RACIAL DIFFERENCES IN INEQUALITY AVERSION:
EVIDENCE FROM REAL WORLD RESPONDENTS IN THE ULTIMATUM GAME

John D. Griffin
David Nickerson
Abigail K. Wozniak

Racial Differences in Inequality Aversion: Evidence from Real World Respondents in the
Ultimatum Game
John D. Griffin, David Nickerson, and Abigail K. Wozniak
NBER Working Paper No. 17097
May 2011
JEL No. J15

ABSTRACT

The distinct historical and cultural experiences of American blacks and whites may influence whether
members of those groups perceive a particular exchange as fair. We investigate racial differences
in fairness standards using preferences for equal treatment in the ultimatum game, where responders
choose to allow a proposed division of a monetary amount or to block it. Although previous research
has studied group differences in the ultimatum game, no study has been able to examine these across
races in America. We use a sample of over 1600 blacks and whites drawn from the universe of registered
voters in three states and merged with information on neighborhood income and racial composition.
We experimentally vary proposed divisions as well as the implied race of the ultimatum game proposer.
We find no overall racial differences in acceptance rates or aversion to unequal divisions. However,
we uncover racial differences in the response to pecuniary returns conditional on inequality of the
division. This is driven by the lowest income group in our sample, which represents the 10th percentile
of the black income distribution. The racial differences are robust across gender and age groups. We
also find that blacks are more sensitive to unfair proposals from other blacks.

John D. Griffin
Department of Political Science
University of Notre Dame
South Bend, IN 46556
John.Griffin@nd.edu

David Nickerson
Department of Economics
University of Notre Dame
441 Flanner Hall
Notre Dame, IN 46556

Abigail K. Wozniak
Department of Economics
University of Notre Dame
441 Flanner Hall
South Bend, IN 46556
and NBER
a_wozniak@nd.edu

I. Introduction

The relationship between whites and African Americans in the United States has included many exchanges in which one group (virtually always African Americans) has been treated unfairly or even unjustly by the other. It is therefore possible that African Americans' historical and ongoing experiences with discrimination may make them more alert to the possibility of exploitation and more sensitive to deviations from fair treatment (Dawson 1995). In addition to this legacy of lopsided exposure to unfairness, the distinct cultural norms of American blacks and whites may dictate different standards of fair treatment across the two groups. In light of these differences, it is reasonable to ask whether blacks and whites respond differently in settings where fairness is a consideration. We use a large sample of blacks and whites drawn from a broad swath of the socio-economic distribution to examine this question. We focus on equal treatment as our standard for fairness and examine preferences for equal treatment – also called inequality aversion – in the classic setting of the ultimatum game.

The importance of fairness considerations in determining behavior in one-on-one and group interactions is now widely recognized by economists. Inequality aversion specifically has been shown to operate in a variety of important contexts, including bargaining in product and labor markets (Babcock and Laschever, 2003), decisions over redistribution (Luttmer 2001), bequests and charitable giving (Fong and Luttmer 2009), and political support for regulation (Lu, Scheve, and Slaughter, 2010). Examining racial differences in inequality aversion is important because it can help us understand whether different preferences for equal treatment might partially explain racial differences in behavior in these settings. For example, Bartels (2008, 133) finds that blacks espouse more egalitarian attitudes on a number of dimensions. Moreover, behavior in these settings may in turn influence relative outcomes. For example, fairness concerns may affect the response to wage offers, which may in turn affect relative wages. Members of groups with strong inequality aversion may be unwilling to accept very low relative wages in the labor market (even if those wages reflect underlying human capital differences) leading to lower relative

employment rates. Similarly, group preferences for equality can influence the features of social redistribution programs through the election process.

In this paper, we report the results of an experiment in which we use the ultimatum game (UG) to test for racial differences in inequality aversion between whites and African Americans in the United States. Preferences for inequality aversion are often studied using the UG. In the game, one participant divides a monetary award into two parts and a second participant either accepts or rejects their offered share of the division. Rejections of non-zero offers are thought to stem from preferences related to fairness, either from a preference for equal splits, also called inequality aversion (Bolton and Ockenfels 2000; Oosterbeek et al. 2004), or from a preference for punishing proposers who act selfishly (Falk, Fehr, and Fischbacher 2005). Our experimental approach enables us to isolate racial differences in responses to unequal treatment and allows us to answer several questions. First, does a minority group that has historically experienced a high level of discrimination and unfair treatment respond differently in the ultimatum game? Second, does this depend on neighborhood income levels or racial integration that respondents experience? Finally, does an unfair proposal elicit different responses among blacks when the proposer is black as opposed to white?

We conduct a survey experiment in which respondents are invited to play the UG by telephone.[1] Participants are selected from the universe of registered voters in three states with large black populations. The states we employ all record registrant race in the official state voter file, which also contains the registrant's home address and telephone number. Using the address data, we merge the voter file information with neighborhood level (Census block group) characteristics from the 2000 U.S. Census. We then draw a random sample stratified on race, block-level median income, and neighborhood racial diversity from the merged data. Balancing our sample on these dimensions allows us to compare the game behavior of comparable whites and African Americans, as well as to make comparisons along other socio-economic lines. This is particularly important in our setting because preferences for equality have been shown to be

---

[1] Others have studied inequality aversion using survey data (Atkinson 1970).

related to income, a characteristic which varies substantially across blacks and whites (Alesina and LaFerrara, 2005). Our final sample contains 1647 respondents.

In our study, all participants play the role of Responder. We focus on the Responder's decision to accept or reject the offer due to our interest in fairness preferences and inequality aversion in particular.[2] Survey Respondents are told that they have been selected to play a game with a proposer (who is hypothetical) who has made an offer to divide a given stakes amount. The stakes to be divided, the amount of the offer, and the name of the "Proposer" are all randomly assigned to respondents. We use commonly-occurring but racially distinct names among white and African American males to imply the race of the Proposer. Our decision to use a research design with this type of deception may raise concerns among experimental economists, and it was not taken lightly. We discuss and defend this decision when we explain the treatment in detail below.

We find that African Americans and whites have similar levels of aversion to unequal treatment defined as their response to a given share allocated by our hypothetical proposers. Two differences in the groups are apparent, however. First, African Americans but not whites are more likely to accept larger offers, after controlling for the share of the total sum being offered. This difference is confined to those with very low incomes. Second, the implied race of the proposer matters. African Americans facing proposers who are implied to be African American increase their acceptance rates to higher offered shares more rapidly than do African Americans facing proposers who are implied to be white. Conversely, blacks facing black proposers decrease their acceptance rates faster as the share falls when proposers are implied to be black.

Our paper makes three contributions. First, we add an important new dimension to the literature on group differences in economic decision-making. We provide the first analysis of racial differences in fairness preferences in a representative U.S. population. Over the last decade economists have used laboratory

---

[2] Non-trivial offers by the first player have been attributed either to strategic or altruistic behavior (Kravitz and Gunto 1992). Altruism is somewhat different from fairness, and we believe the behavior of the second player more directly reflects fairness preferences (Oosterbeek et al 2004.)

experiments to uncover striking and potentially important differences across groups in how they behave in a variety of experimental settings. Differences along lines of gender and ethnicity (or nationality) in particular have received considerable attention. Examples from this literature are many (On gender: Bolton and Katok, 1995; Eckel and Grossman 1998; Solnick 2001; Andreoni and Vesterlund, 2001; Gneezy, Niederle, and Rustichini 2003; Niederle and Vesterlund, 2007. On nationality/ethnicity: Roth et al 1991; Henrich 2000; Chuah et al 2007; Ferraro and Cummings 2007; Ferhstman and Gneezy 2001; Chen and Tang 2009). [3] However, the study of *racial* differences in economic decision-making has so far been quite limited. Two important exceptions are Benjamin, Choi and Strickland (2009), who examine the effects of race priming on subjects' expressed preferences for risk aversion in a university-based sample, and Fong and Luttmer (2009) who use a representative sample to analyze racial differences in charitable giving. Ayers and Siegelman (1995) find that the racial composition of bargaining pairs affects bargaining behavior. The only racial comparison of UG behavior we have been able to locate is ancillary to what is otherwise a comparison of gender in UG play (Eckel and Grossman 2001)[4].

Our second contribution is to implement a sampling methodology that allows us to answer these questions outside the laboratory and in the field. We introduce a new (to economics) technique for obtaining representative samples of experimental subjects—the use of voter file demographic and contact information combined with Census block group data on neighborhood characteristics. Using voter registration files as a sampling frame has been used outside economics to construct representative samples, most widely in health surveys (Adimora et al 2001 for example) but also by political scientists (Gerber and Green 2000; Nickerson 2008; Gerber, Karlan and Bergan 2009) and pollsters.[5]

Our sampling methodology has several advantages over others in the experimental economics literature. First, we are able to construct a data set that is much more representative of the general population than is typically found in laboratory-based experimental studies. Studies of social preferences

---

[3] There are also some studies of age differences (Murnighan and Saxon 2001).
[4] Eckel and Grossman (1998) briefly note that large differences were detected in how students at a predominantly black college played the UG compared to the students from a predominantly white college.
[5] See the discussion on the Field Poll's website: http://www.field.com/fieldpoll/methods.html.

(albeit not the UG) that have compared the behavior of students and non-students have identified

significant differences between the groups (Gordon, Slade, and Schmidt 1986; Fehr and List 2004;

Carpenter, Connolly, and Myers 2008). As one study put it, "problems exist in replicating with nonstudent

subjects behavioral phenomena observed in student samples" (Gordon et al., 1986). Importantly, our

sample contains large numbers of respondents from different parts of the black income distribution. One-

third of our respondents come from Census block groups where the median household income is at the

tenth percentile of the black income distribution. To our knowledge, no other experimental study in a

developed country has such a substantial representation of the poor. Second, respondents to our survey did

not volunteer in response to an advertisement for experiment participants, a form of sample self-selection

that can also detract from representativeness (Doty and Silverthorne, 1975).[6] Finally, our approach allows us

to collect detailed information on a respondent's neighborhood, something that is typically missing from

databases of voluntary, non-student participants and which cannot be reliably collected in a survey format.[7]

The neighborhood level information also allows us to balance our sample according to the types of

environments respondents live in, something that is not possible with other data sets. If neighborhoods are

important determinants or correlates of preferences, then this type of neighborhood-level information can

provide both a crucial set of controls and a useful dimension for sampling.

A final contribution is that our experimental design allows us to consider different notions of

fairness. Specifically, we expand the model developed in Fehr and Schmidt (1999) to incorporate two types

of inequality, which we call relative and absolute. The Fehr and Schmidt model concerns what we call

absolute inequality, where disparities in the amounts paid to the two players generate disutility. In our

model, relative inequality arises when the share of the division deviates from a 50-50 split. Our experiment

varies offer size and stakes randomly across players, so we can determine whether relative and absolute

---

[6] Two recent papers also examine the issue of selection via volunteering to enter a survey pool or experiment. Abraham et al (2009) studies selection into the American Time Use Survey from the Current Population Survey sample, and Falk et al (2011) study selection into lab experiments from college student populations. Interestingly, they come to opposing conclusions about the severity of selection.
[7] Examples of these databases include Knowledge Networks and Time Sharing Experiments in the Social Sciences (TESS).

inequality play separate roles in the acceptance decision. This is not an entirely new question, but to our knowledge this is the first time competing measures of inequality have been tested outside the laboratory using techniques that allow us to experimentally vary both the amount to be divided and the offer size. Oosterbeek et al (2004) consider a similar question using meta-analysis, and a seminal paper on the ultimatum game, Guth et al. (1982) examines behavior across a range of stakes sizes.

A shortcoming of our study is that we can say little about the relationship between our subjects' behavior in the UG and in real world settings. The potential for this connection is a major motivation for our study. However, because the feasibility of conducting an experiment by phone with our sample was previously untested, we elected to keep the experiment extremely simple. We therefore omit more complicated features that would allow us to directly link our findings to racial differences in behavior in other settings.

II. Choice Model

The participants in our study are faced with a choice problem. Specifically, they are offered the possibility of absolute financial gain accompanied by possibly unequal gains accruing to the (hypothetical) proposer. Fehr and Schmidt (1999; hereafter FS) model the utility maximization problem faced by individuals in two-person decision games where player i must decide whether to accept player j's offer as follows:

(1) $u_i(x) = x_i - \alpha_i max|x_j - x_i, 0| - \beta_i max|x_i - x_j, 0|$

If i accepts j's offer, she receives x_i but incurs a utility penalty if $x_i \neq x_j$. FS call this inequality aversion, and their model allows for the utility penalties to differ according to whether i receives more or less than j in the proposed split. The FS formulation also assumes that the utility penalty is a linear function of the absolute difference between what i and j receive.

The choice model we estimate differs from FS in two key ways. First, since there are no instances of experimental splits where the respondent is offered more than an even split, we do not model asymmetric utility losses to inequality that favors the respondent. We do not see this as a limitation of our study, since splits favoring the respondent rarely occur in experimental UG play.[8]

Second, since we experimentally vary both the stakes and the offer among our respondents, we can examine the utility penalty to *relative* inequality in addition to absolute inequality. We embed straightforward concepts of relative and absolute inequality into a simple utility function that is a direct extension of FS.

We assume that an accepted offer in game g generates the following utility for our respondents:

$$(2) \quad v_i(g) = o_{gi} - \gamma_{1i}(s_{gi} - 2o_{gi}) - \gamma_{2i}(0.5 - h_{gi})$$

The vector g contains two elements with information on the game variables: o_{gi} the size of the proposed offer; s_{gi}, the stakes or total to be split between i and j with i receiving o_{gi}. Define h_{gi} to be the share of the stakes o_{gi} represents. o_{gi} and s_{gi} are subject to the constraint that $o_{gi} \leq s_{gi} - o_{gi}$. Following FS, the first term in (2) is simply the utility gained from the offer, while the second term represents the penalty from receiving an unequal distribution of funds. The third term reflects a utility penalty due to *relative* inequality in the split. We express this as deviation from a focal point of 0.5, which represents an even split.[9]

The individual's choice problem is then to accept or reject in order to maximize utility.[10] As modeled, respondents accept an offer if the utility of the financial gain exceeds the penalties associated with the two inequality dimensions.

---

[8] UG offers are usually center around 40% of the stakes, and on average 16% of non-zero offers are rejected, typically when they fall below 20% of the total sum. See Roth (1995) and Camerer (2003) for reviews of the literature on the ultimatum game. For a meta-analysis of 75 UG experiments in 37 studies see Oosterbeek, Sloof, and Kuilen (2004).

[9] The terms in this function can be collected and rearranged. (2) is therefore not a unique way of writing the sum. However, we postulate that the included quantities are the relevant determinants of utility. Written in this way, there is a direct path to multinomial logit estimation of the parameters in (2) given the exogenous variation induced by our experiment.

[10] Alternatively, we could have outlined a model following Bolton and Ockenfels (2000). Their ERC model postulates two elements in the utility function: pecuniary return (equivalent to our *offer*) and the respondent's relative share of the payoff. We could construct the latter from our offer and stakes variables and substitute those for the second and third terms in (2). While this model would be equally interesting, it would not likely alter the substance of our results, since the most stable racial differences we estimate pertain to how changes in the offer size affect behavior, not offer relative to other variables. The individual's choice

Individuals choose $y \in \{0,1\}$ to maximize u subject to some error:

$$(3) \quad u(y; g) = y * v_i(g) + \varepsilon_{iy}$$

We assume individuals choose optimally given v_i, their individual characteristics X, and an extreme-value distributed error, so that $\Pr[y = 1|x] = Pr[u(1; g) \geq u(0; g)|X]$.

We can therefore estimate the following in a standard logit framework:

$$(4) \quad y_{gi}^* = \alpha_0 + \alpha_1 o_i + \alpha_2 (s_{gi} - 2o_i) + \alpha_3 h_i + \mu_{gi} = \alpha' x_{gi} + \mu_{gi}$$

where y_{gi}^* is the unobserved latent variable driving the observed choice y_{gi}, and μ_{gi} is extreme value distributed. Control terms for individual characteristics are omitted for clarity. If we also assume that $\alpha_{ni} = \alpha_{nj} \; \forall n, i, j$, then we can estimate the above equation using standard logit.

One of our main questions of interest is whether blacks and whites respond differently when faced with various combinations of offer, absolute inequality, and relative inequality. Therefore we would also like to estimate the following:

$$(5) \quad y_{gi}^* = \beta_1 black_i + \beta_2 o_{gi} + \beta_3 absolute_{gi} + \beta_4 relative_{gi} + \beta_5 black_i * o_{gi} + \beta_6 black_i *$$
$$absolute_{gi} + \beta_7 black_i * relative_{gi} + \mu_{ig}$$

The estimate of β_1 will tell us whether blacks are more likely than whites generally to accept an offer in the ultimatum game. The interaction coefficients will tell us whether specific game components affect black decisions differently from whites.

The standard logit approach has several advantages. First, we can readily test whether the game components, on average, affect the choices of blacks and whites statistically differently. Second, the likelihood function is computationally simple to optimize.

---

problem remains the same, and estimation of the utility function parameters via multinomial logit would still go through in this case.

A drawback to the standard logit approach is that it assumes that the error term is uncorrelated with the control variables listed above, at least conditional on the inclusion of other individual covariates. This is equivalent to the assumption that $\beta_{ni} = \beta_{nj} \ \forall n, i, j$ in (5). This assumption can fail if there is individual level heterogeneity in the utility function parameters. In that case, we would prefer to estimate the following:

(6) $y_{gi}^* = \beta_i' x_{gi} + \mu_{gi}$

This is possible using the mixed logit approach developed in Revelt and Train (1998) and implemented in Cappelen et al (2007), Barros et al (2005), and Detang-Dessendre et al (2008). The mixed logit approach allows us to estimate the distribution of β_i in our experimental sample after assuming that the underlying individual parameters are drawn from a specific family of distributions.

Specifically, we can estimate the following:

(7) $y_{gi}^* = \beta_i' x_{gi} + \mu_{gi} = \varphi' x_{gi} + \eta_i' x_{gi} + \mu_{gi}$

where the η_i term represents unobserved heterogeneity that depends on included covariates.

The mixed logit approach is appealing in the choice setting because it relaxes the assumption that the relationship between the game components and acceptance choices is identical across individuals up to a location parameter. It also relaxes the common Independence of Irrelevant Alternatives (IIA) assumption, but that is less of a concern here since our participants only face two alternatives.

The main drawback to the mixed logit approach is that the likelihood function is computationally difficult to optimize. Solutions require use of simulation methods to approximate the integrals over the individual utility functions. Also, statistical testing for differences between estimated parameter distributions in the mixed logit framework is not well-developed. Therefore it is not an ideal model for testing one of our central questions of interest – whether blacks and whites exhibit different degrees of inequality aversion.

Finally, it is not clear that unobserved heterogeneity is even a major concern in our setting, where the treatment is randomly assigned to participants instead of respondents selecting into the choice setting

endogenously. For all these reasons, we rely on the standard logit as our main empirical specification. We present a limited set of mixed logit estimates as a robustness check. Ultimately, we view our empirical analysis as fitting best in the category of experiments that test a single model, in the framework of Card, Della Vigna, and Malmendier (forthcoming). We test whether the nominal amount of an offer, as well as its relative inequality features, have differential impacts across blacks and whites. We do not test the Bolton and Ockenfels (2000) framework against Fehr and Schmidt (1999) – which would be equivalent to Card et al's competing models experimental analyses, nor do we directly estimate structural parameters in our main specifications.

III. Experimental Design

The experiment consisted of calling subjects on the telephone and inviting them to play an ultimatum game. The rules were explained and the name and home city of a hypothetical proposer was provided to the subject. The offer was presented to the subject, who could then accept or reject the offer. The subject was then asked to identify the smallest offer the subject would have accepted. The call ended by confirming the contact information of the subject.

The treatment was administered by Eastern Research Services, a professional survey firm. Calls were completed between December 10, 2008, and January 26, 2009, and monitored remotely (via telephone) by the authors and directly by supervisors in the call centers.[11] The callers were carefully trained regarding the rules of the game and the nature of the research.[12] In particular, the callers were trained to read the rules of the game slowly and answer questions about the game.[13] The script used during the course of calling is

---

[11] Nearly all of the calls took place in the evening. Calls were placed during working hours on weekends and a handful of other days in order to increase response rates. Numbers with no answer initially were also retried, and in some specifications we control for the number of times a number was called (we find this makes no difference to our results).

[12] Given the generally positive response from subjects contacted by *Eastern Research Services*, the callers reported very high job satisfaction and found the study a welcome change from consumer research polling. We therefore believe that although this firm primarily conducts consumer research polling, their skills translated well to academic polling.

[13] Two of the authors personally monitored several calls. Our impression was that most subjects understood the rules of the game, although this may not have always been the case, as we discuss below.. On the one hand, this naïve first encounter with the ultimatum game is a quantity of interest. On the other hand it would have been interesting to see how respondents played after a few trial runs.

included in Appendix A.[14]   The script asks the subject if she would like to play a game for research

purposes, describes the financial incentives, explains the rules of the game, confirms that the subject

understands the rules before proceeding, introduces the fictitious proposer, informs her of the proposer's

offer, and records whether she accepts or rejects the offer. The caller then asks the subject for her minimally

acceptable offer and verifies the subject's contact information. A debriefing letter describing the purpose of

the study and the nature of the deception was mailed to participating subjects the day after the call took

place, along with any winnings from accepting an offer.

The call script varied two factors in the treatment. The first factor varied in the script was the terms

of the ultimatum. It is important to our study that both the stakes of the ultimatum game played and the

offer made by the opponent varied. Most prior work on inequality aversion keeps the stakes of the game

constant and varies the offers, assuming that the absolute differences in the stakes will not materially affect a

subject's utility so inequality aversion must be driving the results. Our goal was to randomly vary the stakes

of the game from which particular offers were made to directly manipulate the inequality of the offer. For

instance, an offer of $2 could be made in a $5, $10, $20, $50, or $100 game. To ensure that the results were

not idiosyncratic to the amount of the offer, several different dollar values of offers were tested ($1, $2, $5,

and $10). The frequency distribution of the offer and stake combinations used in the game is presented in

Appendix 2. This distribution concentrated subjects into offer/stakes cells where high variance in

acceptance rates was anticipated and away from cells where it was anticipated that acceptance rates would be

very low or high. Since subjects in each stratum were randomly assigned to each cell, the different

probabilities of receiving a particular offer or stakes does not bias the results in the slightest because subjects

were equally likely to be assigned to each of the conditions.

The second factor varied in the script was the implied race of the proposer. To the extent that the

names are associated with a particular race, we can manipulate a subject's perception of their opponent's

---

[14] To facilitate ease of contact the sample was limited to households containing three or fewer registered voters (excluding roughly 1% of the sample).

race. As explained above, differences between the races in the probability of accepting a given offer may depend on the race of the proposer. In the laboratory, race can be manipulated by purposefully pairing subjects or using confederates. This visual manipulation is not possible in a telephone-based study, so we instead rely on racially polarized names. Appendix 3 provides details of the name construction as well as evidence that the selected names indeed connoted different races. After agreeing to play the game, the subject was told the full name of the fictitious proposer. The proposer's first name was then repeated four more times over the course of the script (see Appendix 1). Upon the final statement of the proposer's name, the amount each player would receive from the proposed split was explicitly stated in order to make the connection between the proposer's name and the inequality of the split very direct.

We want to give particular attention to one aspect of our experimental design because it often raises questions among experimental economists. Our design deceives respondents in telling them that they are paired with a proposer who in fact is hypothetical. While we go to no particular lengths to convince respondents that the proposer is a real person, this is a likely inference on their part.[15] In fact, the design relies on respondents making inferences about the race and gender of the proposer. This type of deception is typically eschewed by experimental economists, largely although not exclusively on the grounds that it can contaminate subject pools by making potential participants in future studies question the truthfulness of experimental instructions.[16] Nevertheless, a number of studies have deceived subjects to some degree (Bertrand and Mullainathan, 2004; audit studies, as summarized in Fix and Turner 1998; Karlan and Zinman 2009). Not coincidentally, these studies all take place in the field, where it sometimes becomes prohibitive to introduce the kind of control that makes it easier to avoid deception in the laboratory.

The objectives of our experiment posed a considerable challenge to developing a non- deceptive design. Because we wanted to reach respondents who rarely appear in laboratory experiments, we needed to

---

[15] Blount (1995) points out that behavior in the ultimatum game is strongly influenced by the respondent's beliefs about whether (a) the proposer is human and therefore making offers intentionally and (b) whether the proposer benefits from her decision.
[16] Deception of the form in which the experimental instructions contain a non-truth or in which the likely inference from the instructions is untrue is typically avoided. Deception in the form of omission is common. For example, experimenters typically provide very limited information about the purpose of an experiment, often to the point that inferences subjects are expected to make about the nature of the experiment are very different from the true objectives. Jamison et al (2008) discuss the origins of the deception prohibition and provide evidence of its effects on the public good of undergraduate subject pools.

take the experiment into the field. While several papers investigate ultimatum game behavior in the field, these studies have taken place in developing countries where lower costs and more common community gathering places dramatically lower the barriers to bringing subjects together in the field to play the game (Cameron 1999; Henrich 2000; Henrich et al 2001), and as this implies they have used convenience sampling procedures . The logistical challenges to bringing together the subjects in our sample are formidable but might have been overcome. One might recruit a group of subjects as proposers and go door-to-door in sampled neighborhoods or use a design like that in Berger and Pope (2010), where proposals are made in a first round and then matched by experimenters to responders in a second round. The former design would either sacrifice the geographic representativeness of our sample or be prohibitively expensive. The latter design generates a real set of proposals that is manipulated by experimenters to randomize the offers along certain dimensions across respondents; experimental instructions reflect this. Yet our objectives pose problems for even this design. First, we wanted to randomize the racial connotations of proposer names in order to signal differences in race without overtly mentioning race as a dimension of interest. Benjamin et al (2010) show that race-related framing can have important effects on respondent behavior, and we felt that it was important to keep explicit mention of race out of the script. If we developed a pool of proposals from real participants, we would have to disclose a participant's name to his responder in the second stage.[17] Even assuming that disclosure of participant identity was approved by our IRB, we find this an unappealing design element. It would certainly select a certain type of subject who is unconcerned about disclosing his participation and behavior. Second, we wanted to explore the effects of offer size independently of the share represented by the offer. This requires a pool of offers that are outside the typical range, since most ultimatum game proposers offer roughly one-third to one-half of the stakes (Oosterbeek et al. 2004).

---

[17] This is also true of the method that takes proposers door-to-door.

Moreover, these would of course have to come from proposers with racially distinct names who were willing to have their behavior disclosed to their respondents.[18]

Harrison and List (2004) describe the ideal experiment as featuring no deception.[19] While we agree that this is ideal, our objectives posed formidable obstacles to meeting this ideal. Moreover, we executed our study in a manner that mitigates the concerns experimental economists have with deception. First, it is not unreasonable to think that the public scrutinizes "cold calls", which are often used for sales pitches, political campaigning, and other solicitations. Since subjects are conditioned to consider a caller's motivations, a norm of non-deception within economics is unlikely to alter the behavior of subjects within a game. Second, while 1600 subjects is large by standards of experimental papers, it is a very small proportion of the universe. It is therefore unlikely that contamination is a major concern for future researchers in this pool. Finally, we debriefed subjects about the deception in a letter following the experiment. A copy of the debriefing letter appears in Appendix 4.

IV. Sample

Experimental games frequently take place in laboratories and rely on undergraduate subject populations.[20] Undergraduate samples pose a particular problem for studying racial differences in inequality aversion because blacks are underrepresented in college populations. College samples also draw from a higher stratum of the socio-economic spectrum where individuals may experience inequality differently from older and less educated persons. Volunteer samples from the wider population suffer from similar problems. As Doty and Silverthorne (1975) note, volunteers in human research "typically have more education, higher occupational status, earlier birth position, lower chronological age, higher need for

---

[18] We also note that our technique of constructing a full name from racially polarized first and last names (as described in Appendix 3) means that actually very few individuals have the names we used in our survey. Power calculations under various assumptions of racial name polarization, response rates, and income stratification are available from the authors upon request.
[19] They write that an experiment is ideal "…in the sense that one is able to observe a subject in a controlled setting but where the subject does not perceive any of the controls as being unnatural and there is no deception being practiced."
[20] See Levitt and List (2009) for a discussion of issues related to typical experimental samples.

approval and lower authoritarianism than non-volunteers" To address these concerns, we sought to randomly sample a broad a spectrum of subjects from the general population.

We use a sampling frame new to the economics literature: voter registration files. Eligible citizens must be registered to vote in 49 states[21], a population that consists of 70% of citizens over the age of 18 (Census Bureau 2009, Table 2-1). According to the Current Population Survey November Supplement (2008), on average, registered voters tend to be slightly older (45 versus 42 years of age), more likely to be married (46% versus 42%) and female (51% versus 48%), and less likely to be non-white (14% versus 17%) and Hispanic (10% versus 13%) than unregistered citizens.[22] Voter files typically include identifying information such as full name, address, gender, date of birth, and whether the person voted in recent elections. Voter files have been used as the sampling frame for studies of voter mobilization (e.g., Gerber and Green 2000; Nickerson 2008) and where voter turnout is a dependent variable of interest (e.g., Gerber, Karlan and Bergan 2009). Sampling registered voters has been found to be slightly superior to random-digit dialing with regards to election forecasting (Green and Gerber 2006). Registered voters are representative of people who participate in politics, and similar to the population as a whole. In the U.S. as a whole, 64.9% of all individuals 18 and over and 71% of citizens 18 and over are registered to vote. Registration rates are slightly higher for whites (73.5% for white, non-Hispanics) than for blacks (69.7%).[23] The registration rates are even closer in states in our sample: in two of the three states, black registration slightly exceeds white registration, and in the third the gap is smaller than the national average (U.S. Census Bureau, 2008).

Our sample comes from randomly selected black and white registered voters with telephone numbers in three states: Georgia, North Carolina, and South Carolina. To be included in the study a state needed to meet four criteria. First, the state's voter file needed to include race as one of the fields

---

[21] North Dakota does not require citizens to register to vote.
[22] Registered and unregistered persons also share the same median categories of education and income in the CPS November Supplement. That said, the overall distribution of education and income tends to be shifted higher for registered voters compared to unregistered citizens.
[23] All figures from U.S. Census Bureau tabulations of 2008 Current Population Survey data.

consistently collected (excluding Virginia and all non-southern states)[24]. Second, the state's voter file needed

to be relatively easy to obtain (excluding Alabama, Florida, and Mississippi). Third, the population could not

have been subject to a recent diaspora (excluding Louisiana, due to Hurricane Katrina). Fourth, variation in

electoral support for President Obama was desired as a proxy for progressive political culture (i.e., Obama

won North Carolina in 2008, narrowly lost Georgia, and lost South Carolina by a wide margin).[25] Finally,

geographically proximate states were selected in order to keep long run determinants of state development

such as weather and immigration patterns as constant as possible. Lists of registered voters were used as the

sampling frame because the process of assembling the sample was less expensive than random digit dialing

with screening questions for race and income. The statewide voter files for all three states were obtained in

early October 2008. The voter files were matched against consumer data files maintained by InfoUSA, one

of the largest consumer data firms in the world, to append reliable telephone numbers and update address

information.[26] After the address information was cleaned, we geocoded the observations and appended

data on Census block group characteristics. Thus, the data used in this experiment come from the voter file,

census block group data, and answers provided during the experiment itself.

Our use of Census block group information makes our sampling approach different from that of

using a representative population sample from a maintained database, like those available through

KnowledgeNetworks and TESS. It allows us to sample respondents *based on their neighborhood characteristics.*

This in turn allows us to stratify our sampling to achieve sample balance along several important

dimensions. This feature of our sampling approach is an important contribution of our study.

Neighborhood level information is not part of the basic demographic variables collected in maintained

databases. Indeed, detailed geographic information is relatively uncommon in microdata. It is also unlikely

that survey respondents can be relied upon to provide this information themselves, since individuals are

---

[24] The Voting Rights Act required some states with a history of disenfranchising black voters to collect this information. For a history, see Canon (1999).

[25] This left us with five potential states. We limited ourselves to three in order to have sufficient sample size within each state to allow for potential analysis at the sub-state level. This excluded Tennessee and Arkansas.

[26] The particular fuzzy matching algorithm used by InfoUSA is proprietary. InfoUSA's data comes from a combination of public sources, shared client data, purchased data, and independently collected data.

unlikely to know such information as their Census block group's median household income. Moreover, we estimate that our sample is considerably larger than the sample that could be drawn using similar restrictions from the KnowledgeNetworks panel.[27]

In order to focus on racial differences in inequality aversion, we selected a sample stratified on several characteristics that differ across blacks and whites but that may also condition behavior in the ultimatum game and reactions to unequal treatment in general. The first of these is income. Given persistent differences in income between blacks and whites, it is important to keep the income of the participants in the experiment as consistent as possible.[28] Thus, our first strata dimension is Census block group median income. To balance income across blacks and whites while providing variance in income, black and white subjects were drawn only from census block groups falling within the following three bands of income based on median household income among blacks for each state: a) 10th percentile to 10th percentile plus $2,000; b) the median plus or minus $1,000; and c) $2,000 below the 90th percentile to the 90th percentile. For example, whites and blacks sampled in South Carolina resided in census block groups with median household incomes between $21 – 23,000, $30 – 32,000, and $48 – 50,000. Despite economic segregation in housing and the Census Bureau's efforts to draw boundaries to account for well-defined neighborhoods and increased homogeneity, there will be variance in income within neighborhoods so the average household income may still differ across blacks and whites in our sample. However, we can be certain that blacks and whites were selected from *neighborhoods* with similar average socio-economic statuses.

The racial diversity of a subject's neighborhood may also moderate her response to the treatment stimulus, so we also stratify on the racial diversity of a respondent's neighborhood. The racial mix of each neighborhood was determined by looking at the percentage of registered voters in each street name – voting precinct group in the major racial categories. We used the street-voting precinct as the neighborhood in this instance since it allowed us to more precisely characterize a subject's true neighborhood. Census block

---

[27] Conversations with KnowledgeNetworks indicated that KN could recruit roughly 400 African-American participants from our three target states. However, once we limited that sample to our neighborhood characteristic strata, the yield falls to less than 100.
[28] In 2008, the median household income for non-Hispanic whites was $55,530 compared to $34,218 for non-Hispanic blacks (DeNavas-Walt, Proctor, and Smith 2009).

groups are larger and may contain distinct neighborhoods or sub-neighborhoods. Drawing loosely on the neighborhood tipping point literature (Card, Mas, and Rothstein 2008), we then categorized neighborhoods as being relatively homogeneous (0-20% and 80-100% black), predominantly white (20-40% black), evenly balanced (40-60% black), or predominantly black (60-80% black). The goal of the categories was to keep the neighborhood context as constant as possible across black and white respondents.

The subject pool was then stratified based on state of residence, the race of the subject (white or black), the three income categories (low, middle, and high for blacks in the state), and the levels of neighborhood diversity (homogeneous, predominantly white, evenly divided, and predominantly black). In all, there are 24 strata in the experiment, with balanced representation from each of the three states within each strata. Subjects within each strata were then randomly assigned to the three facets of the treatment.

Table 1 shows descriptive statistics for our final sample of survey respondents. The sample differs substantially from typical undergraduate populations. The top panel of Table 1 shows that not only is the sample racially diverse, but our respondents are also considerably older. The subjects also come from diverse neighborhoods socio-economically, as the neighborhood levels of education (17% college degree and a median of 11.7 years of education) and income (median household income $35,500 with 16% poverty) from the appended Census block group data suggest. Most importantly for our purposes, the blacks and whites drawn for the experiment are comparable in nearly every measurable capacity. The only statistically significant difference between the black and white variable means in Table 1 is in the percentage of blacks an individual's neighborhood – a difference due to our strategy of stratifying on neighborhood racial composition and using racially homogenous neighborhoods as one strata.

In Tables 2a and 2b we examine how our final sample of respondents compares to the sample of potential respondents the survey firm attempted to contact. Although blacks and whites who participated in the survey are balanced on observable characteristics, participants may differ from potential subjects due to non-random non-response. We have information on the pool of potential subjects the survey firm attempted to contact, so we can compare characteristics of our final sample to those of the wider potential

subject population. Non-response can occur at two points: first, respondents may refuse to answer the phone when contacted, or second, they may refuse to participate in the survey after answering the telephone. We have information on characteristics of all three groups: all potential subjects, those who answered the telephone when contacted by our survey firm but refused to participate, and participants.

The survey firm attempted to dial 5397 valid telephone numbers during the calling period and completed experiments with 1650 subjects for a total response rate of 31%. We later dropped three respondents with invalid phone numbers. Of those subjects who answered the telephone, the refusal rate was only 19% (i.e., 81% of the people took the survey once on the telephone). These completion and refusal rates are very good compared to other public opinion surveys and yielded a broad cross section of the populace.

Unsurprisingly, individuals who chose to answer the phone differ somewhat from the broader pool of potential subjects. This is shown in Table 2, panel A. People answering the telephone were more likely to be white, female, and older, and live in neighborhoods with less income and education. That said, the differences between the ideal sample and the people contacted are not large (e.g., 2 percentage point difference in gender, $2000 in income). Since treatment conditions were randomly assigned, the offers, stakes, and opponent names provided to subjects did not differ between contacted and uncontacted individuals.

Conditional on answering the telephone, there were very few differences between the people participating in the experiment and those refusing to participate, as shown in Table 2, panel B. Compliers and non-compliers were similar with regards to race, age, education, and income. The only statistically significant difference between compliers and non-compliers was that women were more likely to participate in the experiment. Thus, the population of subjects participating in the experiment is fairly representative of black and white registered voters in the three states for the income categories and neighborhood types selected.

V. Results

A. Randomization Checks

Before proceeding to the analysis, it is important to verify that the treatments (names, stakes, and offers) were balanced across subjects—in other words, that our randomization worked properly. We verify this in Table 3, in which we present the estimated mean and standard error of each treatment variable by the 24 strata defined by income, race, and neighborhood diversity. Looking down the columns of means, there is a high degree of consistency in a treatment variable's mean across strata groups. Importantly, the within-stratum mean is typically within one standard error of the total sample mean and in all cases within two. No systematic differences are observed for any of the treatments. We conclude that the randomization procedure was successful, and therefore differences in acceptance rates can be attributed to the treatments rather than the characteristics of the subjects. However, for completeness we present results with controls for the strata and for observable characteristics.

B. Graphical analysis

We begin by presenting a graphical version of our basic analysis in Figures 1(a)-(f). The figures reflect acceptance rates conditional on respondent race and offer size both for the experiment overall and separately by assigned stakes size. The vertical bars represent 95% confidence intervals around the mean acceptance rates.

Figure 1a reports acceptance rates by race and offer averaged over all stakes levels. Acceptance rates are non-trivial but less than one at all offer levels in our experiment. This is reassuring, since it suggests that our stakes-offer combinations span a range over which there is potential for changes in stakes or offer size to lead to changes in respondent behavior. Figure 1a also shows that whites are somewhat more likely than blacks to accept small offers (although the difference is not significant). This relationship is reversed for larger offers (and the difference becomes significant). We seek to better understand this reversal below and in our conclusion. For now, we simply point out that this reversal implies neither race is likely to have a general propensity toward higher acceptance rates. We demonstrate this more formally in our logit analysis.

Figures 1(b) through 1(f) show acceptance rates disaggregated across the five stakes levels in our experiment. Within these sub-experiments, consistent with prior studies acceptance rates generally rise with offer size, and again are non-trivial but less than one.[29] In instances where respondents are offered an even split, roughly 50% accept. This is true across all three stakes levels (5, 10, and 20) in which this split was offered. The figures also show that the black-white reversal in relative acceptance rates generally occurs between $2 and $5 offers, regardless of stakes size.[30]

In Figure 2(a), we plot acceptance rates by race over the share that the proposed split represents. Figures 2(b) through 2(e) plot these separately by offer amount, allowing us to observe how varying the stakes affects the acceptance rate for a given offer. It is clear from the figures that acceptance rates increase with offered share for both races. At higher offers - $5 and $10 - this increase is monotonic and blacks have uniformly higher acceptance rates than whites. For small offers, blacks and whites exhibit similar acceptance rates but there is some non-monotonicity in responses to these offers when they represent very small shares. These figures foreshadow another of our results, namely that blacks and whites differ in their responses to offer size, conditional on the share that the offer represents.

C. Choice Model Estimation

i. Analysis of Choice Model Results

We use standard logit models, as discussed in Section IV, to investigate the determinants of acceptance more formally. We begin by estimating baseline models on our complete sample of respondents. Specifically, we estimate Equations (4) and (5) where *black* is an indicator variable for whether the respondent indicated that she is African American on the voter registration form.[31] Recall that we obtain

---

[29] Interestingly, there are several instances in which acceptance rates decrease when the offer rises from $1 to $2, although the change is not significant in our sample. This is reminiscent of results of studies on the non-linear effects of monetary incentives on performance, as discussed in Gneezy and Rustichini (2000).

[30] The $50 stakes sub-experiment is the only exception to this.

[31] Racial coding was also confirmed by data collected by the consumer data firm where available. Unsurprisingly, people who check "Black" on voter registration forms are also likely to select "Black" in other outlets.

this information from the voter file data, rather than from the respondent directly, so the respondent's race is not being explicitly cued in the experiment.

The *relative* identity introduces some subtlety into the analysis. An artifact of the construction of *relative* is that *offer* can be interpreted as a function of *absolute* and *relative*. This makes interpreting the coefficient on *offer* in model 1(b) somewhat awkward. In light of this, we estimate two specifications for each of our models of interest. Model 1(a) includes only *offer* and *relative* from among the assigned game variables, so no game variable in this specification is a function of any other two variables. The coefficients on *offer* and *relative* estimated by (1a) are cumulative of effects due to their covariance with *absolute*, but their interpretation is more transparent. We then show that these effects are robust to controlling for *absolute* by estimating (1b). Omitting *absolute* is not the only way to address this problem. However, theory and previous experimental work predict an important role for *relative*, and Figures 1(a) to (f) suggest racial differences in the response to *offer*. In order of interest then, *absolute* seems the natural candidate for exclusion.

We estimate versions of the equation (5) specification that exclude interactions of the assigned game variables with *black*. By omitting the race interactions, we can more formally assess whether blacks are more or less likely to accept an ultimatum offer on average across our games. We also document the average effects of changes in our game variables on acceptance decisions. We then address questions about whether blacks and whites respond differently to changes in specific ultimatum game variables.

We also run our models with additional control variables. In principle, our randomization strategy should balance respondents on observed and unobserved characteristics, so no additional covariates need be included for consistent estimation of the parameters in equation (5). Nevertheless, we verify that the estimates obtained from the "no controls" specifications are robust to the inclusion of controls for observed characteristics. Our second specification adds controls for a respondent's Census block group characteristics. These are listed in Table 1. Their inclusion will control for differences across our strata in

tract-level aggregate characteristics other than median household income and racial diversity.[32] A third specification adds controls for "game conditions," which includes the length of the interview in minutes, the time of day the interview was conducted (which embeds the date and is not listed in the table), and the number of times the survey firm attempted to reach the respondent. If for some reason respondents across race-income-neighborhood racial diversity strata differ in the times they were called or their interaction with the survey firm in a way that affects their responses, these variables should account for that.[33] In our fourth specification, we estimate a specification with a full set of fixed effects for our 24 strata. This flexibly captures any differences in game responses that occur systematically across strata.

The results of estimating these 16 specifications are presented in Table 4. Panels A and B present estimates from the permutations of model (5). The sample is not stable across the four specifications, as shown in the bottom rows of the panels. A large number of respondents are missing data for the game conditions variables and, to a lesser extent, for the tract-level demographic controls.[34] These respondents are therefore omitted from the estimation when these variables are included. For this reason, our preferred specification includes the fixed effects strata. The strata fixed effects specification includes all respondents yet controls for group differences across our race-income-neighborhood diversity cells. Results from this specification are presented in bold in Table 4.

There is a high degree of consistency across results from the eight specifications in Panel A. *Black* is an insignificant predictor of acceptance. In all specifications, *relative* has a large, strongly significant and negative impact on acceptance – the smaller the relative share of the stakes offered, the less likely it is accepted. The implied marginal effect of a 10 percentage point decrease in *relative* is a decrease in the likelihood of acceptance of a little more than four percentage points. Moreover, the coefficient on *relative* is unaffected by the addition of the *relative*black* interaction term. This means that blacks and whites respond

---

[32] See McClellan and Skinner (1999) and Geronimus et al. (1996) for an example and discussion of using neighborhood level aggregates to proxy for individual characteristics.
[33] Some of the data appended by the survey firm, like time of interview, is missing for respondents who did not progress sufficiently far in the interview. Some of these respondents did manage to answer the ultimatum game question, so they are part of our sample.
[34] Based on conversations with the survey firm, it appears that respondents who ended the call before the very end of the script was reached (but after answering the ultimatum questions) were sometimes not assigned values for the game variables.

similarly to changes in the relative share that a proposed offer represents, holding offer size constant. We interpret this as evidence that blacks and whites have similar levels of aversion to inequality on average.

Changes in offer size, however, tell a different story. The estimated impact of *offer* is positive, and marginally significant in four cases (although not in our preferred specification) when the interactions with race are excluded. Once race interactions are included, the sign on *offer* changes (in three of the four control specifications). This is because blacks and whites differ significantly in their response to changes in the offer amount, holding relative constant. Estimates in the last four columns of Panel A show that an extra dollar in the offer increases the likelihood of a black respondent accepting by about two percentage points, an effect that is roughly half that of a ten percentage point increase in *relative*.

These results are robust to the inclusion of the *absolute* variable in Panel B. If anything, the patterns identified in Panel A are stronger and more consistent in Panel B. This is because coefficients on *offer* and *relative* no longer reflect the combined effects of those variables and *absolute*. *Black* is again insignificant in all specifications. *Offer* has a positive sign when the race interactions are excluded and a negative sign when they are included but is insignificant in all cases. The impact of *relative* is even more strongly negative than in Panel A. Again there is no significant difference between blacks and whites in the effect of *relative*. The offer effect identified for blacks in Panel A is larger in the Panel B estimates. *Absolute* does not significantly predict acceptance, although it consistently has a positive, small point estimate. In specifications with the race interactions, *black*absolute* has a similarly sized negative point estimate.

In Table 5, we estimate a mixed logit choice model that allows us to relax some of the assumptions behind the standard logit model, as discussed in Section IV. Specifically, we estimate a mixed logit model in which the coefficients on offer, relative, and absolute are allowed to differ across individuals. In other words, these are the "random" coefficients. We omit other controls from the model since the mixed logit is computationally intensive and the Table 4 analysis showed little impact of including additional controls on the standard logit estimates, consistent with our experimental approach. To compare the behavior across

racial groups, we estimate the mixed logit model on the full sample and then on subsamples of blacks and whites.

The mixed logit model generates estimates of the mean and standard deviation of (assumed) normally distributed coefficient vectors. Table 5 shows that the mean of the coefficients on offer and absolute inequality are insignificantly different from zero in the full sample, while the negative coefficient on relative inequality is highly significant, indicating that higher relative inequality reduces the likelihood of accepting an offer. The estimated standard deviations reported in the bottom half of Table 5 shows that in all three cases, the data fail to identify a confidence interval for the estimated standard deviation that excludes zero. This means that although the mean coefficient on relative inequality is different from zero, we have insufficient power to identify to what degree the coefficient differs across individuals in the population. We obtain very similar results to those for the full sample when we estimate the model separately on blacks and whites. The mean of the relative inequality coefficient is very similar across the two groups, and other coefficient means and all standard deviations are insignificant. The exception is the mean coefficient on offer in the black sample. As in the standard logit estimates, an increase in offer increases the probability of acceptance for blacks but not for whites. But as with the coefficients on relative inequality, there is little evidence that this parameter differs importantly across individuals within the black subsample.

The results are qualitatively similar in the mixed and standard logit models. Moreover, the offer effect identified for blacks in the standard logit model is robust to the more flexible mixed logit specification. Since the qualitative results for the (mean) coefficients are similar, we conclude that the standard logit estimates provide an adequate approximation for identifying differences across groups in the means of the underlying parameters. Unfortunately we simply lack the power to precisely estimate a distribution of heterogeneous coefficients.

ii. Understanding the Offer Effect and Robustness Checks

We know of no previous research that identifies an independent effect of offer size on acceptance behavior in the ultimatum game, so we are unable to compare this feature of our results to the work of others. In this section, we investigate the robustness of this finding in order to better understand it. One possible source of this result is that respondents may not have had a good understanding of the game. If blacks overlooked or misunderstood the fairness aspects of the game more frequently than whites, then we might observe an independent effect of offer size for blacks but not for whites.

We can investigate this possibility using answers to Question 3 (Q3) of our survey, which asked respondents to name the smallest offer that they would have found acceptable. As Table 1 documented, nearly 40% of respondents who answered Q3 provided an answer that was incompatible with their behavior in the game moments before. We separate our sample, conditional on answering Q3, into those who gave a response consistent with their behavior and those whose response was inconsistent.[35] We then re-estimate our preferred specifications from Table 4 on these subsamples. It is important to note that we consider the compatibility of Q3 to be only a suggestive proxy for understanding the game. It is possible that respondents who answered Q3 inconsistently nevertheless understood and answered Q2 (the accept/reject item) in a valid way. Nevertheless, we believe that the large number of incompatible responses justifies investigating whether the offer effect is driven by this particular subgroup.

We also investigated the sensitivity of our results to the stakes in the game. Some of our treatments use offer and stakes amounts that are small relative to findings in the literature on the importance of larger amounts for achieving salience with subjects (e.g. Slonim and Roth 1998; Cameron 1999). We therefore repeat our specifications restricting the sample to those who received treatments where the stakes were $20 or more.

The results of these checks are presented in Table 6. Several important points emerge from this analysis. First, respondents who gave inconsistent answers to Q3 exhibit a response to changes in relative

---

[35] Specifically, a consistent answer was defined as a minimum amount less than or equal to an accepted offer OR a minimum amount greater than a rejected offer. Inconsistent responses gave minimum amounts less than or equal to a rejected offer OR greater than an accepted offer.

inequality that is wrong-signed and sometimes statistically significant. Given that their response to increases in relative inequality differs so dramatically from that of respondents who provided consistent answers, we surmise that an inconsistent response to Q3 strongly suggests that the respondent did not understand the game.[36] Also, the coefficient on relative inequality among those with consistent responses is roughly double that in the full sample. This means that among respondents who appear to have understood the game, responses to even splits and near-zero share offers are closer to those found elsewhere in the literature. Very small shares are accepted with very low probability while even splits are accepted with very high probability.

Second, the offer effect identified for blacks in Table 4 is not driven by individuals with poor game comprehension. If anything, the opposite appears to be true. This is apparent from the second two columns in Panel B of Table 6. This model includes all three game variables and their interactions with *black*. Blacks who gave consistent responses to Q3 exhibit a larger, statistically significant response to increases in offer size than do whites with a point estimate that exceeds the Table 4 estimate, while blacks and whites who provided inconsistent responses to the two questions are insignificantly affected by *offer*.

Interestingly, a probit analysis of the determinants of a valid Q3 response turned up no significant predictors.[37] Although it is somewhat puzzling that variables like neighborhood income and education level do not predict game comprehension, there are reasons why this might be the case. Perhaps game comprehension requires catching respondents at a moment when they can briefly pay attention to the caller. If the likelihood of getting a respondent at a "good time" is unrelated to respondent characteristics – i.e. it is just surveyor luck—then again we would find no significant predictors of Q3 response.[38]

While we cannot know the exact reason for poor game comprehension in our experiment, it is helpful for our analysis that comprehension is unrelated to observable respondent characteristics. Given this

---

[36] The implied marginal effect of *share* is greater than 1 for respondents who gave consistent answers. This simply indicates that the slope of *share* is steeper than one over the range observed in the data.

[37] A probit model predicting inconsistent Q3 responses included the following as possible determinants: black dummy, female dummy, household size, age and age squared, state indicators, indicators for neighborhood diversity categories, median household income, and mean years of education at the Census tract level.

[38] In principle it is possible that there was a subset of interviewers at the survey firm who were difficult for respondents to understand. In this case, random allocation of phone numbers to surveyors would mean that respondent demographics are orthogonal to game comprehension. Based on our observation of the callers, we consider this possibility unlikely.

orthogonality, we continue our analysis using the complete sample of Q2 respondents and retaining the

experimental data in the form in which it was collected, rather than dropping those who appear not to have

understood the game. We assume this is a conservative strategy. Because game understanding is unrelated to

observable characteristics, and in particular to race, poor game comprehension on the part of a large

number of respondents likely attenuates our reported results.

The third point to take from Table 6 concerns the importance of salience (large stakes) for our

findings. The rightmost two columns in the Table 6 panels show that our main results hold when our

sample is restricted to those with consistent Q3 response and regardless of the stakes in the assigned

treatment. Specifically, the final two columns show that (i) blacks and whites respond similarly to changes in

relative inequality and (ii) there is a small but robust effect of offer size conditional on relative inequality

among blacks. Sample size becomes an issue here as our standard errors increase as we cut the data. But the

robustness of the point estimates on relative inequality and the interaction of *black*offer* leads us to conclude

that our main findings are not driven by problematic levels of salience nor by poor game comprehension

among some subjects.

The final two columns of Table 6 present our main specifications estimated separately on those

respondents who received offers representing different shares. Specifically, we were interested in trying to

understand what role the very low shares in some of our treatment profiles might have played in our results.

In experiments with subjects divided in both proposers and responders, proposers typically offer one-third

to one-half of the stakes (Forsythe et al. 1994). Offers representing shares much below that range are

uncommon in laboratory studies, although we implement them here to provide a broader survey of ranges

in which black and white responses might differ. We divided our sample into those who received offers

representing common, real-world shares (in the range of 0.25 to 0.5) versus those whose offers represented

lower shares.[39] Our standard errors increase considerably after splitting the sample in this way, but the point

---

[39] No treatments offered shares above 0.5.

estimate on the interaction of black and *offer* is similar in both subsamples, and achieves acceptable

significance levels in Panel A.[40]

To further investigate the source of the offer effect for, we re-estimated our baseline models on

subsets of the data defined by demographic characteristics, to determine if the offer effect might be stronger

in some groups than others. The results of interest from this exercise are presented in Tables 7 and 8.[41]

Table 7 presents our Table 4 specifications estimated separately on subsamples of our data divided

by sex and by age. Our parameter estimates vary little across any of these subgroups. In particular, the point

estimates for both *relative* and *black*offer* are very similar across all four subgroups. The effects are sometimes

insignificant, as we've reduced our sample sizes in these estimates, but the point estimates are stable and

very similar to those for analogous specifications in Table 4.

Table 8 shows the results of estimating our models separately on subsamples defined by income (the

same three income categories used to define our sample strata). Note that we omit strata fixed effects from

the models in Table 8 because (i) their inclusion does not substantively impact the results but does increase

imprecision, and (ii) the relevant strata are not constant across income categories, meaning that the

estimating equation would necessarily change across subsamples.

Panel A reports estimates from the full game variable specification. We first report estimates from

the model without race interactions, primarily for comparison. The results of interest are in the rightmost

three columns of Panel A. These estimates show that the effects of both *offer* and *relative* on acceptance differ

across income groups. The main effect of *relative* is largest for the lowest income group, and is significantly

larger (at the 10% level) than the coefficient on *relative* for the highest income group.

Although the main effect of *relative* is insignificant for the highest income group, this is a result of

retaining respondents with poor game comprehension. When the "inconsistent" group defined in Table 6 is

excluded, the patterns in Panel A of Table 8 are retained but the coefficients increase in magnitude and gain

---

[40] The p-value on *black*offer* in the second column from the right is 0.11.
[41] We performed the same exercise using subsamples defined by neighborhood racial diversity. The offer effect was again limited to blacks and was similar in size across all subsamples (and therefore similar in size to the estimates in Table 4). We found no differences of note in the other parameter estimates.

significance in some cases, consistent with the attenuation bias that error-prone data would introduce. In particular, the lowest income blacks in this exercise were significantly less influenced by *relative* than whites in that group. In the highest income group, the main effect of *relative* was positive and significant, and did not differ significantly across races.

Returning to the results in Table 8, the main effect of *offer* in the lowest income group has a negative coefficient and is weakly significant, but blacks in this group are significantly more likely than whites to accept after a dollar is added to *offer*. The net offer effect for blacks is roughly 0.031 and is significant at the 5% level – very similar to the result we reported for the complete sample in Table 4. The main effect and *black*offer* are both insignificant in estimates using the two higher income groups.

The negative and significant main effect of *offer* in the lowest income group is counterintuitive since, conditional on *relative*, we would expect higher offers to be more likely to be accepted. It is reasonable to worry that the inter-relationship between the three game variables is contributing to this. To assess this possibility, we present versions of the model with either *absolute* or *relative* omitted in Panel B. The main effect of *offer* is either insignificant or positive across all income groups in these models. *Black*offer* is still large and positive for the lowest income group, and the pattern of results for *relative* is also retained in the first three columns in Panel B. Results in Panel B were qualitatively the same when we omitted respondents with poor game comprehension from the estimation. In light of the Panel B results, it seems likely that the negative sign on offer size among whites in the lowest income group in Panel A is explained by the fact that *offer* is a function of *absolute* and *relative*.

We conclude that the offer effect for blacks identified in Table 4 is driven by blacks in the lowest income group.[42] Moreover, blacks and whites in the lowest part of the black income distribution respond differently to the *relative* and *offer* variables, with offer size having relatively more influence on blacks' acceptance decisions and *relative* having more influence for whites. Interestingly, lower income whites and

---

[42] It is possible that blacks in the lowest neighborhood income group are still poorer than whites in the lowest income group. Individual income differences may therefore contribute to this effect. We cannot rule this out with publicly available data, but we are skeptical that this is the entire explanation, since blacks in the other neighborhood income groups may also be poorer than whites in those groups yet we observe no black-white differences in these groups.

blacks both differ from their higher income racial counterparts, but they do so in different ways. For those in the highest income group, *relative* is less important than it is for low income whites and *offer* is less important than it is for low income blacks. Futhermore, we find that the offer effect among blacks in Table 4 is highly robust across male and female subsamples and across subsamples of working age versus older respondents. Our data do not allow us to analyze differences across multi-level demographic groups (e.g. neighborhood income *and* age), but the fact that we find that neighborhood income subsamples are the only instance in which our parameter estimates deviate noticeably from those in the overall sample strongly suggests to us that the most striking black-white differences in ultimatum game behavior occur at the lowest income levels.

iii. The Effect of Proposer Race

The final dimension to our experiment was the random assignment of the implied race of our hypothetical ultimatum proposer. As described in the data section, this was done through the use of racially distinct male names.[43] Tables 9 and 10 explore variation in acceptance rates across proposer-respondent race combinations. Table 9 presents raw acceptance rates for the four proposer-respondent race cells, and acceptance rates across the cells are all very close to the sample average. This is reflected in Table 10, which presents results from a logit model of the respondent's acceptance decision with controls for the four mutually exclusive race combinations. White respondent-white proposer is the omitted category. There is evidence that blacks facing a black proposer are more likely to accept than respondents in the other three scenarios. This is consistent with this cell having the highest observed acceptance rate in Table 9.

We explore the effect of proposer race further by adding interactions with proposer race to the offer-relative model discussed above. We use the offer-relative model, rather than the full game variable model, because (i) the previous analysis showed that adding absolute to the model had little impact on

---

[43] Note that callers for the survey firm were predominantly white. While this is probably helpful from the standpoint of preventing additional variation in the treatment, it is possible that the race of the caller is more salient than the implied race of the proposer. As a result, our variation on this dimension may suffer from reduced impact.

coefficients obtained from the offer-relative model, and (ii) the interactions with proposer race place additional strain on the data, so we try to eliminate unnecessary covariates from the model. We create a dummy variable equal to one if the name of the hypothetical proposer was a high frequency black name.[44] As before, we present results from specifications both with and without respondent race interactions with the randomly assigned variables.

The results are presented in Table 11. A number of findings from previous specifications carry over to the proposer race models. There is no average difference in acceptance rates across blacks and whites; *relative* is an important predictor of acceptance; and the offer effect for blacks is largely robust to the new specifications. With the caveat in mind that we may be asking a lot of the data to estimate these models – particularly the one which includes race interactions – the results in Table 11 reveal some interesting points about proposer race. First, the main effect of proposer race is insignificant. This means that, at least as we have conveyed it through our experiment, proposer race has no impact on overall acceptance rates. Proposer race also does not significantly alter the effects of the game variables *offer* and *relative*, or of respondent race. Finally, the fact that *proposer race * black* is small and insignificant appears at odds with the results in Table 10. However, the model with race interactions shows that blacks respond more strongly to increases in *relative* when the proposer is black. This likely explains why blacks facing black proposers exhibit somewhat higher acceptance rates in the more parsimonious model of Table 10. Therefore the Table 10 result turns out to be something more subtle than a level effect of black proposers when the respondent is black. Rather, blacks facing black proposers increase their acceptance rates more rapidly in response to higher offered shares than do blacks facing white proposers.

V. Conclusion

---

[44] An alternative is to define a proposer race dummy equal to one if the proposer's implied race was the same as the respondent's. We estimated this alternative model. The results are substantively similar to those presented in Table 10, but their interpretation is less transparent. We therefore omit them from the paper. They are available upon request.

We report results from an ultimatum game experiment played with a large sample of black and white respondents in the United States. Our respondents came from a stratified random sample drawn from registered voters in three Southern states. Our sample was representative of blacks and whites from low, middle, and high deciles of the black neighborhood income distribution and balanced on neighborhood diversity in those states. Respondents played the game over the telephone against a (hypothetical) proposer. We varied implied race of the proposer by using distinctively black and white names, and we varied dimensions of inequality by randomizing both the stakes and offer amount across games.

We examine aversion to two types of inequality—relative and absolute—characterized by rejection of an offer of a given level of inequality. We find that both blacks and whites are much more responsive to relative inequality than absolute inequality. We find no differences across races in inequality aversion as measured by rejection of offers with identical inequality characteristics. However, we find that blacks are more likely to accept larger offers conditional on relative inequality, and that this difference is driven in the behavior of blacks and whites in the lowest income category. In the framework of Bolton and Ockenfels (2000), this suggests that the lowest income blacks differ from the lowest income whites in terms of the importance of pecuniary returns for their decisions in this game. We also find an unequal effect of implied proposer race. Blacks increase their acceptance of the offered relative share more rapidly when the perceived race of the Proposer is black.

Our finding that black acceptance of offers is more responsive to shares offered by implied blacks than whites relates to prior work on discrimination (Fershtman and Gneezy 2001). That is, the behavior we observe is consistent with blacks being more sensitive to being treated unfairly by whites than by blacks.

At this point we can only speculate about the source of the offer effect among poorer blacks. One possible explanation relates to concerns about status. Prior work has shown that individuals will attempt to distinguish themselves from groups to which they belong demographically if the reference group is of low status. For instance, Charles, Hurst, and Roussanov (2009) show that blacks and Hispanics spend more than whites on conspicuous goods conditional on income. They argue that this behavior stems from concerns

about reference group status. In our study, poorer blacks showed a tendency to reject small offers ($1 or $2) regardless of the stakes; this tendency might be attributed to the stigma associated with welfare (Gilens 1999). When larger sums are offered, very low income blacks tend more than whites to accept the offer, perhaps because the value of the offer has exceeded the stigma associated with accepting it.

However, there are at least two alternative explanations for the offer effect among blacks. One possibility is that "gifts" might be associated with stronger expectations of future reciprocity among low income blacks as compared to low income whites. This could make accepting small gifts relatively less desirable. Like the stigma explanation, this explanation assumes that utility maximizing behavior in response to social concerns drives racial differences in ultimatum game behavior. A final possibility is that different underlying preferences drive racial differences in behavior. In our case, blacks and whites are both more likely to accept offers that represent larger shares of the stakes (lower relative inequality), suggesting a common preference for more equal splits, but blacks are more likely to impose a nominal lower bar below which any share of a split is unacceptable. While it is theoretically possible that a "reservation" level is part of preferences for blacks more often than for whites, the fact that only low income blacks respond to the game in this way suggests to us that social and status concerns are more likely explanations for the differences we observe.

Our findings suggest many interesting avenues for future work. We find that American blacks and whites differ in the importance they assign to pecuniary returns in the economic exchange of the ultimatum game, conditional on the inequality embodied in the exchange, with pecuniary returns playing a larger role for blacks. Understanding whether the differences detected in this artefactual field experiment have bearing on more substantive decision-making – such as preferences for redistribution, support for social programs, and labor market behavior – is certainly an important area to explore. A second interesting area to explore is the potential for racial differences in other notions of fairness, particularly procedural fairness versus distributional fairness, as examined in Bolton, Brandts, and Ockenfels (2005). Finally, our work suggests that

a much richer understanding of group differences in experimental behavior is possible through the creative

and widespread deployment of experiments in a representative population.

Works Cited:

Abraham, Katharine; Helms, Sara; and Presser, Stanley. 2009. "How Social Processes Distort Measurement: The Impact of Survey Non-response on Estimates of Volunteer Work in the United States." *American Journal of Sociology.* 114(4): 1129-1165.

Adimora, A. A.; Schoenbach, V. J; Martinson, F. E. A.; Stancil, T. R.; Donaldson, K. H. 2001. "Driver's License and Voter Registration Lists as Population-Based Sampling Frames for Rural African Americans." *Annals of Epidemiology.* 11(6): 385-88.

Alesina, A. and Ferrara, E. 2005. "Preferences for Redistribution in the Land of Opportunities." *Journal of Public Economics.* 89(2005): 897-931.

Andreoni, James, and Lise Vesterlund. 2001. "Which Is the Fair Sex? Gender Differences in Altruism," *Quarterly Journal of Economics* CXVI 293–312.

Atkinson, A.B. 1970. "On the Measurement of Inequality." *Journal of Economic Theory* 2: 244-63.

Ayres, I., and Siegelman, P. 1995. "Race and Gender Discrimination in Bargaining For a New Car." *American Economic Review* 85: 304-21.

Babcock, L, and Laschever, S. 2003. *Women Don't Ask: Negotiation and the Gender Divide.* Princeton University.

Barros, C. P.; Proenca, I.; and Vieira, C. J. 2005. "Low-Wage Employment in Portugal: A Mixed-Logit Approach." *IZA Discussion Paper No. 1667.*

Bartels, Larry M. *Unequal Democracy: The Political Economy of the New Gilded Age.* Princeton: Princeton University Press (2008).

Benjamin, D. J., Choi, J. J., and Strickland, A. 2009. "Social Identity and Preferences." *NBER Working Paper W13309.*

Berger, Jonah and Pope, Devin. 2010. "Can Losing Lead to Winning?" *Management Science.*

Bertrand, M. and Mullainathan, S. 2004. "Are Emily and Brendan More Employable that Latoya and Tyrone? Evidence on Racial Discrimination in the Labor Market from a Large Randomized Experiment." *American Economic Review.*

Blount, Sally. "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences." *Organizational Behavior and Human Decision Processes.* 63(2): 131-144.

Bolton, Gary; Jordi Brandts; and Axel Ockenfels. 2005. "Fair Procedures: Evidence from Games Involving Lotteries." *The Economic Journal.* 115(October): 1054-1076.

Bolton, Gary, and Elena Katok. 1995. "An Experimental Test for Gender Differences in Beneficent Behavior," *Economics Letters,* XLVIII 287–92.

Bolton, Gary, and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90(1): 166-193.

Camerer, C. 2003. "Behavioral Studies of Strategic Thinking in Games. *Trends in Cognitive Sciences* 7(5): 225-231.

Cameron, Lisa A. 1999. "Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia." *Economic Inquiry* 37(1): 47-59.

Cappelen, A. W.; Hole, A. D.; Sorensen, E. O.; Tungodden, B. 2007. "The Pluralism of Fairness Ideals: An Experimental Approach." *American Economic Review*, 97(3): 818-827.

Canon, David. 1999. *Race, Redistricting, and Representation: The Unintended Consequences of Majority Black Districts*. Chicago: University of Chicago Press.

Card, David, Alexandre Mas, and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics* 123(1):177-218.

Card, David; Stefano Della Vigna; and Ulrike Malmendier. Forthcoming. "The Role of Theory in Field Experiments." *Journal of Economic Perspectives.*

Carpenter, Jeffrey, Cristina Connolly, and Caitlin Knowles Myers. 2008. "Altruistic Behavior in a Representative Dictator Experiment." *Experimental Economics* 11: 282-98.
Chen, Kang and Fang-Fang Tang. 2009. "Cultural Differences between Tibetans and Ethnic Han Chinese in Ultimatum Bargaining Experiments." European *Journal of Political Economy* 25: 78-84.

Chuah, Swee-Hoon, Robert Hoffman, Martin Jones, and Geoffrey Williams. 2007. "Do Cultures Clash? Evidence from Cross-National Ultimatum Game Experiments." *Journal of Economic Behavior & Organization* 64: 35-48.

Dawson, Michael. 1995. *Behind the Mule: Race and Class in African-American Politics*. Princeton: Princeton University Press.

DeNavas-Walt, C., Proctor, B. D., and Smith, J. C. 2008. "Income, Poverty, and Health Insurance Coverage in the United States: 2007." US Census Bureau, US Department of Commerce.

Detang-Dessendre, C.; Goffette-Nagot, F.; Piguet, V. 2008. "Life Cycle and Migration to Rural and Urban Areas: Estimation of a Mixed Logit Model on French Data." *Journal of Regional Science*, 48(4): 789-824.

Doty, R. L., and Silverthorne, C. 1975. "Influence of Menstrual Cycle on Volunteering Behavior." *Nature* 254: 138–40.

Eckel, C. C., and Grossman, P. J. 1998. "Are Women Less Selfish Than Men?: Evidence from Dictator Experiments." *The Economic Journal* 108(448): 726-35.

Eckel, Catherine C. and Philip J. Grossman. 2001. "Chivalry and Solidarity in Ultimatum Games." *Economic Inquiry* 39(2): 171-88.

Falk, Armin; Fehr, Ernst; and Fischbacher. 2005. "Driving Forces behind Informal Sanctions." *Econometrica*, 73(6): 2017-30.

Falk, Armin; Meier, Stephan; Zehnder, Christian. 2011. "Did We Overestimate the Role of Social Preferences? The Case of Self-Selected Student Samples." *IZA Discussion Paper #5475*.

Fehr, E., and List, J. A. 2004. "The Hidden Costs and Returns of Incentives – Trust and Trustworthiness Among CEOs." *IEW Working Papers 134.*

Fehr, E. and Schmidt, K. M. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 119(3): 817-68.

Ferraro, Paul J. and Ronald G. Cummings. 2007. "Cultural Diversity, Discrimination, and Economic Outcomes: An Experimental Analysis." *Economic Inquiry* 45(2) 217-32.

Fershtman, C. and Gneezy, U. 2001. "Discrimination in A Segmented Society: An Experimental Approach." *The Quarterly Journal of Economics* 116(1): 351-77.

Fix, Michael and Turner, Margery, eds. *A National Report Card on Discrimination in America: The Role of Testing.* Washington, DC: Urban Institute Press, 1998.

Fong, C. M., Luttmer E. F. P. 2009. "Do Race and Fairness Matter in Generosity? Evidence from a Nationally Representative Charity Experiment." Working Paper Series rwp09-014, Harvard University, John F. Kennedy School of Government.

Forsythe, Robert; Joel Horowitz; N.E. Savin; and Martin Sefton. 1994. "Fairness in Simple Bargaining Experiments." *Games and Economic Behavior.* 6(3): 347-369.

Gerber, Alan S., Dean Karlan, and Daniel Bergan. 2009. "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions." *American Economic Journal: Applied Economics* 1(2): 35–52.

Gerber, Alan S., and Donald P. Green. 2000. "The Effects of Canvassing, Direct Mail, and Telephone Contact on Voter Turnout: A Field Experiment." *American Political Science Review* 94(3):653–63.

Geronimus, A. T., Bound, J., and Neidert, L. 1996. "On the Validity of Using Census Geocode Characteristics to Proxy Individual Socioeconomic Characteristics." *Journal of the American Statistical Association* 91(434): 529-37.

Gilens, Martin. 1999. *Why Americans Hate Welfare.* Chicago: University of Chicago Press.

Gneezy, U., and Rustichini, A. 2000. "Pay Enough or Don't Pay at All." *The Quarterly Journal of Economics* 115(3): 791-810.

Gneezy, U., Niederle, M., and Rustichini, A. 2003. "Performance in Competitive Environments: Gender Differences." *The Quarterly Journal of Economics* 118(3): 1049-74.

Gordon, M.E., L.A. Slade, and N. Schmitt. 1986. "The 'Science of the Sophomore' Revisited: From Conjecture to Empiricism." *Academy of Management Review* 11: 191-207.

Green, Donald P., and Alan S. Gerber. 2006. Can Registration-Based Sampling Improve the Accuracy of Midterm Election Forecasts? *Public Opinion Quarterly.* 70(2): 197-223.

Guth, W., R. Schmittberger, and B. Schwarze. 1982. "An Experimental Study of Ultimatum Bargaining." *Journal of Economic Behavior and Organization* 3: 367-88.

Harrison, Glenn and List, John. 2004. "Field Experiments." *The Journal of Economic Literature*. 42(4): 1009-1055.

Henrich, Joseph. 2000. "Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining Among the Machiguenga of the Peruvian Amazon." *The American Economic Review* 90(4): 973-79.

Henrich, Joseph; Boyd, R.; Bowles, S.; Camerer, C.; Fehr, E.; Gintis, H.; and McElreath, R. "In Search of Homo Economicus: Behavioral Experiments in 15 Small Scale Societies." *The American Economic Review*. 91(2): 73-78.

Hole, A. R. 2007. "Fitting Mixed Logit Models by Using Maximum Simulated Likelihood." *Stata Journal*. 7(3):388-401. RePec download of 2008 version: http://econpapers.repec.org/software/bocbocode/s456883.htm.

Jamison, Julian; Karlan, Dean; and Schechter, Laura. 2008. "To Deceive or Not to Deceive: The Effect of Deception on Behavior in Future Laboratory Experiments." *Journal of Economic Behavior and Organization*. 68: 477-488.

Karlan, Dean and Zinman, Jonathan. 2009. "Observing Unobservables: Asymmetries with a Consumer Credit Field Experiment." *Econometrica* 77(6): 1993-2008.

Kerwin K. C.; E. Hurst; and N. Roussanov. 2009. "Conspicuous Consumption and Race." *The Quarterly Journal of Economics* 124(2): 425-67.

Kravitz, D. A. and Gunto, S. 1992. "Decisions and Perceptions of Recipients in Ultimatum Bargaining Games." *Journal of Socio-Economics* 21(1): 65-84.

Levitt, S. and List, J. "Field Experiments in Economics: The Past, the Present, and the Future." *European Economic Review* 53(January): 1-18.

Luttmer, Erzo F. P. 2001. "Group Loyalty and the Taste for Redistribution." *Journal of Political Economy*, 109(3):500-28.

Lu, X.; Scheve, K. F.; and Slaughter, M. J. 2010. "Envy, Altruism, and the International Distribution of Trade Protection." *NBER Working Paper #15700*.

McClellan, M., and Skinner, J. 1999. "Medicare Reform: Who Pays, and Who Benefits?" *Health Affairs* 18(1):48-62.

Murnighan, J. Keith and Michael Scott Saxon. 1998. "Ultimatum Bargaining by Children and Adults." *Journal of Economic Psychology* 19: 415-45.

Nickerson, David W. 2008. "Is Voting Contagious? Evidence from Two Field Experiments," *American Political Science Review* 102(Feb):49-57.

Niederle, M., and Vesterlund, L. 2007. "Do Women Shy Away From Competition? Do Men Compete Too Much?" *The Quarterly Journal of Economics* 122(3): 1067-1101.

Oosterbeek, Hessel, Randolph Sloof, and Gus van de Kuilen. 2004. "Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis." *Experimental Economics* 7: 171-88.

Revelt, D. and Train, K. 1998. "Mixed Logit with Repeated Choices: Households Choices of Appliance Efficiency Level." *The Review of Economics and Statistics*, 80(4): 647-57.

Roth, A.E., V. Prasnikar, M. Okuno-Fujiware, and S. Zamir. 1991. "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study." *American Economic Review* 81: 1068-95.

Roth, A. 1995. "Bargaining Experiments". *The Handbook of Experimental Economics.* Princeton University.

Slonim, Robert and Alvin E. Roth. 1998. "Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic." *Econometrica* 66(3): 569-96.

Solnick, S. 2001. "Gender Differences in the Ultimatum Game." *Economic Inquiry* 39(2): 189-200.

Appendix 1: Survey Script

Hello, may I speak with [subject first name] [subject last name]?

I'm calling on behalf of researchers at the University of Notre Dame. You have been randomly selected to participate in a one question study. If you agree to answer the question, you will be entered in a drawing to win a $500 gift card from Amazon.com. You will also have an opportunity to make some money today.

Great! Just so you know, your personal information and answers will be kept confidential, used only to mail your winnings, and will be discarded when the project is completed. Would you like to participate in the study? [If yes, proceed, if no, thank them for their time].

Q1. You have been selected at random to play a game with [opponent first name] [opponent last name] from [opponent city], who was also selected at random from a statewide sample.

The rules of the game are simple. [Opponent first name] was asked to propose a split of [stakes] with you. If you accept the proposal, you will be paid that amount and [Opponent first name] will keep the rest. If you reject the proposal, neither of you will be paid anything. Are the rules clear? [If yes, proceed, if no, repeat the rules].

Q2. [Opponent first name] has proposed a split in which you receive [offer] of the [stakes] and [Opponent first name] receives [Amount – Offer]. Do you accept or reject the offer?

[Indicate acceptance or rejection].
    1 Accept
    2 Reject
    BOTH Accept and Reject count as a COMPLETE

Thank you.

ASK EVERYONE REGARDLESS OF ACCEPT OR REJECT:
Q3. By the way, what is the smallest amount you would have been willing to accept as a division of [stakes]?

[Enter amount]

Thank you for your participation. I'd like to remind you that you have also been entered in a drawing for a $500 Amazon gift card.

[IF OFFER WAS ACCEPTED, Q2=1] Can I confirm your name and the address where your check should be sent? [45]

[IF OFFER WAS REJECTED, Q2=2] Can I confirm your name and the address where the gift card will be sent if your name is drawn?

---

[45] Authors' note: This was the script as it was read to participants. However, the Iowa floods of summer 2008 forced us to change survey vendors at the last minute, as the vendor we initially contracted with was flooded and unable to do the survey as planned. Eastern Research Services mailed cash exclusively but the script was not updated to reflect this. As this information was only revealed at the end of the phone call, it has no bearing on selection into the sample or game behavior.

Appendix 2: Distribution of treatment profiles

Stakes	Offer	Number
100	1	50
100	2	100
100	5	125
100	10	100
50	1	50
50	2	100
50	5	125
50	10	100
20	1	50
20	2	125
20	5	125
20	10	50
10	1	75
10	2	125
10	5	125
5	1	75
5	2	125

Appendix 3: Construction of Racially Polarized Names

The state voter files used to draw the subject pool provided a natural population by which to calculate racially polarized names. Using this database was consistent with our back story (i.e., both players were randomly selected among registered voters) and also accurately reflected the distribution of names in the subject's social milieu (as opposed to using a national database). To calculate racial polarization of names, we first calculated the probability a person of a give race had a particular name (i.e., what percentage of blacks are named "John" and what percentage of whites are named "Sam"). The racial polarization of a name is then calculated by dividing the likelihood of a person of a black person having the name divided by the probability a white person has the same name (see Bertrand and Mullainathan, 2004).

$$Polarization = \frac{\Pr(White)}{\Pr(Black)}$$

A name was deemed racially polarized if the polarization score was greater than 10 or less than 0.10.[46] In order to select relatively common names, only names with at least 300 occurrences in each of the three state voter files were considered. We also limited the names to be male so as to keep gender constant. Applying these filters, there were 2 white first names, 10 black first names, 22 white first names, and 10 black last names remained candidates. After eliminating names deemed difficult to pronounce (e.g., Hensley), shared by prominent public figures (e.g., Helms), or possibly overly stereotypical (e.g., Tyrone), the authors selected two first names and two last names for each race. We ultimately selected Scott (polarization = 0.072) and Dustin (polarization = 0.097) as the "white" first names, Cedric (32.3) and Andre (21.2) as the "black" first names, Walsh (0.058) and Snyder (0.069) as the "white" first names, and Washington (42.3) and Booker (10.3) as "black" last names. Thus, we have four possible combinations of names for each race.[47] The polarization scores for each first and last name are shown in Tables A3.1 and A3.2, below. These tables also report the results of a survey of 200 individuals via Amazon's Mechanical Turk labor marketplace in which

---

[46] A "black" name was deemed polarized if the polarization score was greater than 10. Similarly, a "white" name was deemed polarized if the polarization score was less than 0.10.

[47] White: Scott Walsh; Scott Snyder; Dustin Walsh; Dustin Snyder. Black: Cedric Washington; Cedric Booker; Andre Washington; Andre Booker.

we asked respondents to rate an individual with each of our eight names as very likely black, more likely black, equally likely black/white, more likely white, or very likely white. The survey responses show strong polarization in the racial connotations of the names used in our survey. We suspect that the use of both a first and last name is critical to conveying this impression. "Cedric" alone may not clearly connote race, but "Cedric Washington" sends a clearer signal.

To give the treatment a little more detail and seem more real, we said the fictitious opponent lived in the state's biggest city.[48] To ensure that subjects were unlikely to know anyone with our created names, we check that no one shared those names in those cities. Across the three states, only six people shared any of the treatment names. Thus, subjects should have viewed the names as believable but would not know anyone sharing the same name.

Subjects of both races were randomly assigned to have a black or white opponent. One of the four created names was then randomly assigned for each race. Thus, each subject was equally likely to have one of the eight possibilities named as an opponent. Multiple names were created for each race to minimize the risk that the results of the experiment were dependent on a particular name. None of the results from any analysis conducted depends on the particular racialized name presented, so the analysis presented in the paper will only pay attention to the race of the opponent provided rather than the particular name used.

---

[48] Charlotte, North Carolina; Atlanta, Georgia; and, Columbia, South Carolina.

Table A3.1 Racial Polarization of Selected Black Names

		Polarization Score [Population Frequency]	Last Names	
			Booker	*Washington*
			0.04 **[2460]**	**0.07** **[430]**
First Names	*Andre*	**0.08** **[599]**	11% likely white 74% likely black	6% likely white 87% likely black
	Cedric	**0.05** **[386]**	9% likely white 83% likely black	2% likely white 87% likely black

Table A3.2 Racial Polarization of Selected White Names

		Polarization Score [Population Frequency]	Last Names	
			Snyder	*Walsh*
			25.6 **[505]**	**30.3** **[344]**
First Names	*Dustin*	**18.2** **[365]**	80% likely white 1% likely black	90% likely white 2% likely black
	Scott	**24.6** **[2555]**	85% likely white 1% likely black	89% likely white 0% likely black

Notes: Polarization score equals the percentage of whites with the name divided by the percentage of blacks with the name in the universe of registered voters in GA, NC, and SC. Number in brackets reports the total number of registered voters in GA, SC, and NC with the name. In the 4 x 4 subgrid, the top percentage is the share of 200 respondents to our survey on Amazon's Mechanical Turk who thought a person with the first/last name combination was very likely or more likely white on a five point scale. The bottom percentage is the share who thought a person with the first/last name combination was very likely or more likely black. The difference between the sum of these shares and one is due to respondents who thought a person with the name was equally likely to be black or white.

Appendix 4: Debriefing Letter to Participants

"Racial Differences in Inequality Aversion"
John D. Griffin, David Nickerson, and Abigail Wozniak, Principal Investigators
University of Notre Dame

Recently, you received a telephone call on behalf of researchers at the University of Notre Dame. Thank you for agreeing to participate in our study. Based on how you played this game, you may have earned a payment. If you did, it is enclosed.

In addition, we would like to briefly explain the purpose and procedure of the game you played. The purpose of the study in which you participated was to examine racial differences in aversion to inequality. In order to accomplish the objectives of the research and to do so at a reasonable cost, it was necessary for us to make a misrepresentation in the game you played. Specifically, the individual you played against was actually a fictional person and the amount you were offered was drawn from a distribution of offers from prior games.

The responses you provided will be used to examine differences in the sensitivity of racial groups in the United States to inequality and how this sensitivity is affected by the race and gender of the individuals with which they interact. Thank you once again for your participation.

If you would like to receive the results of this study, please contact the Principal Investigators at:

John D. Griffin
Department of Political Science
217 O'Shaughnessy
University of Notre Dame
Notre Dame, IN 46556
574-631-7659
John.Griffin@nd.edu

David Nickerson
Department of Political Science
217 O'Shaughnessy
University of Notre Dame
Notre Dame, IN 46556
574-631-7016
dnickers@nd.edu

Abigail Wozniak
Department of Economics and Econometrics
434 Flanner Hall
University of Notre Dame
Notre Dame, IN 46556
574-631-6208
awozniak@nd.edu

To read more about citizens' attitudes concerning inequality, see:

Eckel, Catherine C. and Philip J.Grossman. 2001. "Chivalry and Solidarity in Ultimatum Games." *Economic Inquiry* 39(2) 171-88.

Table 1: Descriptive Statistics within Black and White Subsamples

Variable	Black				White			
	Mean	*Std Dev*	*Min*	*Max*	**Mean**	*Std dev*	*Min*	*Max*
Characteristics from Voter File Record								
Age	57.11	*15.25*	*19*	*99*	59.47	*16.68*	*19*	*95*
Female	0.64	*0.48*	*0*	*1*	0.57	*0.50*	*0*	*1*
Household size	1.42	*0.61*	*1*	*3*	1.38	*0.55*	*1*	*3*
Neighb'd % Black	0.62	*0.25*	*0.2*	*1*	0.36	*0.25*	*0*	*0.80*
Characteristics from Merged Census Tract Data								
% Single parents	11.44	*5.28*	*0*	*34*	10.47	*5.60*	*1*	*35*
% In poverty	16.25	*10.40*	*0*	*54*	15.43	*10.40*	*0*	*62*
% Single unit	67.64	*19.66*	*5*	*100*	66.29	*18.77*	*3*	*100*
% College grads	16.65	*12.44*	*0*	*77*	16.51	*12.77*	*0*	*81*
% Homeowners	68.77	*19.32*	*2*	*97*	69.95	*18.82*	*2*	*95*
% Urban	62.75	*42.84*	*0*	*100*	54.02	*44.86*	*0*	*100*
% Blue collar	49.81	*15.99*	*10*	*88*	49.73	*15.60*	*6*	*86*
% Professionals	24.89	*12.13*	*0*	*76*	25.26	*11.80*	*0*	*63*
% White collar	33.07	*11.04*	*4*	*63*	32.95	*10.77*	*5*	*63*
% Unemployed	3.69	*2.83*	*0*	*24*	3.18	*2.72*	*0*	*16*
% Hispanic	2.14	*4.19*	*0*	*44*	1.86	*3.21*	*0*	*44*
% Asian	0.45	*1.26*	*0*	*11*	0.52	*1.57*	*0*	*21*
% 65 and over	22.21	*8.82*	*3*	*48*	22.95	*8.35*	*3*	*56*
Med HH income	35.5	*13.9*	*20.0*	*65.0*	35.6	*13.9*	*20.0*	*65.0*
Mean years educ	11.67	*1.18*	*8*	*16*	11.65	*1.14*	*8*	*16*
Game and Interview Variables								
Interview length	1.47	*0.84*	*0*	*5*	1.36	*0.92*	*0*	*6*
Times tried	4.77	*4.46*	*0*	*20*	4.35	*3.98*	*0*	*17*
Stakes	41.75	*35.65*	*5*	*100*	41.56	*35.78*	*5*	*100*
Offer	4.00	*2.99*	*1*	*10*	3.97	*2.98*	*1*	*10*
Share	0.18	*0.15*	*0.01*	*0.5*	0.18	*0.15*	*0.01*	*0.5*
Accept	0.37	*0.48*	*0*	*1*	0.34	*0.47*	*0*	*1*
Acceptable min	10.26	*16.72*	*0*	*100*	8.27	*15.60*	*0*	*100*
Acceptable share	0.27	*0.28*	*0*	*1*	0.21	*0.26*	*0*	*1*
Invalid min flag	0.40	*0.49*	*0*	*1*	0.42	*0.49*	*0*	*1*
N	818				829			

Notes: Data collected by Eastern Research Services via phone interviews for the authors, December 2008-January 2009. Median household income in $1000s. Share = offer / stakes. Acceptable share = acceptable min / stakes.

Table 2: Comparison of selected characteristics across subsamples

A. Contacted versus non-contacted subsamples

	Contacted	Not Contacted
Black	0.49	0.55*
Female	0.59	0.57
Age	58.68	51.86*
Median hh income	35536	37733*
Mean years of educ	11.68	11.80*
Stakes	42.0	41.8
Offer	4.01	4.00
Accept	0.35	-
N	2003	5666

B. Participants versus non-participants in the contacted subsample

	Participated	Did not participate
Black	0.50	0.48
Female	0.60	0.55
Age	58.3*	60.44*
Median hh income	35526	35581
Mean years of educ	11.66	11.72
Stakes	41.7	43.58
Offer	4.0	4.02
Accept	0.35	
N	1647	356

Notes: Contact defined as having day of interview recorded by survey firm. Participation defined as answering accept/reject ultimatum offer (Q2). * indicates different from the mean in the contacted (participated) subsample at the 5% level.

Table 3: Distribution of treatment variables across strata

Strata	Stakes		Offer		Share		Race Treatment		N
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	
1	39.04	4.22	4.01	0.35	0.19	0.02	0.55	0.06	73
2	40.25	4.05	3.92	0.34	0.19	0.02	0.51	0.06	79
3	42.43	4.23	4.24	0.34	0.18	0.02	0.53	0.06	70
4	46.75	4.03	4.09	0.31	0.18	0.02	0.52	0.05	91
5	43.49	4.24	3.86	0.34	0.16	0.02	0.59	0.06	73
6	37.21	4.16	3.89	0.38	0.19	0.02	0.48	0.06	61
7	44.36	4.33	4.30	0.35	0.18	0.02	0.54	0.06	70
8	37.66	4.14	3.95	0.38	0.19	0.02	0.58	0.06	64
9	39.43	4.81	3.32	0.38	0.15	0.02	0.49	0.07	53
10	36.17	4.09	3.80	0.41	0.18	0.02	0.60	0.06	60
11	45.58	5.01	3.87	0.42	0.16	0.02	0.52	0.07	60
12	46.02	4.58	4.45	0.41	0.17	0.02	0.48	0.06	64
13	43.44	4.13	4.09	0.33	0.18	0.02	0.51	0.06	77
14	47.63	4.23	4.93	0.37	0.19	0.02	0.46	0.06	80
15	43.24	4.18	4.03	0.34	0.18	0.02	0.54	0.06	74
16	40.71	4.09	4.27	0.36	0.19	0.02	0.40	0.06	70
17	40.66	3.85	3.28	0.31	0.15	0.02	0.38	0.06	76
18	46.08	4.23	4.01	0.33	0.16	0.02	0.40	0.05	83
19	46.16	4.58	4.45	0.39	0.18	0.02	0.52	0.06	69
20	42.19	4.48	3.47	0.36	0.15	0.02	0.47	0.06	64
21	32.24	4.53	3.57	0.35	0.22	0.02	0.41	0.07	58
22	43.77	5.11	3.91	0.44	0.16	0.02	0.45	0.07	53
23	32.27	3.98	3.55	0.35	0.20	0.02	0.41	0.06	66
24	36.10	4.13	3.85	0.37	0.18	0.02	0.42	0.06	59
Total	41.66	0.88	3.99	0.07	0.18	0.004	0.49	0.01	1647

Notes: Strata defined by race of respondent (two categories), three block level income categories, and four neighborhood racial diversity categories. Stakes and offer are randomly assigned within strata. Share equals offer/stakes. Race treatment randomly assigned within strata-stakes-offer cells.

Table 4: Standard probit models of ultimatum choice

A. Offer-Relative Specification

	Respondent race interactions omitted				Respondent race interactions included			
Black	0.035	0.023	-0.008	**0.14**	-0.037	-0.04	-0.03	**0.066**
	(1.50)	(0.91)	(0.26)	**(1.59)**	(0.53)	(0.55)	(0.39)	**(0.61)**
Offer	0.007	0.008	0.009	**0.007**	-0.002	-0.002	0.001	**-0.002**
	(1.71)†	(1.93)†	(1.87)†	**(1.70)†**	(0.69)	(0.32)	(0.09)	**(0.40)**
Relative	-0.44	-0.45	-0.60	**-0.43**	-0.44	-0.43	-0.50	**-0.43**
	(5.60)**	(5.70)**	(6.20)**	**(5.45)****	(3.92)**	(3.81)**	(4.15)**	**(3.83)****
Black*Offer					0.018	0.02	0.017	**0.018**
					(2.21)*	(2.34)*	(1.78)†	**(2.20)***
Black*Relative					-0.005	-0.05	-0.13	**0.001**
					(0.03)	(0.29)	(0.72)	**(0.01)**
Additional controls…								
Tract controls		X	X			X	X	
Game conditions			X				X	
Strata dummies				**X**				**X**
Observations	1647	1598	1500	1647	1647	1598	1500	1647

B. Full Specification

	Respondent race interactions omitted				Respondent race interactions included			
Black	0.035	0.023	-0.008	**0.14**	-0.074	-0.084	-0.08	**0.027**
	(1.50)	(0.90)	(0.27)	**(1.60)**	(0.94)	(1.06)	(0.88)	**(0.23)**
Absolute	0.001	0.001	0.001	**0.001**	0.001	0.001	0.002	**0.001**
	(1.08)	(1.29)	(1.30)	**(1.17)**	(1.44)	(1.67)†	(-1.70)	**(1.54)**
Offer	0.005	0.005	0.006	**0.005**	-0.007	-0.007	-0.005	**-0.007**
	(1.03)	(1.16)	(1.08)	**(1.00)**	(1.00)	(1.02)	(0.68)	**(1.04)**
Relative	-0.54	-0.58	-0.74	**-0.54**	-0.63	-0.66	-0.80	**-0.64**
	(4.40)**	(4.61)**	(5.08)**	**(4.38)****	(3.63)**	(3.73)**	(4.04)**	**(3.66)****
Black*Absolute					-0.001	-0.001	-0.001	**-0.001**
					(0.95)	(1.07)	(-1.02)	**(1.01)**
Black*Offer					0.022	0.024	0.022	**0.022**
					(2.42)*	(2.59)**	(2.06)*	**(2.43)***
Black*Relative					0.17	0.16	0.10	**0.19**
					(0.71)	(0.66)	(0.34)	**(0.78)**
Additional controls…								
Tract controls		X	X			X	X	
Game conditions			X				X	
Strata dummies				**X**				**X**
Observations	1647	1598	1500	1647	1647	1598	1500	1647

Notes: Dependent variable is indicator for acceptance of proposed split. All equations estimated via probit; implied marginal effects calculated at the means reported. Absolute value of robust z statistics in parentheses. † significant at 10%; * significant at 5%; ** significant at 1%.

Table 5: Mixed logit estimates of random coefficient choice model

	Full Sample	Whites	Blacks
Coefficient Mean			
Offer	0.033	-0.003	0.064
	(1.54)	(-0.09)	(2.16)*
Absolute	-0.005	-0.016	0.001
	(-0.39)	(-0.45)	(0.28)
Relative	-2.81	-2.26	-2.94
	(-2.69)**	(-1.66)†	(-2.48)*
Coefficient SD			
Offer	0.003	0.001	0.003
	(0.99)	(0.01)	(0.01)
Absolute	0.028	0.048	0.00
	(0.86)	(0.69)	(0.02)
Relative	3.13	2.58	2.73
	(0.94)	(0.42)	(0.76)

Notes: Dependent variable is indicator of respondent's choice between accepting and rejecting ultimatum. Model was estimated using mixed logit routine in Stata developed by A. R. Hole (2008) with nreps set to 500. No other covariates included. Z-statistics in parentheses.

Table 6: Standard probit models separately by Q3 response, stakes size, and share size

A. Offer-Relative Specification

Q3 response relative to Q2:	Consistent	Not Consistent	Consistent	Not consistent	Consistent	Consistent	All	All
Additional restrictions:	-	-	-	-	Small Stakes	Large Stakes	Share in 0.25-0.50	Share < 0.25
Black	0.04	0.30	-0.15	0.19	-0.35	0.07	-0.10	-0.31
	(0.28)	(1.99)*	(0.89)	(1.07)	(0.95)	(0.25)	(0.44)	(0.16)
Offer	0.01	0.002	-0.002	-0.011	-0.002	0.008	-0.004	-0.001
	(1.58)	(0.42)	(0.24)	(1.36)	(0.03)	(0.67)	(0.28)	(0.22)
Relative	-1.18	0.29	-1.26	0.22	-1.13	-0.88	-0.62	-0.11
	(9.48)**	(2.97)**	(6.87)**	(1.37)	(2.07)*	(2.84)**	(1.80)†	(0.36)
Black*Offer			0.02	0.03	0.04	0.01	0.04	0.02
			(1.89)†	(1.96)*	(0.55)	(0.71)	(1.57)	(1.85)†
Black*Relative			0.12	0.11	0.25	-0.26	0.20	0.02
			(0.48)	(0.54)	(0.34)	(0.58)	(0.42)	(0.05)
Observations	951	663	951	663	293	658	438	1209

B. Full Game Variable Specification

Q3 response relative to Q2:	Consistent	Not Consistent	Consistent	Not consistent	Consistent	Consistent	All	All
Additional restrictions:	-	-	-	-	Small Stakes	Large Stakes	Share in 0.25-0.50	Share < 0.25
Black	0.04	0.13	-0.20	0.21	-0.09	-0.16	-0.07	-0.28
	(0.31)	(1.40)	(1.08)	(1.12)	(0.2)	(0.49)	(0.24)	(0.96)
Absolute	0.002	-0.001	0.003	-0.001	0.05	0.004	-0.04	0.001
	(2.71)**	(0.93)	(2.89)**	(0.75)	(0.91)	(3.23)**	(0.96)	(1.31)
Offer	0.002	0.004	-0.013	-0.007	-0.04	-0.01	0.01	-0.009
	(-0.24)	(0.85)	(1.43)	(0.72)	(0.6)	(0.94)	(0.55)	(1.04)
Relative	-1.60	0.40	-1.92	0.36	-2.37	-2.18	1.02	-0.71
	(7.83)**	(2.57)*	(6.14)**	(1.39)	(1.58)	(4.57)**	(0.58)	(1.31)
Black*Absolute			-0.002	0.00	0.01	-0.003	0.008	-0.002
			(-1.43)	(0.36)	(0.16)	(1.5)	(0.14)	(1.33)
Black*Offer			0.03	0.016	0.04	0.02	0.03	0.027
			(2.29)*	(1.43)	(0.4)	(1.32)	(0.89)	(2.26)*
Black*Relative			0.62	0.03	-0.03	0.66	-0.14	0.89
			(1.48)	(0.10)	(0.02)	(0.93)	(0.06)	(1.15)
Observations	951	663	951	663	293	658	438	1209

Notes: Q2 is the respondent's choice of whether to accept or reject the ultimatum offer. Q3 is the respondent's statement about the minimum amount s/he would have accepted. Roughly 40% of respondents who answered Q3 give minimum amounts inconsistent with their choices in Q2. All specifications contain a full set of strata fixed effects. 36 respondents did not answer Q3. Dependent variable is indicator for acceptance of proposed split. All equations estimated via probit; implied marginal effects calculated at the means reported. Absolute value of robust z statistics in parentheses. † significant at 10%; * significant at 5%; ** significant at 1%.

Table 7: Standard probit models separately by age and sex

A. Offer-Relative Specification

Subsample:	Women	Men	Ages 18-55	Ages 55+
Black	0.067	-0.008	0.153	-0.137
	(0.48)	(0.04)	(0.83)	(1.08)
Offer	-0.01	0.007	0.003	-0.003
	(1.31)	(0.78)	(0.32)	(0.42)
Relative	-0.38	-0.519	-0.534	-0.335
	(2.58)**	(2.88)**	(2.80)**	(2.40)*
Black*Offer	0.016	0.025	0.019	0.02
	(1.55)	(1.81)+	(1.38)	(1.90)+
Black*Relative	-0.014	-0.001	-0.064	0.044
	(0.07)	(0.00)	(0.24)	(0.22)
Observations	992	655	715	932

B. Full Game Variable Specification

Subsample:	Women	Men	Ages 18-55	Ages 55+
Black	-0.02	-0.139	0.095	0.031
	(0.13)	(0.73)	(0.48)	(0.22)
Absolute	0.001	0.001	0.002	0.001
	(1.28)	(0.97)	(1.47)	(0.87)
Offer	-0.015	0.003	-0.003	-0.006
	(1.79)+	(0.33)	(0.28)	(0.80)
Relative	-0.605	-0.726	-0.859	-0.48
	(2.67)**	(2.60)**	(2.92)**	(2.25)*
Black*Absolute	-0.002	0.001	-0.002	-0.001
	(1.66)+	(0.43)	(0.91)	(0.74)
Black*Offer	0.025	0.023	0.024	0.024
	(2.18)*	(1.45)	(1.58)	(2.08)*
Black*Relative	0.375	-0.139	0.221	0.22
	(1.2)	(0.34)	(0.55)	(0.71)
Observations	992	655	715	932

Notes: Q2 is the respondent's choice of whether to accept or reject the ultimatum offer. Q3 is the respondent's statement about the minimum amount s/he would have accepted. Roughly 40% of respondents who answered Q3 give minimum amounts inconsistent with their choices in Q2. All specifications contain a full set of strata fixed effects. 36 respondents did not answer Q3. Dependent variable is indicator for acceptance of proposed split. All equations estimated via probit; implied marginal effects calculated at the means reported. Absolute value of robust z statistics in parentheses. † significant at 10%; * significant at 5%; ** significant at 1%.

Table 8: Standard probit models by income category

A. Full Specification

	Black 10th Percentile	Black Median	Black 90th Percentile	Black 10th Percentile	Black Median	Black 90th Percentile
Black	0.054	0.03	0.017	-0.20	-0.129	0.16
	(1.33)	(0.74)	(0.37)	(1.59)	(0.90)	(1.03)
Absolute	0.002	-0.001	0.001	0.002	0.00	0.001
	(1.68)†	(0.96)	(1.22)	(1.72)†	(0.01)	(0.87)
Offer	0.006	0.008	0.00	-0.018	0.005	-0.006
	(0.91)	(1.00)	(0.05)	(1.76)†	(0.49)	(0.43)
Relative	-0.71	-0.27	-0.63	-0.87	-0.58	-0.383
	(3.55)**	(1.30)	(2.67)**	(3.03)**	(1.93)*	(1.18)
Black*Absolute				-0.001	-0.002	0.00
				(0.73)	(1.00)	(0.07)
Black*Offer				0.049	0.006	0.009
				(3.28)**	(0.35)	(0.50)
Black*Relative				0.284	0.61	-0.57
				(0.71)	(1.46)	(1.19)
Observations	614	560	473	614	560	473

B. Offer-Relative and Absolute-Offer Models

|	Offer-Relative Model			Absolute-Offer Model		
	Black 10th Percentile	Black Median	Black 90th Percentile	Black 10th Percentile	Black Median	Black 90th Percentile
Black	-0.15	-0.05	0.15	-0.12	0.05	-0.007
	(1.38)	(0.45)	(1.14)	(1.61)	(0.72)	(0.09)
Absolute				-0.001	-0.002	0.00
				(0.98)	(2.10)*	(0.02)
Offer	-0.011	0.005	0.00	-0.004	0.016	0.002
	(1.13)	(0.57)	(0.04)	(0.41)	(1.75)†	(0.17)
Relative	-0.49	-0.587	-0.17			
	(2.67)**	(3.07)**	(0.80)			
Black*Absolute				0.00	0.00	-0.002
				(0.26)	(0.07)	(1.17)
Black*Offer	0.04	-0.002	0.009	0.04	-0.006	0.017
	(3.22)**	(0.16)	(0.60)	(3.20)**	(0.45)	(1.14)
Black*Relative	0.06	0.29	-0.53			
	(0.22)	(1.09)	(1.71)†			
Observations	614	560	473	614	560	473

Notes: Dependent variable is indicator for acceptance of proposed split. All equations estimated via logit; implied marginal effects calculated at the means reported. Models include only listed covariates. Absolute value of robust z statistics in parentheses. † significant at 10%; * significant at 5%; ** significant at 1%.

Table 9: Raw acceptance rates by race of respondent and implied race of proposer

	Proposer:	
Respondent:	**Black**	**White**
Black	0.39	0.35
	(0.02)	(0.02)
White	0.35	0.33
	(0.02)	(0.02)

Notes: Data collected by Eastern Research Services for the authors over December 2008-January 2009. Standard errors of mean acceptance rates within the cells are in parentheses.

Table 10: Probability of acceptance across proposer-respondent race combinations

Respondent-Proposer Race:	**Marginal Effect**	
Black-Black	0.184	0.065
	(2.00)*	(2.02)*
Black-White	0.14	0.022
	(1.5)	(0.66)
White-Black	0.026	0.022
	(0.77)	(0.64)
Strata Fixed Effects	Yes	No
Observations	1647	1647

Notes: Dependent variable is indicator for acceptance of proposed split. All equations estimated via probit; implied marginal effects reported and robust z-statistics in parentheses. White-White is the omitted category. P-value of F-test that the three reported coefficients are jointly different from zero is 0.33 in fixed effects column; 0.24 in no fixed effects column. * significant at 5%.

Table 11: Standard probit model with proposer race interactions

Proposer Race (PR) Definition:	Proposer Race=1 if Proposer Black	
Black	0.022	-0.14
	(0.65)	(1.49)
Offer	0.008	0.00
	(1.33)	(0.02)
Relative	-0.37	-0.52
	(3.37)**	(3.53)**
PR*Black	0.022	0.21
	(0.45)	(1.39)
PR*Offer	-0.002	-0.005
	(0.19)	(0.46)
PR*Relative	-0.13	0.20
	(0.81)	(0.87)
PR	0.065	-0.02
	(0.88)	(0.18)
Black*Offer		0.015
		(1.33)
Black*Relative		0.33
		(1.50)
PR*Black*Offer		0.006
		(-0.36)
PR*Black*Relative		-0.65
		(2.03)*

Notes: Dependent variable is indicator for acceptance of proposed split. All equations estimated via probit; implied marginal effects calculated at the means reported and absolute value of robust z-statistics in parentheses. N equals 1647 in all specifications. † significant at 10%; * significant at 5%; ** significant at 1%.

FIGURES 1a-1e. Figures show mean acceptance rates by offer amount separately for blacks (squares) and whites (diamonds). Vertical lines show 95% confidence intervals around mean estimates.
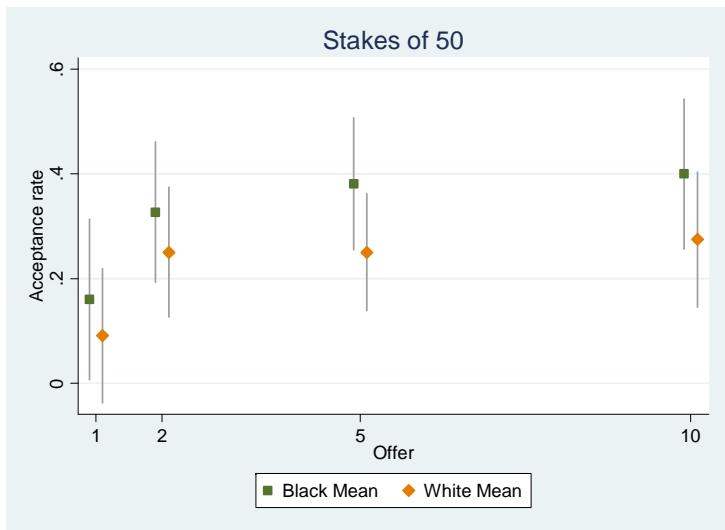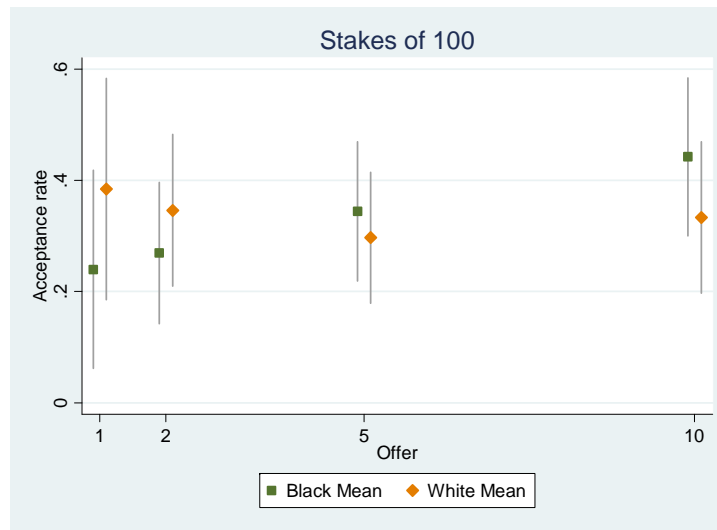

1a.
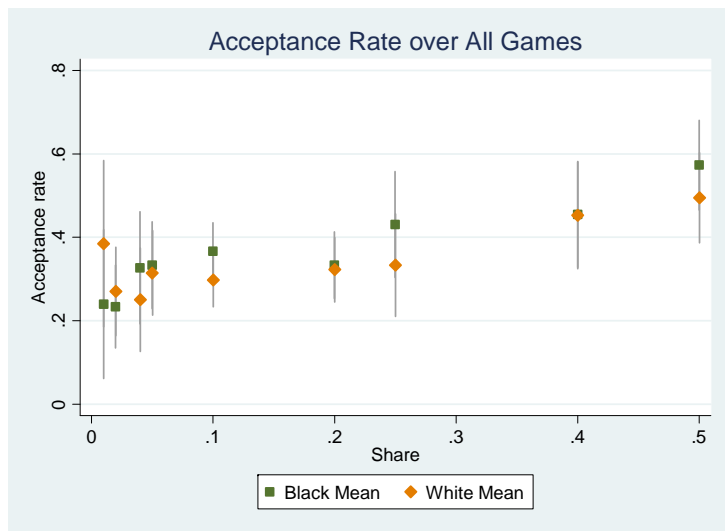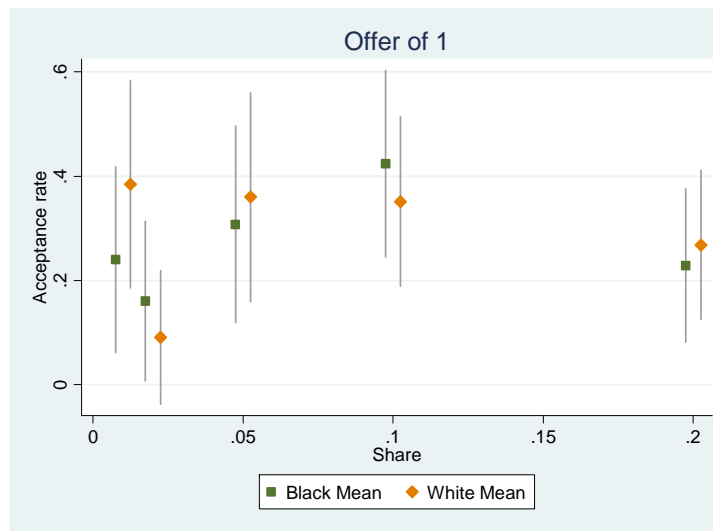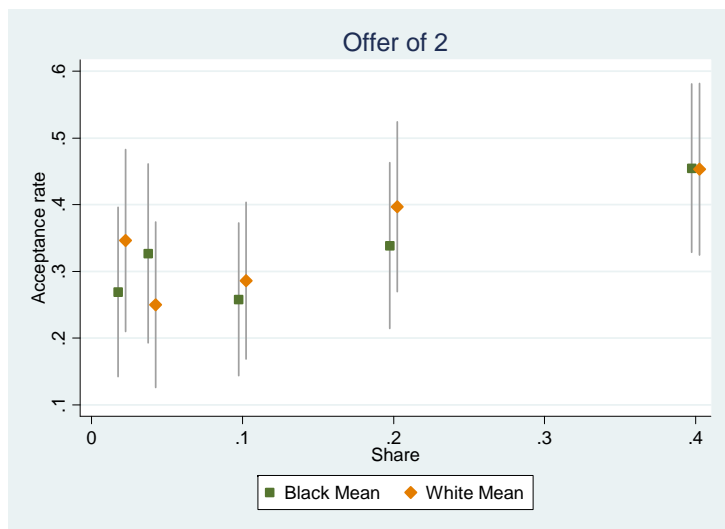
1b.

1c.

1d.

Stakes of 50

Stakes of 100
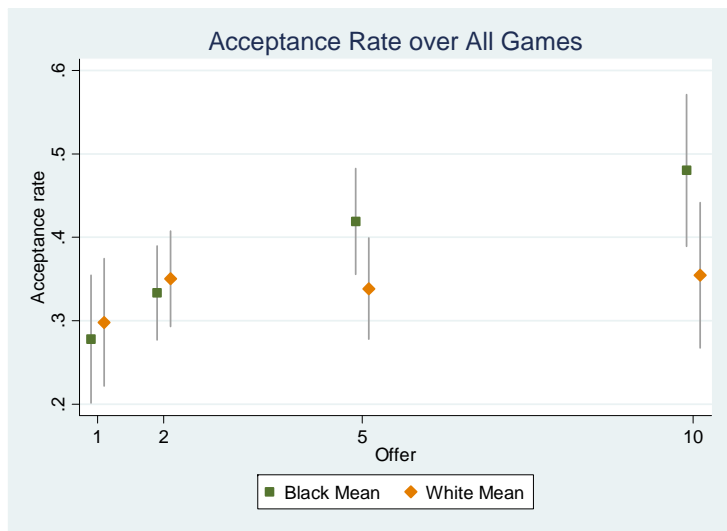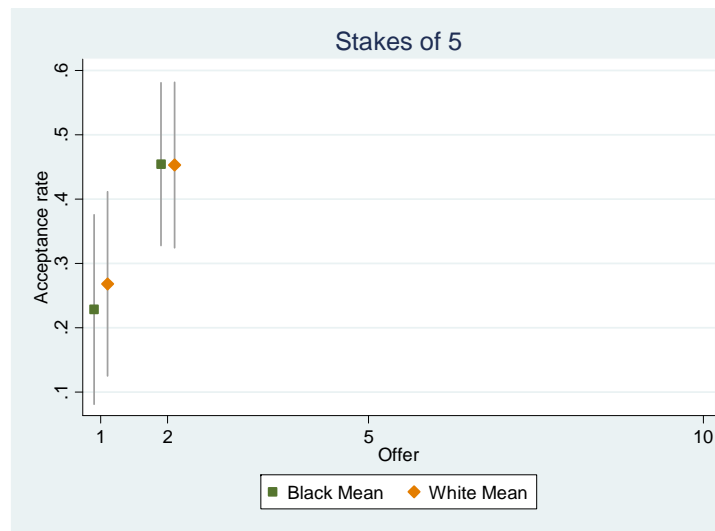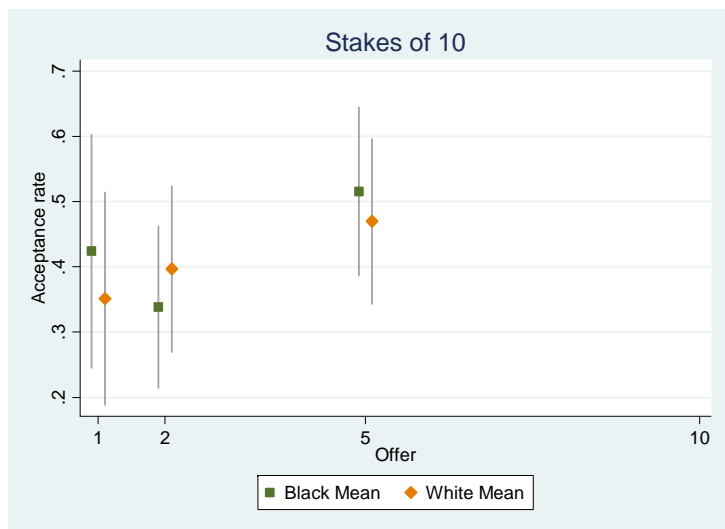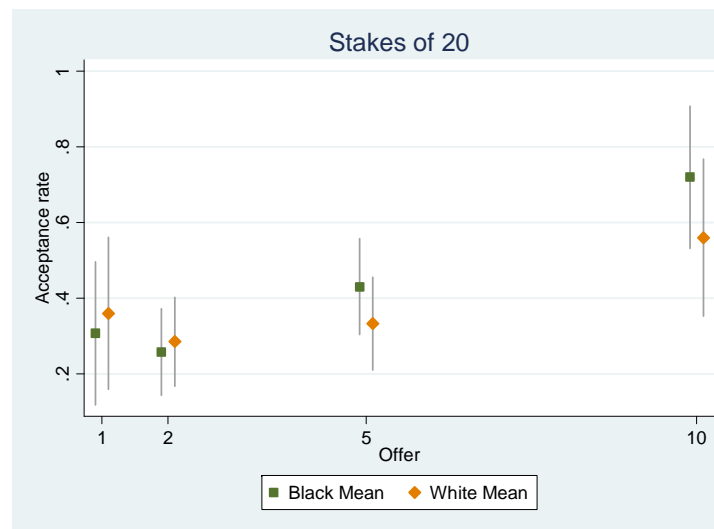
1e.

1f.

FIGURES 2a-2e. Figures show mean acceptance rates by share separately for blacks (squares) and whites (diamonds). Vertical lines show 95% confidence intervals around mean estimates.
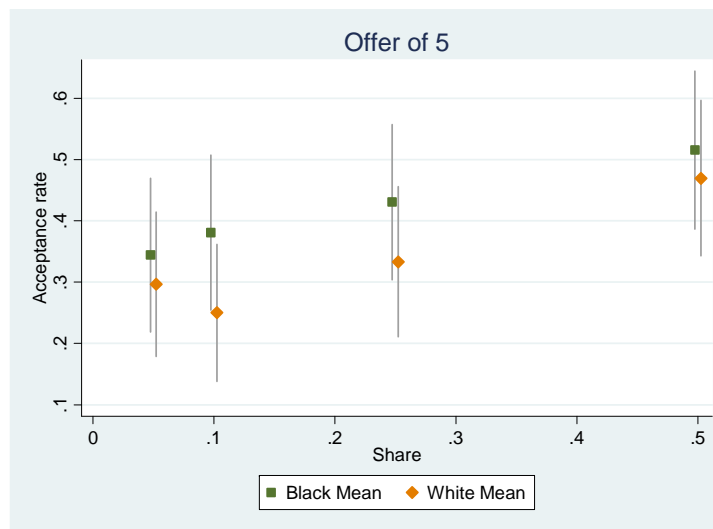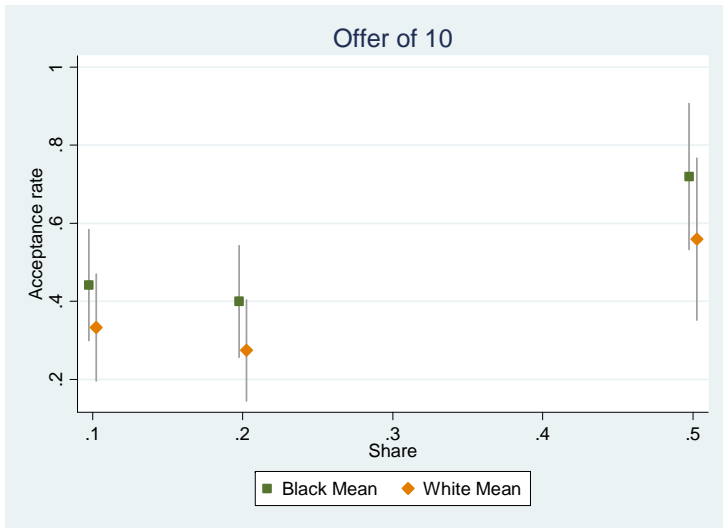

2a.

2b.

2c.

2d.

2e.