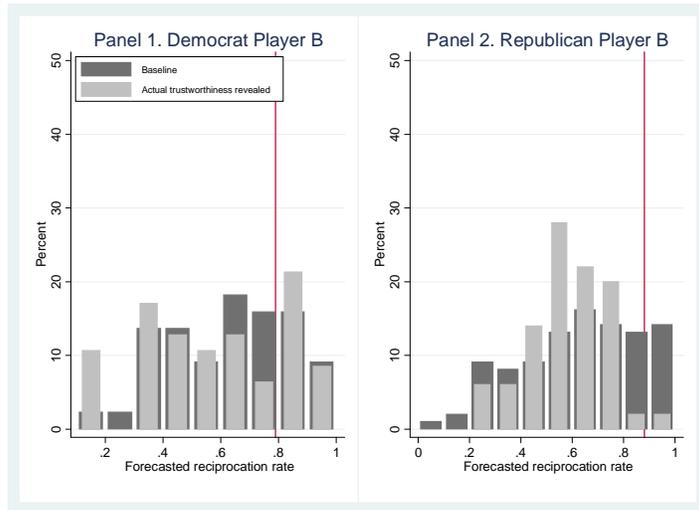
Figure 5: This figure shows the empirical distribution of beliefs declared in the main treatments (dark gray) and the beliefs treatments (light gray) when Player A is Republican and Player B is Democrat (Panel 1) and Republican (Panel 2). In Panel 1, the distributions are not statiscally different (Kolmogorov-Smirnov test for equality of distributions = 0.96). In Panel 2, the beliefs are more skewed in the belief treatment compared with the main treatment; and if anything, Players A become more pessimistic about a Republican Player B's reciprocation rate (Kolmogorov-Smirnov test p-value < 0.05). The red line represents the revealed reciprocation rate from the corresponding treatment in the previously run main treatments.

test p-value < 0.05).

From these results we can conclude that revealing previous cooperation rates has a positive effect on beliefs about Player B's trustworthiness only when Player A is Democrat. When Player B is Republican, on the other hand, revealing information has no, or even a negative, effect on beliefs. In particular, revealing information about actual reciprocation rates changes little the beliefs about the other party's members' trustworthiness. This results are consistent with the evidence presented in Friesen, Campbell, and Kay (2014).

a.

| Democrat's partner identity is… | Democrat | Republican |
|---|---|---|
| Fraction of Trust (main treatment) | 0.63 | 0.40 |
| # of Players A who Trust/Total | 63/100 | 19/47 |
| Fraction of Trust (beliefs treatment) | 0.76 | 0.35 |
| # of Players A who Trust/Total | 35/46 | 17/48 |
| Difference | 0.13 | -0.05 |

***: significant at 1% level

b.

| Republican's partner identity is… | Democrat | Republican |
|---|---|---|
| Fraction of Trust (main treatment) | 0.66 | 0.58 |
| # of Players A who Trust/Total | 29/44 | 57/99 |
| Fraction of Trust (beliefs treatment) | 0.51 | 0.16 |
| # of Players A who Trust/Total | 24/47 | 8/50 |
| Difference | -0.15 | -0.42*** |

***: significant at 1% level

Table 8: This table shows trust for Democrat (panel a) and Republican (panel b) Players A in the main and belief treatments.

## 6.2 Effect of revealing previous reciprocation rates on trust

We have seen beliefs positively change only for Democrats. Does this increase trust behavior? Focusing first on Democrat Players A, revealing reciprocation rates induces higher trust in Democrat Players B, from 63% (63 out of 100) to 76% (35 out of 46), although the difference is not statistically significant at conventional levels (chi-squared test p-value = 0.12). Revealing actual reciprocation rates does not affect trust when Player B is Republican. Trust rates go from 40% (19 out of 47) to 35% (17 out of 48, chi-squared p-value = 0.62).

For Republican Players A who face Democrat Players B, even though trust rates decrease across treatments, they are not statistically different (66%, 29 out of 100

in the main treatment; and 51%, 24 out of 47, in the beliefs treatment, chi-squared test p-value = 0.15). Finally, for Republican Players A facing Republican Players B, the trust rate decrease by a significant amount: 58% (57 out of 99) in the main treatment to 16% (8 out of 50) in the beliefs treatment (chi-squared test p-value < 0.01). Table 8 summarizes these results.

Trust does not seem to respond to the revelation of the actual reciprocation rates in previous games. We expected that, at least for Democrat Players A, trust rates would increase when reciprocation rates are revealed. We conclude from this exercise that manipulating incorrect beliefs by revealing the true state of the world is not easy. Thus, beliefs about reciprocity across partisanship may not be very malleable.

## 6.3   Conclusion

Political polarization is an important phenomenon that motivates many recent popular and academic articles in economics (see, e.g., Fisman, Jakiela, and Kariv 2014). The emergence of polarization is often explained as being driven by those with different political identities having dislike or even hatred for one other. We, however, offer another possibility: people have different beliefs about how trustworthy people of different ideologies are.

To explore both of these possibilities, we explicitly studied the relationship between political ideology and trust among a representative US population. In particular, we had paired subjects play a simple trust game with either the same or opposite political identity partner. We found that there are partisan identity-based differences in trusting rates. Whereas Republicans do not exhibit different trust rates between partners of different partisan identities, Democrats trust partners of their own partisan identity more than Republicans. The mechanism that explains this difference seems to be driven by beliefs about partner trustworthiness and not by a taste for discrimination based on partisan identity. The source of trusting behavior is important—whether it comes primarily from beliefs or taste—not only because trust is defined in terms of beliefs, but also because beliefs reflect stereotypes, whereas preferences reflect individual taste for discrimination. We also found that although

there was no difference in trustworthiness as a function of partisan identity, overall, individuals held beliefs that were much more pessimistic than actual trustworthiness, regardless of the partisan identity of both the trustee and trustor.

It is encouraging that polarization is more likely driven by incorrect beliefs than distaste for others: It seems that it would be easier to change incorrect beliefs than taste-based preferences. However, these incorrect beliefs are apparently not easy to change; we were only able to modestly nudge beliefs towards reality. We leave this challenge to future research: discovering the most effective means by which to correct beliefs. Both theory-building and empirical analysis on this issue should prove helpful.

We also note that we only explored the two primary political ideologies in the United States. It would be interesting to explore additional ideologies, as well as those in different countries. We leave this exploration to future research.

# References

[1] Abramowitz, Alan. (2007). "Constraint, Ideology and Polarization in the American Electorate: Evidence from the 2006 Cooperative Congressional Election Study." Presented at the Annual Meeting of the American Political Science Association, Chicago, IL.

[2] Abramowitz, Alan. (2010). The Disappearing Center: Engaged Citizens, Polarization, and American Democracy. New Haven, CT: Yale University Press.

[3] Abramowitz, Alan, and Kyle Saunders. (2005). "Why Can't We All Just Get Along? The Reality of a Polarized America." The Forum: A Journal of Applied Research in Contemporary Politics 3(2): 1–22.

[4] Anderson, L. R., Mellor, J. M., & Milyo, J. (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, J. and J. Miller (2002) "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 70 (2), 737-753.

[6] Akerlof, G. A., & Kranton, R. E. (2000). Economics and identity. *Quarterly Journal of Economics*, 715-753.

[7] Akerlof, G. A., & Kranton, R. E. (2005). Identity and the Economics of Organizations. *Journal of Economic Perspectives*, 9-32.

[8] Arrow, K. (1974). The Limits of Organization. NY: Norton.

[9] Ashraf, N., Bohnet, I., & Piankov, N. (2006). "Decomposing trust and trustworthiness." *Experimental Economics*, 9(3), 193-208.

[10] Becker, Gary, (1957) "The Economics of Discrimination" (Chicago: University of Chicago Press, 1957).

[11] Becker, Gary, (1993) "The Evidence against Banks Doesn't Prove Bias," Business Week, April 19.

[12] Berg, Joyce, Dickhaut, John, et McCabe, Kevin (1995). "Trust, Reciprocity, and Social History." *Games and Economic Behavior*, vol. 10, no 1, p. 122-142.

[13] Bohnet, I., & Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 55(4), 467-484.

[14] Bolle, F. (1998). "Rewarding trust: An experimental study." *Theory and Decision*, 45(1), 83-98.

[15] Brewer, Marilynn, and Kathleen P. Pierce. (2005). "Social Identity Complexity and OutgroupTolerance." *Personality and Social Psychology Bulletin* 31(3): 428–37.

[16] Camerer, Colin, and Keith Weigelt (1988). "Experimental Tests of a Sequential Equilibrium Reputation Model." *Econometrica*, 56, 1–36.

[17] Carlin, R. E., & Love, G. J. (2013). The Politics of Interpersonal Trust and Reciprocity: An Experimental Approach. *Political Behavior*, 35(1), 43-63.

[18] Chambers, J. R., & Melnyk, D. (2006). Why do I hate thee? Conflict misperceptions and intergroup mistrust. *Personality and Social Psychology Bulletin*, 32(10), 1295-1311.

[19] Charness, G., & Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, 74(6), 1579-1601.

[20] Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics*, 817-869.

[21] Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *The American Economic Review*, 431-457.

[22] Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior,* 46(2), 260-281.

[23] Dufwenberg, M., & Gneezy, U. (2000). "Measuring beliefs in an experimental lost wallet game." *Games and Economic Behavior*, 30(2), 163-182.

[24] Eckel, C. C., & Wilson, R. K. (2004). Is trust a risky decision?. *Journal of Economic Behavior & Organization*, 55(4), 447-465.

[25] Falk, A., & Zehnder, C. (2013). A city-wide experiment on trust discrimination. *Journal of Public Economics*, 100, 15-27.

[26] Fehr, E. (2009). On the economics and biology of trust. *Journal of the European Economic Association*, 7(2-3), 235-266.

[27] Fershtman, C., & Gneezy, U. (2001). Discrimination in a segmented society: An experimental approach. *Quarterly Journal of Economics*, 351-377.

[28] Fisman, R., P. Jakiela, and S. Kariv (2014). "The Distributional Preferences of Americans," NBER Working Paper No. 20145, May 2014.

[29] Fisman, R., S. Kariv, and D. Markovits (2007) "Individual Preferences for Giving," *American Economic Review*, 97 (5), 1858–1876.

[30] Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 25-42.

[31] Friesen, J. P., Campbell, T. H., & Kay, A. C. (2014). The Psychological Advantage of Unfalsifiability: The Appeal of Untestable Religious and Political Ideologies. *Journal of Personality and Social Psychology*, Nov 17.

[32] Garbarino, E., & Slonim, R. (2009). The robustness of trust and reciprocity across a heterogeneous US population. *Journal of Economic Behavior & Organization*, 69(3), 226-240.

[33] Geanakoplos, J., Pearce, D., & Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior*, 1(1), 60-79.

[34] Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, 96(5), 1029.

[35] Haidt, J. and Sam Adams (2014) "Ten Causes of America's Rising Polarization," mimeo.

[36] Houser, D., Schunk, D., & Winter, J. (2010). "Distinguishing trust from risk: An anatomy of the investment game." *Journal of Economic Behavior & Organization*, 74(1), 72-81.

[37] Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology a social identity perspective on polarization. *Public Opinion Quarterly*, 76(3), 405-431.

[38] Iyengar, S., & Westwood, S. J. (2014). Fear and Loathing Across Party Lines: New Evidence on Group Polarization. *American Journal of Political Science*.

[39] Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics*, 1251-1288.

[40] Mason, L. (2014) "I Disrespectfully Agree": The Differential Effects of Partisan Sorting on Social and Issue Polarization. *American Journal of Political Science.*

[41] Ortmann, A., Fitzgerald, J., & Boeing, C. (2000). Trust, reciprocity, and social history: A re-examination. Experimental Economics, 3(1), 81-100.

[42] Pew Research Center (2012), "A Closer Look at the Parties in 2012," **??** .

[43] Pew Research Center (2014), "Political Polarization in the American Public," June 2014.

[44] Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American Economic Review,* 659-661.

[45] Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American Economic Review,* 1281-1302.

[46] Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review,* 23(3), 393-404.

[47] Sapienza, Paola, Toldra-Simats, Anna, & Zingales, Luigi (2013). Understanding trust. *The Economic Journal,* 123(573), 1313-1332.

[48] Williamson, O. E. (1993). Calculativeness, trust, and economic organization. *Journal of Law and Economics,* 453-486.

# 7　Appendix

## 7.1　Survey instrument, baseline condition: Without revealing partner's political identity

Political Ideology

* For the following questions, you will be paid in Amazon gift certificates according to how you choose to answer them.

Q1 You will receive a payment according to your decision in the following scenario: You have a total of $5 to divide between yourself and another survey participant in any way you want (in increments of $1).

_____ Decide how many dollars you hold (1)

_____ Decide how many dollars you pass (2)

Q2 You will receive a payment based on your decision in the following scenario: You will be matched to another survey participant. You need to decide between the following two options: 1) You and the other participant each receive $5 2) You let the other participant choose. He/she will decide between one of two options: i) You receive $0 and he/she receives $14 or ii) Each of you receives $10. Please enter your decision:

■      I choose option 1) (1)

■      I choose option 2) (2)

Q3 You will receive a payment based on your decision in the following scenario: You will be matched to another survey participant. The other participant can choose for each of you to receive $5 or instead he/she can let you decide between one of two options: 1) You and the other participant each receive $10 2) You receive $14 and the other participant receives $0 In case the other participant lets you choose, please enter your decision:

■      I choose option 1) (1)

■      I choose option 2) (2)

Q4 You will receive an additional $3 if you guess the correct percentage range of participants that choose option 1) for the above question: Between:

■      0 and 9% (1)

■      10 and 19% (2)

■      20 and 29% (3)

■   30 and 39% (4)

■   40 and 49% (5)

■   50 and 59% (6)

■   60 and 69% (7)

■   70 and 79% (8)

■   80 and 89% (9)

■   90 and 100% (10)

Q5 Please answer the following questions:

A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? (1)

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (2)

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (3)

Q6 What is your political orientation?

■   Very Liberal (1)

■   Liberal (2)

■   Moderate (3)

■   Conservative (4)

■   Very Conservative (5)

■   Don't know (6)

Q7 What is your annual household income?

■   less than $10,000 (1)

■   $10,001 to $20,000 (2)

■   $20,001 to $50,000 (3)

■   $50,001 to $75,000 (4)

■   $75,001 to $100,000 (5)

■   $100,001 to $150,000 (6)

■     $150,001 to $250,000 (7)

■     $250,001 to $350,000 (8)

■     more than $350,000 (9)

Q8 Based on your political views, would you consider yourself to be:

■     A Democrat (1)

■     A Republican (2)

■     An Independent (3)

■     Other (4)

* Click continue to finish your survey.You will receive your total final payment in the coming week. Thank you for participating!

## 7.2   Survey instrument, treatment conditions: Revealing partner's political identity

Political Ideology

* For the following questions, you will be paid in Amazon gift certificates according to how you choose to answer them.

Q1 You will receive a payment according to your decision in the following scenario: You have a total of $5 to divide between yourself and another survey participant in any way you want (in increments of $1).

_____ Decide how many dollars you hold (1)

_____ Decide how many dollars you pass (2)

Q2 You will receive a payment based on your decision in the following scenario: You will be matched to another survey participant **who considered him or herself to be [POLITICAL IDENTITY]**. You need to decide between the following two options: 1) You and the other participant each receive $5 2) You let the other

41

participant choose. He/she will decide between one of two options: i) You receive $0 and he/she receives $14 or ii) Each of you receives $10. Please enter your decision:

- ■     I choose option 1) (1)
- ■     I choose option 2) (2)

Q3 You will receive a payment based on your decision in the following scenario: You will be matched to another survey participant **who considered him or herself to be [POLITICAL IDENTITY]**. The other participant can choose for each of you to receive $5 or instead he/she can let you decide between one of two options: 1) You and the other participant each receive $10 2) You receive $14 and the other participant receives $0 In case the other participant lets you choose, please enter your decision:

- ■     I choose option 1) (1)
- ■     I choose option 2) (2)

Q4 You will receive an additional $3 if you guess the correct percentage range of [**POLITICAL IDENTITY**] participants that choose option 1) for the above question: Between:

- ■     0 and 9% (1)
- ■     10 and 19% (2)
- ■     20 and 29% (3)
- ■     30 and 39% (4)
- ■     40 and 49% (5)
- ■     50 and 59% (6)
- ■     60 and 69% (7)
- ■     70 and 79% (8)
- ■     80 and 89% (9)
- ■     90 and 100% (10)

Q5 Please answer the following questions:

A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? (1)

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (2)

In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (3)

Q6 What is your political orientation?

- ■ Very Liberal (1)
- ■ Liberal (2)
- ■ Moderate (3)
- ■ Conservative (4)
- ■ Very Conservative (5)
- ■ Don't know (6)

Q7 What is your annual household income?

- ■ less than $10,000 (1)
- ■ $10,001 to $20,000 (2)
- ■ $20,001 to $50,000 (3)
- ■ $50,001 to $75,000 (4)
- ■ $75,001 to $100,000 (5)
- ■ $100,001 to $150,000 (6)
- ■ $150,001 to $250,000 (7)
- ■ $250,001 to $350,000 (8)
- ■ more than $350,000 (9)

Q8 Based on your political views, would you consider yourself to be:

- ■ A Democrat (1)
- ■ A Republican (2)
- ■ An Independent (3)
- ■ Other (4)

* Click continue to finish your survey.You will receive your total final payment in the coming week. Thank you for participating!

| Pre-screen Survey | Experimental Survey | | | | Total |
| --- | --- | --- | --- | --- | --- |
| | Democrat | Republican | Independent | Other | |
| Republican | 39 | 182 | 16 | 1 | 238 |
| % of pre-screen | 16.4 | 76.5 | 6.7 | 0.4 | 100.0 |
| Democrat | 221 | 14 | 9 | 3 | 247 |
| % of pre-screen | 89.5 | 5.7 | 3.6 | 1.2 | 100.0 |

Table 9: Differences in partisanship between Elab pre-screen survey and experimental survey.

## 7.3 Switchers

This section analyses the behavior of those who declared different partisanship in the E-lab pre-screen survey compared to the experimental survey. This analysis is important because the E-lab pre-screen survey was not part of any experiment. Everybody answered the same set of questions in the same order and no payoffs were derived from doing so. In our experimental survey, on the other hand, we asked for partisanship after participants decided whether to trust and reciprocate conditional on their partner's partisan identity. Table 9 shows the numbers of those who switched. In general, the majority stuck to their pre-screen partisanship: 89% of Democrats (221 out of 247) and 76% of Republicans (182 out of 238).

Depending on the identity of the matching partner, individuals may feign partisanship affiliation following reprobate behavior. Table 10 shows the number of individuals who switched partisanship per treatment. Overall, participants who switch to the opposite party tend to do it when they are not matched with someone with the same partisan identity, although the difference is only significant for (pre-screened) Democrats. Precisely, none of the 100 Democrats switched to Republican, and only 4 did to either Independent or Other (those who switch account for 4% in this treatment, chi-squared p-value = 0.029) when matched with a Democrat. From the 99 Republicans matched with a Republican, 16 switched to Democrat and 6 to Independent or Other (those who switched account for 22% in this treatment, chi-squared p-value = 0.785).

| | Democrats (Pre-screen survey) | | | | | Republicans (Pre-screen survey) | | | |
|---|---|---|---|---|---|---|---|---|---|
| Treatment | Democrat | Republican | Independent/Other | Total | | Democrat | Republican | Independent/Other | Total |
| matched anyone | 85 | 10 | 5 | 100 | | 17 | 72 | 6 | 95 |
| % of total per treatment | 85 | 10 | 5 | 100 | | 17.9 | 75.8 | 6.3 | 100 |
| matched Democrat | 96 | 0 | 4 | 100 | | 6 | 33 | 5 | 44 |
| % of total per treatment | 96 | 0 | 4 | 100 | | 13.6 | 75 | 11.4 | 100 |
| matched Republican | 40 | 4 | 3 | 47 | | 16 | 77 | 6 | 99 |
| % of total per treatment | 85.1 | 8.5 | 6.4 | 100 | | 16.2 | 77.8 | 6.1 | 100 |
| Total | 221 | 14 | 12 | 247 | | 39 | 182 | 17 | 238 |

Table 10: Differences in partisanship per treatment between Elab pre-screen survey and experimental survey.

One important concern that arises is whether declaring different partisan identity is in part caused by their previous behavior and the partisan identity of the partner. For example, Republicans may want to deliver a statement that even Democrats trust Republicans. Our data is consistent with this: 16 of the 39 Republicans who switched to Democrats after the experiment were matched with Republicans. 81% of them (13 out of 16) trusted a Republican Player B, which is significantly larger than the 53% of the 77 who did not switch to Democrat (41 out of 77, chi-squared p-value = 0.039).[14] Another reason for the switch may be to feign partisanship when the participant did not trust. Consistent with this, we observe Republicans who switched to Democrat and were matched with an anonymous Player B (17 out of 95) trusted the anonymous partner significantly less often than those who did not switch (72 out of 95): 35% compared to 61%, chi-squared p-value = 0.054.[15] This seemingly strategic behavior is not particular to Republicans. Democrats may also feign partisan identity to make Republicans "look bad." Democrats who were matched with a Republican, for example, who switched to Republican (4 out of 47)

---

[14]Note that there were 6 Republicans who switched to either Independent or Other and were matched with a Republican. With these, we account for the 99 individuals who were matched with Republicans in the experimental survey (16+77+6=99).

[15]The remaining 6 (17+72+6=95) switched from declaring being Republican in the E-lab pre-screen survey to declaring that they were either Independent or Other in the experimental survey.

never trusted a Republican Player B, while those who did not switch (40 out of 47, the remaining 3 switched to either Independent or Other) did so 42% of the time. Although this difference is marginally significant at conventional levels (chi-squared p-value = 0.096), it suggests that individuals may be trying to feign partisan identity to undermine the other party's image.