G Model
JEBO-2996; No. of Pages 18

ARTICLE IN PRESS

Journal of Economic Behavior & Organization xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game

John Griffin [a], David Nickerson [b], Abigail Wozniak [c,*]

[a] Department of Political Science, University of Colorado, United States
[b] Department of Political Science, University of Notre Dame, United States
[c] Department of Economics, University of Notre Dame, NBER and IZA, United States

ARTICLE INFO

Article history:
Received 13 October 2011
Received in revised form 9 August 2012
Accepted 8 September 2012
Available online xxx

JEL classification:
J15
D03
D63
C72
C91

Keywords:
Ultimatum game
Racial differences
Fairness
Inequality aversion
Artefactual field experiments

ABSTRACT

The distinct historical and cultural experiences of American blacks and whites may influence whether members of these groups perceive a particular exchange as fair. We investigate racial differences in fairness standards using preferences for equal treatment in the ultimatum game. We focus on whether responders choose to accept a proposed division of a monetary amount or to block it. We use a sample of over 1600 blacks and whites drawn from the universe of registered voters in three states merged with information on neighborhood income and racial composition. We experimentally vary proposed divisions as well as the implied race of the proposer. We find that acceptance in both groups is strongly influenced by the level of inequity in a proposed division, but blacks are also influenced by the offer amount while whites are not. This is driven by the lowest income group in our sample, which represents the 10th percentile of the black income distribution. We are able to reject that blacks and whites in this group share a common, simple utility function. We also find that blacks are more sensitive to unfair proposals from other blacks.

1. 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, we 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 widely recognized by economists. Inequality aversion has been shown to operate in a variety of contexts, including bargaining in product

* Corresponding author. Address: 441 Flanner Hall, University of Notre Dame, Notre Dame, IN 46556, United States. Tel.: +1 574 631 6208.
E-mail address: a_wozniak@nd.edu (A. Wozniak).

0167-2681/$ – see front matter
http://dx.doi.org/10.1016/j.jebo.2012.09.010

and labor markets (Babcock and Laschever, 2003), decisions over redistribution (Luttmer, 2001), bequests and charitable giving (Fong and Luttmer, 2011), and political support for regulation (Lü et al., 2012). Examining racial differences in inequality aversion 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 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 to test for racial differences in inequality aversion between whites and African Americans in the United States. 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 et al., 2005). Our approach enables us to isolate racial differences in responses to unequal treatment and thereby answer several questions. First, does a minority group that has experienced a high level of discrimination and unfair treatment respond differently in the ultimatum game? Second, does this depend on the neighborhood income levels or racial integration that respondents experience? Finally, does an unfair proposal elicit different responses when the proposer shares the responder's race as opposed to when the responder does not?

We conduct a survey experiment in which respondents are invited to play the ultimatum game by phone.[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 related to income, a characteristic which varies substantially across blacks and whites (Alesina and La Ferrara, 2005). Our final sample contains 1647 respondents.

All participants in our study play the role of responder in the ultimatum game. We focus on the responder's decision to accept or reject the offer in order to focus on inequality aversion.[2] Respondents are told that they have been selected to play a game with a proposer who 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 proposer race.

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 the proposer. 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 stakes being offered. This difference is driven by black respondents from the lowest income neighborhoods. Second, the implied race of the proposer matters for black respondents. African Americans facing African American proposers are more sensitive to the offered shares than any other proposer-respondent race combination.

Our paper makes three contributions. First, 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 experiments to uncover striking and potentially important group differences in a variety of experimental settings. Differences along lines of gender and ethnicity (or nationality) 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 et al., 2003; Niederle and Vesterlund, 2007. On nationality/ethnicity: Roth et al., 1991; Henrich, 2000; Henrich et al., 2001; Chuah et al., 2007; Ferraro and Cummings, 2007; Fershtman and Gneezy 2001; Chen and Tang, 2009).[3] However, the study of *racial* differences in economic decision-making has so far been limited. Benjamin et al. (2010), who examine the effects of race priming on subjects' expressed preferences for risk aversion in a university-based sample, and Fong and Luttmer (2011) who use a representative sample to analyze racial differences in charitable giving. Ayres and Siegelman (1995) find that the racial composition of bargaining pairs affects bargaining behavior. Eckel and Grossman (2001) touch on racial differences among subjects in an experiment designed to address other questions.

Our second contribution is to implement a sampling methodology that allows us to answer these questions for a broad cross-section of the population. 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. Voter registration files have been used outside economics as a sampling frame from

[1] Others have studied inequality aversion using survey data (Atkinson, 1970).
[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).
[3] There are also some studies of age differences (Murnighan and Saxon, 1998).

which 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 et al., 2009) and pollsters.[4] This method has advantages over others in the experimental economics literature. Our sample contains large numbers of respondents from key 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. Our approach also 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.[5]

A final contribution is that our experimental design allows us to study responses to different types of inequality in ultimatum game offers. We combine two types of inequality modeled in the literature into a single model (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). Our experiment varies offer size and stakes randomly across players, so we can determine whether deviations from even shares (as in Bolton–Ockenfels) or unequal nominal divisions (as in Fehr–Schmidt) are more important for the acceptance decision. 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. Andersen et al. (2011) show that the stakes size matters in the ultimatum game, as respondents trade off rising nominal offer amounts for inequality at higher stakes levels.

2. 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.[6] 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.[7] The callers were carefully trained regarding the rules of the game and the nature of the research.[8] In particular, the callers were trained to read the rules of the game slowly and answer questions about the game.[9] The script used during the course of calling is available in Appendix 1.[10] 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 are detailed 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.

[4] See the discussion on the Field Poll's website: http://www.field.com/fieldpoll/methods.html.

[5] Examples of these databases include Knowledge Networks and Time Sharing Experiments in the Social Sciences (TESS). Other studies, including Bellemare et al. (2008), have successfully used maintained online subject pools to conduct lab style experiments in representative populations.

[6] Home city was chosen to be a large city in the respondent's own state.

[7] 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).

[8] 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.

[9] 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.

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

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 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 raises questions among experimental economists. Our design deceives respondents in telling them that they are paired with a proposer who in fact is hypothetical. The design relies on respondents making inferences about the race and gender of the proposer. Deception in experimental instructions is typically eschewed by experimental economists, largely although not exclusively on the grounds that it may contaminate subject pools by making participants question the truth of instructions in future studies.[11] A number of studies have nevertheless 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.

Harrison and List (2004) describe the ideal experiment as featuring no deception.[12] While we agree that this is ideal, we also believe that potential costs and benefits of a design should be considered before it is adopted or rejected. In the case of our experiment, the potential benefits included learning about a population that (a) is of tremendous importance to social policy and (b) is very difficult to bring to a lab or otherwise reach in large numbers given the current samples in maintained online subject pools. The major cost to the deceptive component of our design is the potential for future subjects in similar experimental settings to discount the instructions they receive.

In light of what we viewed as large potential benefits to the study, we elected to carry out our design in a way that minimized the costs from deception. There were three steps we took to reduce any potential negative impact on future studies from deception in our own. First, while respondents likely infer that the hypothetical proposer is a real person, we go to no particular lengths to convince them of this.[13] Our statement to subjects was the following: "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." Beyond this, we say little about the proposer. Our statement is false, but it is perhaps less false than training an actor as a study confederate, for example, although the latter is a common practice in psychology experiments (Mendes et al., 2008).

The second step we took was to conduct our experiment in a setting in which trust levels among subjects are already low – the "cold call." It is not unreasonable to think that the public attaches less weight to information in "cold calls," since these are often used for sales pitches, political campaigning, and other – possibly fraudulent – solicitations.[14] Since subjects in this situation are conditioned to place less trust in a caller, a norm of non-deception within economics is unlikely to alter the behavior of subjects within a game *in the context of a cold call*. We also believe that deception in a research experiment conducted during a cold call is unlikely to affect subject behavior in other settings. For example, insurance and travel packages are frequent vehicles for fraudulent telemarketing, but it is unlikely that this affects consumer behavior toward legitimate firms in these businesses. From the perspective of the subject, the most important details of our call were truthful (i.e. call from university researchers with real money prizes for playing a game). We debriefed subjects about the nature of the deception in a letter following the experiment. See Appendix 4. We discuss additional advantages to using our sample in the next section. We believe these add to the benefits of doing this study while holding constant the main costs. Our use of the cold call setting may raise some questions about the salience of our experiment for subjects. We can partially address this by examining robustness of our findings across various subgroups in the data. We discuss this in detail when we present the results.

[11] 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.

[12] 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."

[13] 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.

[14] According to a 2008 news release, the FTC was pursuing over 180 telephone fraud cases by the late 2000s. The agency estimated the fraud in those cases affected over half a million consumers. (FTC Press Release May 20, 2008. "FTC Announces Operation Tele-PHONEY, Agency's Largest Telemarketing Sweep.")

Please cite this article in press as: Griffin, J., et al., Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game. J. Econ. Behav. Organ. (2012), http://dx.doi.org/10.1016/j.jebo.2012.09.010

3. Sample

Experimental games frequently take place in laboratories and rely on undergraduate subject populations.[15] 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 approval and lower authoritarianism than non-volunteers". Studies of social preferences that have compared the behavior of students and non-students have identified significant differences between the groups (Gordon et al., 1986; Fehr and List, 2004; Carpenter et al., 2008). As one study put it, "problems exist in replicating with nonstudent subjects behavioral phenomena observed in student samples" (Gordon et al., 1986). 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,[16] a population that consists of 70% of citizens over the age of 18 (Census Bureau, 2009, Table 2-1). 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 et al., 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. 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.[17] 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 listed 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 consistently collected (excluding Virginia and all non-southern states).[18] 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). 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.[19] 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.[20] 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. We define our *black* indicator variable as equal to one if the respondent indicated that she is African American on the voter registration form and zero otherwise.[21] Since we obtain this information from the voter file data, rather than from the respondent directly, the respondent's race is not being explicitly cued in the experiment.

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

---

[15] See Levitt and List (2009) for a discussion of issues related to typical experimental samples.

[16] North Dakota does not require citizens to register to vote.

[17] 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.

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

[19] 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.

[20] 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.

[21] 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.

Please cite this article in press as: Griffin, J., et al., Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game. J. Econ. Behav. Organ. (2012), http://dx.doi.org/10.1016/j.jebo.2012.09.010

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 unlikely to know such information as their Census block group's median household income. Moreover, our sample is considerably larger than the sample that could be drawn using similar restrictions from the KnowledgeNetworks panel.[22]

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.[23] 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 $2000; (b) the median plus or minus $1000; and (c) $2000 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 groups are larger and may contain distinct neighborhoods or sub-neighborhoods. Drawing loosely on the neighborhood tipping point literature (Card et al., 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 Table 2A and B 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 fail 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

[22] 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.

[23] In 2007, the median household income for non-Hispanic whites was $55,530 compared to $34,218 for non-Hispanic blacks (DeNavas-Walt et al., 2008).

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.

	Contacted	Not contacted
A. Contacted versus non-contacted subsamples		
Black	0.49	0.55[a]
Female	0.59	0.57
Age	58.68	51.86[a]
Median HH income	35,536	37,733[a]
Mean years of educ.	11.68	11.80[a]
Stakes	42.0	41.8
Offer	4.01	4.00
Accept	0.35	–
N	2003	5666
	Participated	**Did not participate**
B. Participants versus non-participants in the contacted subsample		
Black	0.50	0.48
Female	0.60	0.55
Age	58.3[a]	60.44[a]
Median HH income	35,526	35,581
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).
[a] Indicates different from the mean in the contacted (participated) subsample at the 5% level.

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% 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.

4. A model of respondent choice

We assume that an accepted offer generates utility for our respondents according to the following model:

$$v(g) = o_g - \gamma_1 relative_g - \gamma_2 absolute_g$$
$$\text{where} \quad relative_g = \left(0.5 - \frac{o_g}{s_g}\right) \quad \text{and} \quad absolute_g = (s_g - 2o_g). \tag{1}$$

The vector g contains two elements with information on the game variables: o_g, the amount of the proposed offer and s_g, the stakes to be split with the responder receiving o_g. o_g and s_g are subject to the constraint that $o_g \le s_g - o_g$. The model is a hybrid of Fehr–Schmidt (1999) and Bolton–Ockenfels (2000). The first term in (1) is utility gained directly from the offer, while the second and third terms represent penalties from receiving an unequal distribution of funds. The $absolute_g$ term reflects a utility penalty to the responder due receiving less than the proposer in nominal terms, following Fehr–Schmidt (1999). Following Bolton–Ockenfels (2000), the $relative_g$ term reflects utility loss due to receiving less in percentage terms. We express this as deviation from a focal point of 0.5, which represents an even split. We assume the utility associated with rejection is zero. Our experimental manipulation of both stakes and offer size allows us to separately identify the two inequality aversion parameters.

The individual's choice problem is to accept or reject the offer in order to maximize utility. We assume individuals choose $y \in \{0, 1\}$, where 1 indicates acceptance, to maximize u subject to a decision-making error that follows an extreme-value distribution:

$$u(y; g) = y * v(g) + \varepsilon_g \tag{2}$$

It follows that,

$$Pr(y = 1|X) = Pr(v(g) > -\varepsilon) = \frac{\exp(\alpha' X)}{1 + \exp(\alpha' X)},$$

where $\alpha' X$ is a linear equation describing determinants of the decision to accept. Since our experiment randomly assigns all arguments of the utility function, we can estimate the following logit model of the acceptance decision:

$$y = \alpha_1 o_g + \alpha_2 relative_g + \alpha_3 absolute_g + \varepsilon_g \tag{3}$$

From this it is straightforward to estimate the parameters of the utility function as $\gamma_1 = \alpha_2/\alpha_1$ and $\gamma_2 = \alpha_3/\alpha_1$.

A main question of interest is whether blacks and whites respond differently when faced with various combinations of offer, absolute inequality, and relative inequality. Therefore we are interested in allowing for heterogeneity in the parameters of the utility function across races. Specifically:

$$v_{i=B,W}(g) = o_g - \gamma_{1i=B,W} relative_g - \gamma_{2i=B,W} absolute_g \tag{4}$$

To examine this heterogeneity empirically, we first estimate a logit model allowing for interactions with the utility components with race:

$$y_i = \beta_1 o_g + \beta_2 relative_g + \beta_3 absolute_g + \beta_4 black_i * offer_g + \beta_5 black_i * relative_g + \beta_6 black_i * absolute_g + \varepsilon_{ig} \tag{5}$$

From this we can derive the following parameters of the utility function allowing for racial heterogeneity:

$$\gamma_{1W} = -\frac{\beta_2}{\beta_1}, \quad \gamma_{2W} = -\frac{\beta_3}{\beta_1}, \quad \gamma_{1B} = -\frac{\beta_2 + \beta_5}{\beta_1 + \beta_4}, \quad \text{and} \quad \gamma_{2B} = -\frac{\beta_3 + \beta_6}{\beta_1 + \beta_4}. \tag{6}$$

Note that this derivation requires that β_1 and $\beta_1 + \beta_4$ are non-zero. We discuss the empirical support for this requirement in the following section.

Please cite this article in press as: Griffin, J., et al., Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game. J. Econ. Behav. Organ. (2012), http://dx.doi.org/10.1016/j.jebo.2012.09.010

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.

5. Results

5.1. 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.

5.2. Graphical analysis

We begin by presenting a graphical version of our basic analysis in Fig. 1a–e. 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. Fig. 1a reports mean acceptance rates by race at all values of share in our experiment. Acceptance rates range from about 0.25 at lower shares to about 0.5 at higher shares. This is reassuring, since it suggests that our randomly assigned stakes-offer combinations span a range over which there is potential for changes in stakes or offer size to lead to changes in acceptance.

Fig. 1b–e shows acceptance rates disaggregated across the four nominal offer amounts in our experiment. Acceptance rates rise with share conditional on offer size for both races, with a possible exception when only $1 is offered. This suggests respondents receive less utility from relatively unequal offers, even when the nominal gains are equal across offers. The figures also show that blacks have uniformly higher acceptance rates than whites at higher offers of $5 and $10. This suggests that blacks and whites may differ in their responses to offer size, conditional on the share the offer represents.

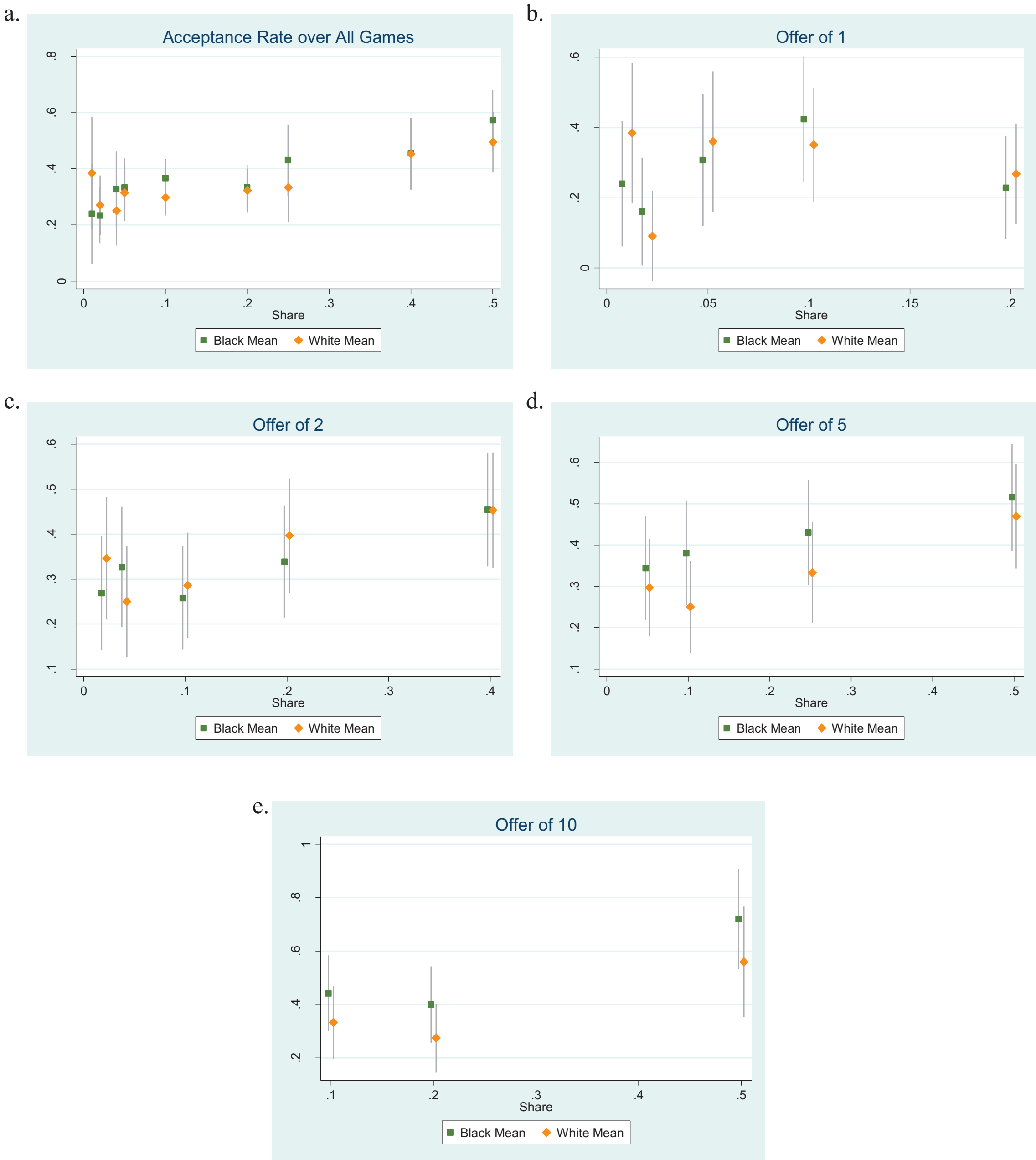

Fig. 1. (a–e) Shows mean acceptance rates by share separately for blacks (squares) and whites (diamonds). Vertical lines show 95% confidence intervals around mean estimates.

5.3. Choice model estimation

5.3.1. Analysis of choice model results

We first carry out several model selection tests to determine which specifications best represent the data. Specifically, we estimate variants of Equation 5 in which one or more parameters are constrained to be equal to zero. We summarize the models we tested and report the results from specification tests in Table A1. We report both *F*-tests and LR tests that various models are nested within one another. In all models where both absolute and relative inequality terms are included, we are unable to reject a more parsimonious model that excludes absolute inequality. An alternative to including absolute inequality directly is to control for the stakes amount using dummy variables, as in Andersen et al. (2011). This avoids the problem of interpretation inherent to specifications where relative and absolute inequality are co-determined. We cannot

Table 4
Linear probability models of acceptance choice.

	Respondent race interactions omitted				Respondent race interactions included			
Offer	0.007*	0.008*	0.006	0.007	−0.001	−0.001	0.00	−0.002
	(1.66)	(1.85)	(1.56)	(1.64)	(0.22)	(0.12)	(0.01)	(0.42)
Relative	−0.448***	−0.454***	−0.508***	−0.438***	−0.410***	−0.392***	−0.429***	−0.434***
	(5.54)	(5.64)	(6.36)	(5.36)	(4.47)	(4.29)	(4.85)	(3.75)
Black * offer					0.016**	0.017***	0.013**	0.018**
					(2.51)	(2.63)	(2.08)	(2.22)
Black * relative					−0.076	−0.124	−0.163*	−0.008
					(0.90)	(1.44)	(1.95)	(0.05)
Additional controls. . .								
Tract controls		X	X			X	X	
Game conditions			X				X	
Strata dummies				X				X
R^2	0.03	0.06	0.23	0.03	0.03	0.06	0.23	0.04
Observations	1647	1598	1500	1647	1647	1598	1500	1647

Notes: Dependent variable is indicator for acceptance of proposed split. All equations estimated via OLS. Absolute value of robust *t*-statistics in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

reject that the more parsimonious model in [6] fits the data as well as the Andersen et al. model (specification [7]) under either the *F*-test or the LR test.[24]

We conclude that, conditional on relative inequality, neither absolute inequality nor stakes affect a respondent's acceptance decision. We therefore emphasize specifications in which only the offer size and relative inequality terms are included. Our preferred model is therefore specification [6] in Table A1. As indicated by the tests at the bottom of the table, we cannot reject that specification [6] is nested in the full version of Equation 5. However, we *can* reject that a model that excludes interactions of the game variables with the black respondent indicator (as in [1]) is nested in [6].

Our first estimates of acceptance choice models are presented in Table 4. Although we use logit models in our model selection (and later in our calculation of the parameters in Eq. (6)), Table 4 presents results from linear probability models of the acceptance decision. We experimented with both probit and logit non-linear models of the acceptance decision. We find that the choice of functional form in the estimating equation makes no qualitative and almost no quantitative difference for our results. This is expected given that our dependent variable has a mean well within the unit interval. On the other hand, calculating the impact of interaction terms in a non-linear model is much less straightforward than in a linear model (Norton et al., 2004). We therefore report the results from the linear model for ease of interpretation.

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 Eq. (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.[25] 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.[26] 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 are presented in Table 4. 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.[27] These respondents are therefore omitted from the estimation when these variables are included. For this reason, our preferred specification includes the strata fixed effects. The strata fixed effects specification includes all respondents yet controls for group differences across our race-income-neighborhood diversity cells.

24 The Andersen et al. (2011) study found significant effects of stakes, but these were at much higher stakes levels in real terms than the ones we use in our study.

25 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.

26 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.

27 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.

Table 5
Estimates of utility function parameters under assumption of a uniform offer effect, $\beta_4 = 0$.

Specification	No racial heterogeneity		Racial heterogeneity allowed	
γ_1 (γ_{1W} under racial heterogeneity)	62.10 (40.64) $p = 0.13$	119.54 (134.42) $p = 0.37$	56.39 (39.05) $p = 0.15$	109.90 (127.46) $p = 0.39$
γ_{1B}			67.65 (45.21) $p = 0.14$	129.32 (146.13) $p = 0.38$
γ_1 (γ_{1B} under racial heterogeneity)		−0.15 (0.22) $p = 0.53$		−0.14 (0.25) $p = 0.58$
γ_{2B}				−0.15 (0.26) $p = 0.55$

Notes: Estimates computed using Stata's nlcom command following logit estimates of Eq. (5) (including or excluding absolute inequality terms, as indicated). Strata dummies included and $N = 1647$ in all specifications.

There is a high degree of consistency across the first four columns in Table 4. 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 effect of a 10% point decrease in *relative* is a decrease in the likelihood of acceptance of more than four percentage points. By contrast, *offer* has a significant but modest positive effect on the likelihood of acceptance.

This changes when the effects of *offer* and *relative* are allowed to vary by race of the respondent. The estimates are again stable across the four sets of additional controls, but they differ from those in the first four columns of Table 4. The right four columns of Table 4 show that the main effect of *offer* is smaller than in the non-interacted specifications and is insignificantly different from zero. In the non-linear versions of this model, β_1 in Eq. (5) is estimated to be zero. Also, coefficients on *black * offer* show that blacks and whites differ significantly in their response to changes in the offer amount, holding relative constant. Together these estimates imply that $\beta_1 + \beta_4$ in Eq. (6) is non-zero and positive in the analogous non-linear model. The effects of *offer* on the likelihood of acceptance among black respondents are also economically large. Estimates from our preferred specification, shown in the rightmost column of Table 4, show that an extra dollar in the offer increases the likelihood of a black respondent accepting by almost 2% points, an effect that is roughly half that of a 10% point increase in *relative*. The coefficient on *relative* is unaffected by the addition of the *relative * black* interaction term. This means that blacks and whites respond 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.

Our interpretation of Table 4 results is that the utility function in Eq. (4) appropriately describes behavior for blacks but not for whites, since our estimates indicate that the utility function parameters in Eq. (6) are undefined for whites. Nevertheless, in order to roughly compare our estimates with others in the literature, we calculate the parameters in Eq. (6) under the constraint that $\beta_4 = 0$ in Eq. (5). Although our model testing indicated that absolute inequality is unimportant for acceptance decisions, we lift the restriction that $\beta_3 = \beta_6 = 0$ so that we can also generate estimates for γ_2 in models both with and without racial heterogeneity.

The results are reported in Table 5. Because the effect of the two forms inequality is expressed relative to the effect of offer size, these estimates can be interpreted as indicating willingness-to-pay to avoid inequality. In a model without racial heterogeneity and omitting absolute inequality, we estimate that γ_1 is equal to 62.1. This is nearly significant, with a *p*-value of 0.13. When we add absolute inequality to the model, our estimate of γ_1 doubles but becomes more imprecise. Our estimate of γ_2 is economically trivial and insignificantly different from zero. Models that allow for racial heterogeneity show estimates of γ_1 andγ_2 that are statistically indistinguishable between blacks and whites. This is not surprising given that Table 4 shows responses to relative inequality that do not differ across races and that we imposed equal offer effects across blacks and whites. Responses to absolute inequality are small and statistically insignificant for both blacks and whites. These are reflected in small and insignificant estimates of γ_2 for each group. We do not believe these parameter estimates accurately reflect racial differences in responses to the ultimatum game because, as shown above, the data reject the utility function model imposed to estimate these parameters in our sample of whites. For blacks, the parameter estimates are valid. However, the parameter estimates we obtain are broadly consistent with others in the literature. Bellemare et al. (2008) and Goeree and Holt (2000) estimate models similar to ours but excluding a separate role for relative inequality.[28] Those studies find an inequity aversion parameter regarding one's own disadvantage of 1.58 (s.e. = 0.61) in Bellemare et al. and 0.84 (s.e. = 0.16). Given the large standard errors on our estimates, these are of the same magnitude as those we report in

[28] Models in both Bellemare et al. (2008) and Goeree and Holt (2000) constrain γ_1 to be zero in Eq. (1). However, models in both papers allow absolute inequality that favors the proposer to have a distinct role from inequality that favors the responder. In our experiment, no responders received offers in their favor, so we do not model this type of inequity aversion.

Please cite this article in press as: Griffin, J., et al., Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game. J. Econ. Behav. Organ. (2012), http://dx.doi.org/10.1016/j.jebo.2012.09.010

Table 6
Linear probability models of acceptance choice by subgroup.

Subsample	Q3 consistent	Q3 not consistent	Large stakes	Small stakes	Women	Men	Younger than 55	Older than 55
Offer	−0.002	−0.007	−0.002	0.005	−0.01	0.007	0.003	−0.003
	(0.23)	(1.44)	(0.19)	(0.13)	(1.38)	(0.76)	(0.31)	(0.49)
Relative	−1.081***	0.135	−0.433*	−0.341	−0.381**	−0.514***	−0.522***	−0.344**
	(8.37)	(1.47)	(1.79)	(0.88)	(2.47)	(2.82)	(2.80)	(2.33)
Black * offer	0.021*	0.019*	0.015	0.015	0.016	0.025*	0.017	0.020*
	(1.84)	(1.93)	(1.35)	(0.29)	(1.58)	(1.80)	(1.34)	(1.90)
Black * relative	0.082	0.300*	−0.144	0.014	−0.021	0.008	−0.047	0.046
	(0.45)	(1.70)	(0.43)	(0.03)	(0.10)	(0.03)	(0.19)	(0.21)
R^2	0.13	0.08	0.04	0.04	0.04	0.08	0.07	0.04
Observations	951	663	1,116	531	992	655	715	932

Notes: Q2 is the respondent's choice of whether to accept or reject the ultimatum offer. Dependent variable is indicator for acceptance of proposed split. 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. All equations estimated via OLS. Absolute value of robust *t*-statistics in parentheses.

* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Table 5. A difference is that those studies focus on absolute inequality, while our study shows that relative inequality is the more important dimension when both can be controlled.

5.3.2. Understanding the offer effect and robustness checks

In this section, we investigate the robustness of our finding that the effect of *offer* differs between black and white respondents. 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 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.[29] We then re-estimate our preferred specification from the final column of Table 4 on these subsamples. 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 points emerge from this analysis. First, respondents who gave inconsistent answers to Q3 exhibit a response to changes in relative inequality that is wrong-signed and not statistically significant. Given that their response to increases in relative inequality differs markedly from that of respondents who provided consistent answers, we surmise that an inconsistent response to Q3 means that the respondent did not understand the game. Also, the coefficient on relative inequality among those with consistent responses is more than double its size in the full sample. Alternatively, these may be subjects for whom the game had very low salience, potentially because of the cold call setting as discussed in Section 2. These respondents may have put little though into answers they gave.[30]

However, the second point to take from our subgroup analysis is that the offer effect identified for blacks in Table 4 is not driven by individuals with inconsistent Q3 responses. The first two columns in Table 6 show that the impact of *offer* on blacks is insignificantly different across the two Q3-based subsamples. The main consequence of poor game comprehension, as indicated by the Q3 response, seems to be atypical responses to relative inequality. The response to increasing nominal offer size is unaffected.

[29] 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.

[30] The coefficient on *relative* in the Q3 consistent sample is estimated to be (statistically) greater than one. This is an out of sample outcome, but increments of 1 in the share are also out of sample. Non-linear models give a similar out of sample prediction in this subsample. We simply interpret this to mean that very small shares are accepted with very low probability while shares near one are accepted with very high probability. This is consistent with the literature.

Please cite this article in press as: Griffin, J., et al., Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game. J. Econ. Behav. Organ. (2012), http://dx.doi.org/10.1016/j.jebo.2012.09.010

Table 7
Linear probability models of acceptance choice by neighborhood income strata.

Sample	Neighborhood median income at black 10th percentile	Neighborhood median income at black median	Neighborhood median income at black 90th percentile
Offer	−0.005	0.007	−0.006
	(0.58)	(0.80)	(0.60)
Relative	−0.347**	−0.542***	−0.316*
	(2.34)	(3.47)	(1.77)
Black * offer	0.033***	−0.006	0.02
	(3.22)	(0.54)	(1.61)
Black * relative	−0.211	0.178	−0.231
	(1.48)	(1.24)	(1.44)
R^2	0.05	0.03	0.03
Observations	614	560	473

Notes: Dependent variable is indicator for acceptance of proposed split. All equations estimated via OLS. Models include only listed covariates. Absolute value of robust *t*-statistics in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

Interestingly, a probit analysis of the determinants of a valid Q3 response turned up no significant predictors.[31] Although it is somewhat puzzling that background variables like neighborhood income and education level do not predict game comprehension as measured by Q3, 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 we would find no significant predictors of Q3 response.[32]

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 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 view this as 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 is that our results are robust to varying levels of salience (high or low stakes). Specifically, the effects of relative inequality and of offer size among blacks are very similar across respondents who received large versus small stakes, and these in turn are statistically indistinguishable from those in our preferred estimates from Table 4. 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.[33]

To further investigate the source of the offer effect for blacks, 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 from this exercise are presented the rightmost four columns of Table 6. These report results from separate estimation of our acceptance model on subsamples defined 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 age and gender groups. The effects are sometimes insignificant, as sample sizes are reduced in the subgroup estimation, but the point estimates are stable and very similar to those in our preferred Table 4 specification.[34]

As a final cut of the data, Table 7 shows the results of estimating our models separately on subsamples defined by income, using the three income categories used to define our sample strata. We omit strata fixed effects from the models in Table 7 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.

[31] 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.

[32] It is also possible that there was a subset of interviewers at the survey firm who were less able to get a respondent's full attention. 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.

[33] We also estimated our acceptance model separately on respondents who received offers representing atypically low shares versus typical shares of 0.25–0.5, which are defined as typical in (Forsythe et al., 1994). Again we find no qualitative difference in results across these subgroups. We conclude that our results are not driven by responses to atypically low shares.

[34] 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.

Please cite this article in press as: Griffin, J., et al., Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game. J. Econ. Behav. Organ. (2012), http://dx.doi.org/10.1016/j.jebo.2012.09.010

Table 8
Linear probability model of acceptance choice with proposer race interactions.

Sample	Full	Neighborhood median income at black 10th percentile	Neighborhood median income at black median	Neighborhood Median income at black 90th percentile
Offer	0.00	−0.012	0.011	0.005
	(0.01)	(0.98)	(0.93)	(0.30)
Relative	−0.522***	−0.607**	−0.559**	−0.297
	(3.48)	(2.46)	(2.14)	(1.04)
Black * offer	0.016	0.043**	0.008	−0.007
	(1.38)	(2.27)	(0.40)	(0.29)
Black * relative	0.335	0.274	0.554	0.128
	(1.47)	(0.76)	(1.43)	(0.28)
Black	−0.15	−0.266	−0.095	−0.078
	(1.47)	(1.59)	(0.54)	(0.39)
Proposer race (PR)	−0.021	−0.093	0.177	−0.128
	(0.20)	(0.55)	(0.96)	(0.66)
PR * black	0.212	0.217	0.053	0.453
	(1.45)	(0.92)	(0.21)	(1.65)
PR * offer	−0.005	0.003	−0.015	−0.008
	(0.46)	(0.16)	(0.81)	(0.35)
PR * relative	0.193	0.265	−0.09	0.365
	(0.83)	(0.71)	(0.22)	(0.83)
PR * black * offer	0.005	0.001	−0.017	0.03
	(0.34)	(0.06)	(0.62)	(0.95)
PR * black * relative	−0.662**	−0.39	−0.506	−1.274**
	(2.04)	(0.76)	(0.89)	(2.03)
R^2	0.03	0.06	0.04	0.05
Observations	1647	614	560	473

Notes: Variable PR = 1 if proposer race is implied to be black, zero if implied to be white. Dependent variable is indicator for acceptance of proposed split. All equations estimated via OLS and absolute value of robust *t*-statistics in parentheses.
* Significant at 10%.
** Significant at 5%.
*** Significant at 1%.

The results show important differences across neighborhood income strata in the response to offer size. The main effect for *offer* is economically and statistically insignificant in all three groups, consistent with our preferred Table 4 estimate. The effect of relative inequality is also reasonably stable across neighborhood income groups and similar to Table 4 estimate. However, the effect of *offer* for blacks is large and statistically significant only among blacks in the lowest neighborhood income group. The offer effect for blacks is not present in either of the two neighborhood income groups. In unreported results, we interact neighborhood income group with all variables in Table 7 specifications. We find that the offer effect for blacks in the lowest neighborhood income group is statistically different from that for blacks in the higher income groups at the 1% level.

We conclude that the offer effect for blacks identified in Table 4 is driven by blacks in the lowest income group.[35] 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.

5.3.3. *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 Section 2, this was done through the use of racially distinct male names.[36] We explore the effect of proposer race further by adding interactions with proposer race to our preferred model from Table 4, excluding the strata fixed effects. We create a dummy variable equal to one if the name of the hypothetical proposer was a high frequency black name.[37] The results are presented in Table 8. Because we found important heterogeneity in the impact of *offer* by neighborhood income strata, we also estimate the model with proposer race interactions on the subgroups in Table 7.

[35] 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.
[36] 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.
[37] 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, but their interpretation is less transparent.

Many of our earlier results carry over to this expanded specification. First, the main effect of *offer* is small and statistically insignificant both in the full sample and in the separate income group samples. Second, the effect of relative inequality is of a similar magnitude to results in earlier tables and is significant in all but one subsample. And third, the significant offer effect for blacks is still limited to the lowest neighborhood income group. In the expanded model, the effect of the *black* * *offer* interaction is economically trivial and statistically insignificant in both higher income groups.

The new results in Table 8 concern the effects of proposer race. The main effect of and interactions with proposer race are insignificant in all but one case: the interaction of black, proposer race, and relative inequality. The sign on this triple interaction is consistently negative and the impacts large in magnitude across all samples in Table 8. The estimates indicate that sensitivity to relative inequality about doubles among blacks facing black proposers. Effects might be somewhat larger among blacks in the highest neighborhood income group. We conclude that blacks in general are more sensitive to relative inequality imposed by other blacks.

6. Conclusion

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 by the behavior of blacks in the lowest neighborhood 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 a more general effect of implied proposer race on black respondents. Specifically, blacks are about twice as sensitive to relative inequality (defined as the effect of an increase in relative inequality of a proposal) when the perceived race of the proposer is black. This is true for blacks of all neighborhood income levels in our sample. This is consistent with blacks being more sensitive to being treated unfairly by whites than by blacks, a finding consonant with results in Fershtman and Gneezy (2001) and Mendes et al. (2008).

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 example, Charles et al. (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 share, as shown in the panels of Fig. 1. 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 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 et al. (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.

Please cite this article in press as: Griffin, J., et al., Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game. J. Econ. Behav. Organ. (2012), http://dx.doi.org/10.1016/j.jebo.2012.09.010

Table A1
Model selection.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Offer	X	X	X	X	X	X	X
Relative	X		X	X	X	X	X
Absolute		X	X	X	X		
Black				X	X		
Black * offer					X	X	X
Black * relative					X	X	X
Black * absolute					X		
Stakes dummies included	No	No	No	No	No	No	Yes
H_0: absolute = 0		$p = 0.001$	$p = 0.24$	$p = 0.24$			
H_0: black = 0				$p = 0.50$			
H_0: absolute = 0, black = 0, and black * absolute = 0					$p = 0.49$		
H_0: Stakes dummies jointly = 0							$p = 0.26$
LR test that [6] nested in [5]						$p = 0.29$	
LR test that [1] nested in [3]	$p = 0.27$						
LR test that [3] nested in [4]			$p = 0.13$				
LR test that [1] nested in [4]	$p = 0.18$						
LR test that [1] nested in [6]	$p = 0.028$						
LR test that [6] nested in [7]						$p = 0.28$	

Notes: All models estimated via logit and include strata dummies. $N = 1647$. Strata dummies excluded from LR test specifications to force degrees of freedom to differ.

Acknowledgements

We thank Bill Evans, Dan Hungerman, Lars Lefgren, Sandra Black, Jennifer Richeson, and seminar participants at Princeton University, Vanderbilt University, and the University of Notre Dame for helpful comments. Wozniak thanks the Industrial Relations Section at the Princeton Economics Department for financial support during the course of this project. All errors are our own.

Appendix A.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jebo.2012.09.010.

References

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, 385–388.
Alesina, A., La Ferrara, E., 2005. Preferences for redistribution in the land of opportunities. Journal of Public Economics 89, 897–931.
Andersen, S., Seda, E., Gneezy, U., Hoffman, M., List, J., 2011. Stakes matter in ultimatum games. American Economic Review 101, 3427–3439.
Andreoni, J., Vesterlund, L., 2001. Which is the fair sex? Gender differences in altruism. Quarterly Journal of Economics 116, 293–312.
Atkinson, A.B., 1970. On the measurement of inequality. Journal of Economic Theory 2, 244–263.
Ayres, I., Siegelman, P., 1995. Race and gender discrimination in bargaining for a new car. American Economic Review 85, 304–321.
Babcock, L., Laschever, S., 2003. Women Don't Ask: Negotiation and the Gender Divide. Princeton University Press, Princeton, New Jersey.
Bartels, L., 2008. Unequal Democracy: The Political Economy of the New Gilded Age. Princeton University Press, Princeton.
Bellemare, C., Kröger, S., van Soest, A., 2008. Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. Econometrica 76, 815–839.
Benjamin, D.J., Choi, J.J., Strickland, A., 2010. Social identity and preferences. American Economic Review 100, 1913–1928.
Bertrand, M., Mullainathan, S., 2004. Are Emily and Brendan more employable that Lakisha and Jamal? Evidence on racial discrimination in the labor market from a large randomized experiment. American Economic Review 94, 991–1013.
Blount, S., 1995. When social outcomes aren't fair: the effect of causal attributions on preferences. Organizational Behavior and Human Decision Processes 63, 131–144.
Bolton, G., Katok, E., 1995. An experimental test for gender differences in beneficent behavior. Economics Letters 48, 287–292.
Bolton, G., Ockenfels, A., 2000. ERC: a theory of equity, reciprocity, and competition. American Economic Review 90, 166–193.
Bolton, G., Brandts, J., Ockenfels, A., 2005. Fair procedures: evidence from games involving lotteries. Economic Journal 115, 1054–1076.
Cameron, L.A., 1999. Raising the stakes in the ultimatum game: experimental evidence from Indonesia. Economic Inquiry 37, 47–59.
Canon, D., 1999. Race, Redistricting, and Representation: The Unintended Consequences of Majority Black Districts. University of Chicago Press, Chicago.
Card, D., Mas, A., Rothstein, J., 2008. Tipping and the dynamics of segregation. Quarterly Journal of Economics 123, 177–218.
Carpenter, J., Connolly, C., Myers, C.K., 2008. Altruistic behavior in a representative dictator experiment. Experimental Economics 11, 282–298.
Charles, K., Hurst, E., Roussanov, N., 2009. Conspicuous consumption and race. Quarterly Journal of Economics 124, 425–467.
Chen, K., Tang, F., 2009. Cultural differences between Tibetans and ethnic Han Chinese in ultimatum bargaining experiments. European Journal of Political Economy 25, 78–84.
Chuah, S., Hoffman, R., Jones, M., Williams, G., 2007. Do cultures clash? Evidence from cross-national ultimatum game experiments. Journal of Economic Behavior and Organization 64, 35–48.
Dawson, M., 1995. Behind the Mule: Race and Class in African-American Politics. Princeton University Press, Princeton.

DeNavas-Walt, C., Proctor, B.D., Smith, J.C., 2008. Income, Poverty, and Health Insurance Coverage in the United States: 2007. US Census Bureau, US Department of Commerce.
Doty, R., Silverthorne, C., 1975. Influence of menstrual cycle on volunteering behavior. Nature 254, 138–140.
Eckel, C., Grossman, P., 1998. Are women less selfish than men? Evidence from dictator experiments. Economic Journal 108, 726–735.
Eckel, C., Grossman, P., 2001. Chivalry and solidarity in ultimatum games. Economic Inquiry 39, 171–188.
Falk, A., Fehr, E., Fischbacher, U., 2005. Driving forces behind informal sanctions. Econometrica 73, 2017–2030.
Fehr, E., List, J., 2004. The Hidden Costs and Returns of Incentives – Trust and Trustworthiness among CEOs. IEW Working Papers 134.
Fehr, E., Schmidt, K., 1999. A theory of fairness, competition, and cooperation. Quarterly Journal of Economics 119, 817–868.
Ferraro, P., Cummings, R., 2007. Cultural diversity, discrimination, and economic outcomes: an experimental analysis. Economic Inquiry 45, 217–232.
Fershtman, C., Gneezy, U., 2001. Discrimination in a segmented society: an experimental approach. Quarterly Journal of Economics 116, 351–377.
Fix, M., Turner, M. (Eds.), 1998. A National Report Card on Discrimination in America: The Role of Testing. Urban Institute Press, Washington, DC.
Fong, C., Luttmer, E.F.P., 2011. Do race and fairness matter in generosity? Evidence from a nationally representative charity experiment. Journal of Public Economics 95, 372–394.
Forsythe, R., Horowitz, J., Savin, N., Sefton, M., 1994. Fairness in simple bargaining experiments. Games and Economic Behavior 6, 347–369.
Gerber, A., Green, D., 2000. The effects of canvassing, direct mail, and telephone contact on voter turnout: a field experiment. American Political Science Review 94, 653–663.
Gerber, A., Karlan, D., Bergan, D., 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, 35–52.
Geronimus, A., Bound, J., Neidert, L., 1996. On the validity of using census geocode characteristics to proxy individual socioeconomic characteristics. Journal of the American Statistical Association 91, 529–537.
Gilens, M., 1999. Why Americans Hate Welfare. University of Chicago Press, Chicago.
Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in competitive environments: gender differences. Quarterly Journal of Economics 118, 1049–1074.
Goeree, J., Holt, C., 2000. Asymmetric inequality aversion and noisy behavior in alternating-offer bargaining games. European Economic Review 44, 1079–1089.
Gordon, M., Slade, L., Schmitt, N., 1986. The 'science of the sophomore' revisited: from conjecture to empiricism. Academy of Management Review 11, 191–207.
Green, D., Gerber, A., 2006. Can registration-based sampling improve the accuracy of midterm election forecasts? Public Opinion Quarterly 70, 197–223.
Guth, W., Schmittberger, R., Schwarze, B., 1982. An experimental study of ultimatum bargaining. Journal of Economic Behavior and Organization 3, 367–388.
Harrison, G., List, J., 2004. Field experiments. Journal of Economic Literature 42, 1009–1055.
Henrich, J., 2000. Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. American Economic Review 90, 973–979.
Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., McElreath, R., 2001. In search of homo economicus: behavioral experiments in 15 small scale societies. American Economic Review 91, 73–78.
Jamison, J., Karlan, D., Schechter, L., 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, D., Zinman, J., 2009. Observing unobservables: asymmetries with a consumer credit field experiment. Econometrica 77, 1993–2008.
Kravitz, D., Gunto, S., 1992. Decisions and perceptions of recipients in ultimatum bargaining games. Journal of Socio-Economics 21, 65–84.
Levitt, S., List, J., 2009. Field experiments in economics: the past, the present, and the future. European Economic Review 53, 1–18.
Lü, X., Scheve, K., Slaughter, M.J., 2012. Inequity aversion and the international distribution of trade protection. American Journal of Political Science 56 (3), 638–655.
Luttmer, E.F.P., 2001. Group loyalty and the taste for redistribution. Journal of Political Economy 109 (3), 500–528.
McClellan, M., Skinner, J., 1999. Medicare reform: who pays, and who benefits? Health Affairs 18 (1), 48–62.
Mendes, W., Major, B., McCoy, S., Blasovich, J., 2008. How attributional ambiguity shapes physiological and emotional responses to social rejection and acceptance. Journal of Personality and Social Psychology 94 (2), 278–291.
Murnighan, J., Saxon, M., 1998. Ultimatum bargaining by children and adults. Journal of Economic Psychology 19, 415–445.
Nickerson, D., 2008. Is voting contagious? Evidence from two field experiments. American Political Science Review 102 (February), 49–57.
Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? Quarterly Journal of Economics 122 (3), 1067–1101.
Norton, E., Wang, H., Ai, C., 2004. Computing interaction effects and standard errors in logit and probit models. The Stata Journal 4 (2), 154–167.
Oosterbeek, H., Sloof, R., van de Kuilen, G., 2004. Cultural differences in ultimatum game experiments: evidence from a meta-analysis. Experimental Economics 7, 171–188.
Roth, A., Prasnikar, V., Okuno-Fujiware, M., Zamir, S., 1991. Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study. American Economic Review 81, 1068–1095.
Slonim, R., Roth, A., 1998. Learning in high stakes ultimatum games: an experiment in the Slovak Republic. Econometrica 66 (3), 569–596.
Solnick, S., 2001. Gender differences in the ultimatum game. Economic Inquiry 39 (2), 189–200.

Please cite this article in press as: Griffin, J., et al., Racial differences in inequality aversion: Evidence from real world respondents in the ultimatum game. J. Econ. Behav. Organ. (2012), http://dx.doi.org/10.1016/j.jebo.2012.09.010