Unfair rules for unequal pay: Wage discrimination and procedural justice

Scott Bokemper
Yale University

Peter DeScioli
Stony Brook University

Reuben Kline
Stony Brook University

Abstract

Do people judge some forms of wage discrimination to be more unfair than others? We report an experiment in an online labor market in which participants were paid based on discriminatory rules. We test hypotheses about fairness based on procedural justice, divisiveness, and affective polarization between partisans. Workers transcribed text and then learned that they earned more or less money than other workers for doing the same job. We manipulated whether the unequal pay was based on their political party, eye color, or an arbitrary choice between two doors. Consistent with the divisiveness hypothesis, participants judged discriminatory pay to be less fair when it was based on a stable characteristic, political party or eye color, compared to a transient choice (between doors). We find mixed evidence about how affective polarization exacerbates the unfairness of partisan discrimination. We discuss implications for the procedural justice of wage discrimination.

Introduction

Imagine that you work all day long laying bricks, filing paperwork, or grading exams only to find out that your employer paid someone else twice as much money for the same job. Many people in this position feel that they have been treated unfairly (Fehr and Schmidt 1999; Sweeney 1990; Sweeney et al. 1990). However, workers might be more accepting of lower pay if they judge the procedure for determining pay to be fair, such as paying higher wages to more qualified, talented, or senior employees. In contrast, people show outrage and political activism when procedures are unfair such as paying lower wages to women and minorities (Blau and Kahn 1992; Blau and Kahn 2000; Leicht 2008). People's sense of procedural justice also shapes legislation such as equal employment laws that prohibit certain types of pay discrimination. But questions remain about what makes a procedure fair or unfair, and whether some discriminatory rules are judged to be more unfair than others. Here we examine people's moral judgments about pay discrimination by using an online labor market to test what makes people perceive unequal pay as more or less unfair.

Procedural justice shapes how people react to the ways that leaders make decisions, police handle suspects, courts determine guilt, employers set wages, and other important events

in our personal, economic, and political lives (Dickson, Gordon, and Huber 2009, 2015; Gerber et al. 2012; Gibson 1989; Hibbing and Theiss-Morse 2001; Tyler and Blader 2003; Tyler, Casper, and Fisher 1989). Procedural justice pertains to the *process* by which decisions are made as opposed to the *outcome,* which is the purview of distributive justice.

We investigate procedural justice in the context of pay discrimination. Pay discrimination occurs when an employer deliberately pays some workers lower wages based on categories that are irrelevant for merit or performance such as race, gender, or age. Pay discrimination violates procedural justice because the employer uses a rule or procedure to determine pay that is not acceptable, particularly transgressing the expectation that procedures should be impartial (Lind and Tyler 1988; Shaw et al. 2012; Shaw 2013; Tyler and Blader 2003). In this way, pay discrimination is analogous to racial profiling in which the police decide who to search based on race. Indeed, previous research found that participants judged racial profiling by police or airport security to be an unfair violation of procedural justice (Sindhav et al. 2006; Tyler and Wakslak 2004). Moreover, for both wage discrimination and racial profiling, the *procedure*, i.e., treating people in one category differently than others, is distinct from the *outcome*, i.e., the distribution of wages or police searches. The procedural justice literature finds that people care about both procedures and outcomes, so this implies that for pay discrimination, people will care not only about the distribution of wages but also about whether the procedure for determining wages is impartial and fair.

The principle of equity, "equal pay for equal work," is a popular and intuitive notion rooted in our social psychology (Fehr and Schmidt 1999; Sweeney 1990; Sweeney et al. 1990). Nevertheless, inequalities in pay are common, and many people accept unequal wages when the differences result from procedures for paying workers based on experience, longevity, effort, talent, performance, market demand, or other criteria deemed fair within a given community. At the same time, people reject other possible rules such as paying lower wages to the elderly, a practice that is prohibited by U.S. law. Hence, for some inequalities people show outrage and political activism, whereas in other cases people accept or at least tolerate unequal pay. Yet, it is unclear which types of inequality people view as more unfair than others. That is, holding constant the inequality in pay for the same work, why do some forms of pay discrimination provoke more outrage than others?

We propose that a procedure's *divisiveness* is a key factor underlying fairness judgments. A rule or procedure is divisive when it consistently disfavors the same individuals, as opposed to rules that only temporarily or sporadically disfavor a particular group. The divisiveness hypothesis draws on ideas from evolutionary game theory. In a theoretical model, Grafen (1987) found that individuals most resist rules for solving coordination games when the rule consistently goes against their interests. These divisive rules resolve coordination problems based on the stable characteristics of players so that they consistently disadvantage the same individuals. For this reason, disadvantaged players resist divisive rules even if this leads to costly failures to coordinate. Grafen (1987) gave an example of a rule that benefits younger people at the expense of older people. Because the older people will never benefit from the rule, they are expected to strongly oppose it. In contrast, a rule that benefits older people at the expense of younger people is less strongly opposed by the young because they stand to benefit from the rule in the future.

2

Divisiveness could influence procedural justice because procedures often function to solve coordination problems (Camerer 2003; Chwe 2001; Cronk and Leech 2012; Schelling 1960; Thomas et al. 2014), such as when citizens need to agree on the legitimacy of a leader's decision or the guilt of a suspect. In general terms, a procedure can serve as a convention for resolving the coordination problems posed by disagreement. For instance, if an employer and worker disagree about the worker's wage, then they can use conventional procedures to resolve the impasse, such as rules that scale pay based on experience, performance, longevity, or outside offers.

The divisiveness hypothesis holds that the strategic logic of divisiveness shapes our intuitive judgments of fairness, such that people judge procedures as less fair when they consistently deprive the same individuals. Importantly, this does not mean that people consciously reason about coordination problems when they make judgments of fairness. Rather, the hypothesis is that the strategic logic of divisiveness shapes people's intuitions about fairness, analogous to how geometry guides our perceptions of depth by shaping visual cognition rather than through people's conscious calculations (Pinker, 1997). This theory builds on a growing literature arguing that the logic of coordination shapes the psychology of language (Lewis, 1969; Pinker, 2007), emotions (Reed et al. 2014; Tybur et al. 2013), reasoning about mental states (Thomas et al. 2014), moral judgment (DeScioli and Kurzban 2013), property rights (DeScioli and Wilson 2011; DeScioli and Karpoff 2015; Maynard Smith 1982), and leadership (Van Vugt 2006).

Finally, we examine whether people judge divisive rules to be even more unfair when they transgress a factional divide in society, specifically the tense opposition between Democrats and Republicans in the United States. Democrats and Republicans are prominent competing factions in American politics with a long history of conflict and antagonism. Moreover, researchers have argued that Americans have become increasingly polarized along party lines, including a rise in affective polarization with hostile feelings toward the opposing party (Green, Palmquist, and Schickler 2004; Iyenger, Sood, and Lelkes 2012; Mason 2015). This hostility even extends outside of politics, such as partisans who say they would disapprove if their son or daughter married someone who is a member of the opposing party (Iyengar, Sood, and Lelkes 2012). Given the antagonism between parties, people might judge partisan discrimination to be especially unfair because it inflames partisan conflict. Hence, people might judge that partisan discrimination is unfair not only because it is divisive, based on a stable characteristic, but also because it incites partisan hostility.

The present experiment
We test theories about unfair procedures by manipulating workers' actual wages for completing jobs in an online labor market. In different treatments, workers who did the same job received pay that was equal or unequal to other workers, and if unequal, they were either advantaged or disadvantaged by different payment rules. Workers first completed the job by transcribing text, next they learned the payment rule that determined wages, and then they judged the fairness of the payment rule.

In a between-subject design, we presented workers with one of three different unequal payment rules based on their choice of a door, eye color, or political partisanship. We selected the choice of a door (yellow or green) as a familiar, arbitrary decision like choosing between doors in a game show, which is transient choice rather than a stable characteristic of a person. We selected eye color as a stable characteristic of a person, and we selected political party (Democrat or Republican) as a stable characteristic that is additionally fraught by affective polarization. Hence, the door rule is low in divisiveness, the eyes rule and party rule are high in divisiveness, and the party rule adds an association with partisan hostility.

Hence, the experiment has a 3 (relative pay: disadvantaged, advantaged, equal) x 3 (payment rule: door, eyes, party) design. Table 1 summarizes the hypotheses and experimental predictions. First, based on a large literature (e.g., Dawes et al. 2007; Fehr & Schmidt, 1999; Shaw & Olson 2012; Starmans, Sheskin, and Bloom, 2017), the equity hypothesis predicts that participants will judge unequal pay for the same work as less fair than equal pay, especially when they are disadvantaged by the rule (holding constant the amount of pay; disadvantaged and equal pay are both a 30 cent bonus). Second, the self-interest hypothesis is also drawn from previous work (DeScioli et al. 2014; Fehr and Schmidt 1999; Sweeney 1990; Sweeney et al. 1990), and self-interest predicts that participants will judge a payment rule to be less fair when they are disadvantaged compared to advantaged by it.

Table 1
Experimental Predictions for Fairness Judgments

Hypothesis	Predictions
Equity	Disadvantaged < equal
	Advantaged < equal
Self-interest	Disadvantaged < advantaged
*Divisiveness	Eyes rule < door rule
	Party rule < door rule
*Affective polarization	Party rule < eyes rule

Note. Summary of experimental predictions, where < means "less fair than." The * indicates the distinctive hypotheses about payment procedures that are the main focus of the experiment. The predictions imply all else equal (e.g., "disadvantaged < equal" means within each payment rule). Note that these hypotheses are not mutually exclusive.

The main focus of this experiment are the next hypotheses in Table 1 about discriminatory payment rules. The divisiveness hypothesis predicts that participants will judge

4

the eyes rule and party rule to be less fair than the door rule. This is because the choice of door is transient, whereas eye color and partisanship are stable characteristics so that rules based on them consistently deprive the same individuals, making these rules divisive and hence more intensely opposed (Grafen 1987). This prediction applies especially to the disadvantaged participants, and to a lesser extent the advantaged (given the expected mitigating effect of self-interest previously discussed).

Last, the affective polarization hypothesis predicts that the party rule will be judged as less fair than eyes rule. The party rule is not only divisive but is also associated with a prominent and increasingly hostile divide between Democrats and Republicans. Aware of these social tensions, participants will judge procedures to be more unfair when they further provoke ongoing conflicts. Note that this polarization hypothesis does not require that someone is an active or strong partisan themselves, but only that they are aware of partisan conflict in society (just as someone does not need to be elderly or disabled to be offended by discrimination against the elderly or disabled).

Methods

We recruited 541 participants (45% female; age: $M = 32.5$, $SD = 10.7$) to complete a short study (~10 min) using the online labor market MTurk (Berinsky, Margolis, and Sances 2014). We applied Mturk filters requiring that participants were located in the United States and had a 98% approval rate for previous work. Participants received a small payment (50 cents) and could earn additional money from the task which would be paid to them as a bonus. We report all measures, manipulations, and exclusions in the experiment. The sample size of approximately 60 per condition was determined before analysis to provide sufficient power to detect medium effect sizes in the key comparisons. Data collection was not continued after the analysis began.

Participants first answered demographic items, reported their political party, selected their eye color, and chose between a yellow door and green door (see full materials in the Appendix). For political party, participants answered a forced-choice question about whether they were a strong, moderate, or weak Democrat or Republican; if none applied, participants were prompted to choose "the closest option," since many independents lean more toward one party (Layman & Carsey, 2002). For eye color, participants made a forced-choice between blue/green and brown/other.

Next, participants transcribed two paragraphs for an additional payment. They read that the exact nature of the payment would be shown after they finished the work. Then participants were shown the payment rule used to determine their earnings for transcription and the amount they made. Participants next answered the question "In your opinion, how fair was the Payment Rule used for the transcription task?" by rating fairness on a 7-point scale (coded as *very unfair* = -3, *neither fair nor unfair* = 0, and *very fair* = +3). Participants were debriefed and later paid their Mturk bonuses.

In a 3 (relative pay: advantaged, disadvantaged, and equal) x 3 (payment rule: door, eyes, or political parties) between-subject design, we manipulated participants' pay relative to other workers and the payment rules that determined earnings. For the political party rule, participants read:

> "*Participants who chose Democrat [Republican] receive 30 cents for completing the transcription task. Participants who chose Republican [Democrat] receive 60 cents for completing the transcription task.*"

The survey software (Qualtrics) automatically completed the text with the participant's reported political party and according to whether they were randomly assigned to the advantaged or disadvantaged condition. The door rule and eyes rule were similar except it said their pay was determined by their choice of door or eye color.

We also included an equal pay condition alongside each unequal payment rule.[1] Instead of the text above, participants read:

> "*All participants receive the same payment of 30 cents for completing the transcription task.*"

We used the same equal pay condition as separate baselines for each unequal payment rule. For equal pay, participants received 30 cents which was the same amount as the disadvantaged participants. We set the equal pay to be the same as disadvantaged pay so that any difference in fairness judgments can be attributed to the payment rule rather than the absolute amount of earnings.

Results

Figure 1 and Table 2 show the results. We first examine the predictions from previous theories about equity and self-interest. We then report the main analysis testing predictions from the divisiveness and affective polarization hypotheses.

Equity and self-interest

We test how relative pay affects fairness judgments within each payment rule. For the door rule, participants judged that payments were less fair when they were disadvantaged compared to equal pay, $t(119) = 5.31$, $p < .001$, $d = 0.96$. This shows a concern for equity particularly because the disadvantaged payment was the same as for those who receive equal pay (30 cents). In contrast, advantaged participants' judgments did not differ from equal pay, $t(118) = 1.26$, $p = .10$, $d = 0.23$, showing little concern for equity. Consistent with self-interest, participants who were advantaged rated the same rule as more fair than those who were disadvantaged by it, $t(119) = 3.75$, $p < .001$, $d = 0.68$.

For the eyes rule, participants judged the payments as less fair when disadvantaged compared to equal pay, $t(121) = 11.10$, $p < .001$, $d = 2.00$, and when advantaged compared to equal pay, $t(117) = 7.55$, $p < .001$, $d = 1.39$, which shows concern for equity. And again showing self-interest, participants judged the payment rule to be more fair when they were advantaged then disadvantaged by it, $t(116) = 2.45$, $p = .01$, $d = 0.45$.

For the political party rule, participants again showed concern for equity: They judged the rule as less fair when disadvantaged compared to equal pay, $t(118) = 13.34$, $p < .001$, $d = 2.43$, and when advantaged compared to equal pay, $t(118) = 5.07$, $p < .001$, $d = 0.93$. And evincing

---

[1] We repeated the equal pay condition three times, even though the stimuli were identical, to provide independent baselines for each of the three unequal payment rules (based on door, eyes, or party). Although somewhat inefficient, this also provided three replications of the baseline case for comparison to the unequal conditions.

self-interest, participants judged the rule as more fair when advantaged than disadvantaged, $t(117) = 5.98$, $p < .001$, $d = 1.10$.
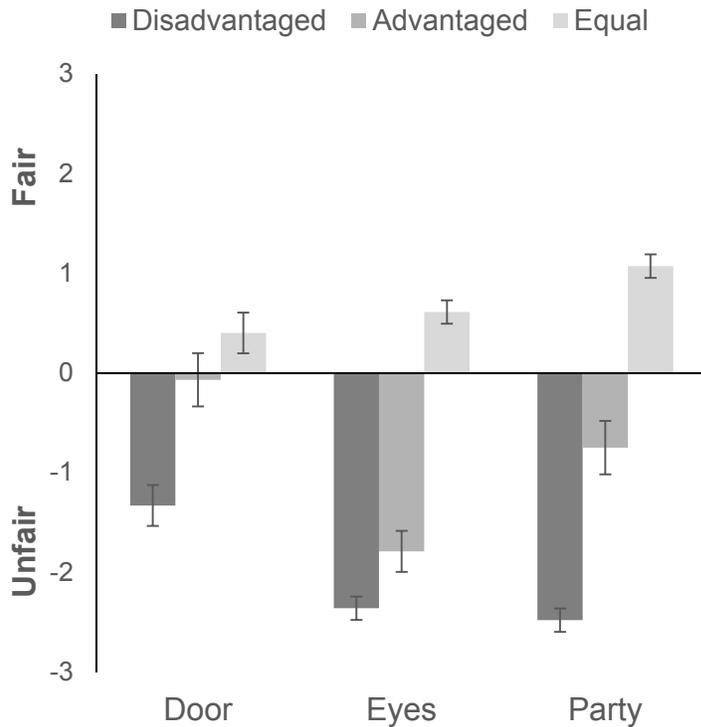

Figure 1. Mean (*SE*) unfairness ratings by payment rule (door, eyes, political party) and relative pay (disadvantaged, advantaged, equal). Unfairness was rated on a 7-point scale (-3 = *very unfair*, 0 = *neither fair nor unfair*, and 3 = *very fair*).

Divisiveness and affective polarization

To test the divisiveness and affective polarization hypotheses, the main analysis compares the different payment rules, especially for disadvantaged participants. When disadvantaged, participants judged the eyes rule as less fair than the door rule, $t(120) = 4.38$, $p < .001$, $d = 0.79$, and they judged the party rule as less fair than the door rule, $t(119) = 4.87$, $p < .001$, $d = 0.89$. These results support the divisiveness hypothesis. Participants' judgments of the eyes rule and party rule did not differ, $t(119) = 0.74$, $p = .23$, $d = 0.13$, contradicting the affective polarization hypothesis.

When advantaged, participants judged the eyes rule as less fair than the door rule, $t(115) = 5.06$, $p < .001$, $d = 0.94$, and they judged the political party rule as less fair than the door rule, $t(117) = 1.79$, $p = .04$, $d = 0.33$.[2] Unexpectedly, however, advantaged participants judged the party rule as *more* fair than the eyes rule, $t(114) = 3.08$, $p = .001$, $d = 0.57$, which is the opposite direction of our initial prediction based on affective polarization. Apparently, participants thought it was more acceptable to earn more money than political rivals, compared to those with different eye color. Although we hadn't considered this possibility, it does have an intuitive (post

---

[2] In further post hoc analyses, we found that this difference (advantaged: party vs. door) did not reach significance with a nonparametric test or when adding covariates for sex and age (see Appendix).

hoc) interpretation, which is that partisan antipathy reduces someone's sympathy for a partisan rival who is deprived of equal pay.

Table 1
Fairness Ratings for Payment Rule

	n	*M*	*(SD)*
Door			
Disadvantaged	61	-1.32	(1.60)
Advantaged	60	-0.07	(2.07)
Equal	60	0.40	(1.97)
Eyes			
Disadvantaged	61	-2.36	(0.91)
Advantaged	57	-1.79	(1.56)
Equal	62	0.61	(1.88)
Party			
Disadvantaged	60	-2.48	(0.91)
Advantaged	59	-0.75	(2.06)
Equal	61	1.07	(1.85)

Note. Fairness scale from *very unfair* = -3 to *very fair* = +3.

How does partisan strength relate to judgments of fairness for partisan discrimination?

We also conducted a post hoc analysis of partisan strength. Although we did not have initial predictions about partisan strength, it seems likely that someone's loyalty to their party would affect how offended they are by partisan discrimination. To take an initial look, we divided participants who said they were weak partisans from those who were moderate or strong partisans. (Note that weak partisans would also include independents who chose the closest party in the forced-choice partisanship question.) We repeat the key comparisons between payment rules, separately for weak and moderate/strong partisans.

Table 3 shows the results (see Appendix for descriptive statistics). For disadvantaged participants, both groups of partisans judged the party rule as less fair than the door rule, as before. However, the moderate/strong partisans judged that the party rule was less fair than the eyes rule, differing from before and this time consistent with the affective polarization hypothesis; in contrast, weak partisans continued to show no difference as in the initial analysis. For advantaged participants, both groups of partisans showed no significant difference between the party rule and door rule (although marginal for weak partisans), which is consistent with the tenuous effect found in the main analysis. And both groups judged that the party rule is *more* fair than the door rule, as in the main analysis. In sum, moderate/strong partisans, when disadvantaged, judged that the party rule is less fair than the eyes rule, which lends some support to the affective polarization hypothesis, although cautious support given that this analysis is post hoc.

Table 3

Comparisons with Party Rule by Partisan Strength

Comparison	Moderate/Strong Partisan			Weak Partisan		
	t	*p*	*d*	*t*	*p*	*d*
Disadvantaged						
Door vs. Party	4.40	0.00	1.01	2.21	0.02	0.65
Eyes vs. Party	1.95	0.03	0.49	0.83	0.21	0.23
Advantaged						
Door vs. Party	0.75	0.23	0.21	1.55	0.06	0.38
Eyes vs. Party	2.18	0.02	0.56	2.13	0.02	0.57

Discussion

The main focus of this experiment was to test whether disadvantaged participants' judgments of fairness differ depending on the basis of pay discrimination. We found that disadvantaged participants judged that lower pay was less fair when it was based on their political party or eye color, compared to their choice between doors. This finding shows that participants judged some forms of wage discrimination to be more unfair than others. This is consistent with the procedural justice literature which finds that people are concerned not only with outcomes but also the particular procedures that determine the outcomes (Tyler and Blader 2003). More specifically, this result supports the divisiveness hypothesis which holds that people judge rules and procedures to be less fair when they consistently deprive the same individuals.

We found mixed evidence for the affective polarization hypothesis. Participants judged partisan discrimination as highly unfair but not as significantly more unfair than discrimination based on eye color, which is not explicitly associated with polarized factions in society. However, we found in a post hoc analysis that disadvantaged participants who were moderate or strong partisans did judge the party rule as less fair than the eyes rule. But we note that this experiment was not designed to examine partisan strength, although this preliminary evidence suggests that it could be a promising direction for future research on partisan discrimination.

We note that another possible interpretation of the eyes rule is that some participants saw it as a correlate of race. We acknowledge this possibility which can be addressed in future research by examining discrimination based on other traits such as handedness. Even so, we point out that eye color is only roughly linked to race in the United States (where the study was conducted) since brown eyes are widespread across races in the United States. Moreover, even if this occurred, we intended the eyes rule to be *less* explicitly associated with polarized factions than political partisanship. For this purpose, we think it is informative to compare the eyes rule to the party rule, even while we agree that future research should continue to examine additional stable characteristics.

For the advantaged participants, we observed a similar pattern of differences across payment rules, with one exception: Participants thought it was more fair when they were advantaged over their political rivals, compared to people with different color eyes, contrary to our initial expectation. This observation suggests that tense polarization can reduce an advantaged group's sympathy for the disadvantaged group's unfair treatment.

Finally, the effects of relative pay were consistent with previous research on equity and self-interest (DeScioli et al. 2014; Fehr and Schmidt 1999; Sweeney 1990; Sweeney et al. 1990). Participants showed a concern for equity by judging unequal pay to be less fair than equal pay (while the amount of pay was held constant). And, they showed a self-interested bias by judging the same payment rule to be more fair when they were advantaged rather than disadvantaged by it.

Importantly, these patterns of results are necessarily based on a small number of cases looking at three particular rules. We think the current experiment provides revealing tests of the current theories, but we do not recommend drawing firm conclusions about them, which would ultimately depend on a large number of tests that examine a variety of rules in future research.

The divisiveness hypothesis could have important policy implications. For example, it may help explain why protected classes in anti-discrimination laws tend to be defined by stable characteristics such as race, sex, age, religion, and national origin. Discrimination based on these characteristics is divisive because it consistently disadvantages the same people. Divisiveness might also affect other political domains such as conflicts over economic redistribution. For instance, policies that increase taxes on the wealthy are more divisive when the wealthy are consistently the same people, while economic mobility is expected to reduce divisiveness and hence conflict surrounding progressive taxes (Moene and Wallerstein, 2001; Rehm et al. 2012).

In sum, we found in the case of wage discrimination that even among illegitimate rules—all three payment rules in this experiment—participants judged some rules to be a more serious breach of fairness than others. We suggest that researchers can leverage this variation in unfairness to test alternative theories about procedural justice. Moreover, these theories could have important implications for which forms of inequality are more likely to provoke public opposition, political mobilization, and legislative action, and which forms might escape notice even if they have damaging effects. The present results offer some initial evidence that one key factor may be a rule's divisiveness based on its connection to stable characteristics rather than transient choices. Finally, these findings about fairness and wage discrimination may hint at deeper connections between coordination games, moral judgment, and procedural justice.

References

Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention to Self-administered Surveys." *American Journal of Political Science*, *58*, 739–753.

Blau, F. D., & Kahn, L. M. (1992). "Race and Gender Pay Differentials." *National Bureau of Economic Research*.

Blau, F. D., & Kahn, L. M. (2000). "Gender Differences in Pay." *National Bureau of Economic Research*.

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*. New York, NY: Russell Sage Foundation.

Chwe, M. S.-Y. (2001). *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton, NJ: Princeton University Press.

Cronk, L., & Leech, B. L. (2012). *Meeting at Grand Central: Understanding the Social and Evolutionary Roots of Cooperation*. Princeton, NJ: Princeton University Press.

Dawes, C.T. Fowler, J.H., Johnson, J., McElreath, R. & Smirnov, O. (2007). "Egalitarian Motives in Humans." *Nature*, *446*, 794-796.

DeScioli, P., & Karpoff, R. (2015). "People's Judgments about Classic Property Law Cases." *Human Nature*, *26*, 184–209.

DeScioli, P., & Kurzban, R. (2011). "A Solution to the Mysteries of Morality." *Psychological Bulletin*, *139*, 477–496.

DeScioli, P., Massenkoff, M., Shaw, A., Petersen, M. B., & Kurzban, R. (2014). "Equity or Equality? Moral Judgments Follow the Money." *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20142122.

DeScioli, P., & Wilson, B. J. (2013). "The Territorial Foundations of Human Property". *Evolution and Human Behavior*, *32*, 297–304.

Dickson, E. S., Gordon, S. C., & Huber, G. A. (2009). "Enforcement and Compliance in an Uncertain World: An Experimental Investigation". *The Journal of Politics*, *71*, 1357–1378.

Dickson, E. S., Gordon, S. C., & Huber, G. A. (2015). "Institutional Sources of Legitimate Authority: An Experimental Investigation." *American Journal of Political Science*, *59*, 109–127.

Fehr, E., & Schmidt, K. M. (1999). "A Theory of Fairness, Competition, and Cooperation." *The Quarterly Journal of Economics*, *114*, 817–868.

Gerber, A. S., Huber, G. A., Doherty, D., & Dowling, C. M. (2012). "Is There a Secret Ballot? Ballot Secrecy Perceptions and Their Implications for Voting Behaviour." *British Journal of Political Science*, *43*, 77-102.

Gibson, J. L. (1989). "Understandings of Justice: Institutional Legitimacy, Procedural Justice, and Political Tolerance." *Law and Society Review*, *23*, 469–496.

Grafen, A. (1987). "The Logic of Divisively Asymmetric Contests: Respect for Ownership and the Desperado Effect." *Animal Behaviour*, *35*, 462–467.

Green, D. P., Palmquist, P., & Schickler, E. (2004). *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters.* New Haven, CT: Yale University Press.

Hibbing, J. R., & Alford, J. R. (2004). "Accepting Authoritative Decisions: Humans as Wary Cooperators." *American Journal of Political Science*, *48*, 62–76.

Hibbing, J. R., & Theiss-Morse, E. (2001). "Process Preferences and American Politics: What People Want Government To Be." *American Political Science Review, 95*, 145–154.

Iyengar, S. Sood, G., & Lelkes, Y. (2012). "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly, 76*, 405-431.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). "Can Race Be Erased? Coalitional Computation and Social Categorization." *Proceedings of the National Academy of Sciences, 98,* 15387–15392.

Layman, G.C., & Carsey, T.M. (2002). "Party Polarization and 'Conflict Extension' in the American Electorate." *American Journal of Political Science, 46,* 786-802.

Leicht, K. T. (2008). "Broken Down by Race and Gender? Sociological Explanations of New Sources of Earnings Inequality." *Annual Review of Sociology*, *34*, 237–255.

Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.

Mason, L. (2015). "'I Disrespectfully Agree': The Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science*, *59*, 128-145.

Maynard Smith, John. (1982). *Evolution and the Theory of Games*. Cambridge, England: Cambridge University Press.

Maynard Smith, John, & G.A.Parker. (1976). "The Logic of Asymmetric Contests." *Animal Behaviour*, *24*, 159–175.

Moene, K. O., & Wallerstein, M. (2001). "Inequality, Social Insurance, and Redistribution." *American Political Science Review, 95*, 859–874.

Pinker, S. (1997). *How the Mind Works*. New York, NY: W.W. Norton & Company.

Pinker, S. (2007). *The Stuff of Thought: Language as a Window into Human Nature*. New York, NY: Viking.

Reed, L. I., DeScioli, P., & Pinker, S. A. (2014). "The Commitment Function of Angry Facial Expressions." *Psychological Science*, *25*, 1511–1517.

Rehm, P., Hacker, J. S., & Schlesinger, M. (2012). "Insecure Alliances: Risk, Inequality, and Support for the Welfare State." *American Political Science Review*, *106*, 386–406.

Schelling, T. C. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.

Shaw, A. (2013). "Beyond 'to Share or Not to Share' The Impartiality Account of Fairness." *Current Directions in Psychological Science*, *22*, 413–17.

Shaw, A., DeScioli, P., & Olson, K. R. (2012). "Fairness versus Favoritism in Children." *Evolution and Human Behavior*, *33*, 736–45.

Shaw, A. & Olson, K.R. (2012). "Children Discard a Resource to Avoid Inequity." *Journal of Experimental Psychology*: *General*, *141*, 382-395.

Starmans, C., Sheskin, M., & Bloom, P. (2017). "Why People Prefer Unequal Societies." *Nature Human Behaviour*, *1*, 0082.

Sweeney, P. D. (1990). "Distributive Justice and Pay Satisfaction: A Field Test of an Equity Theory Prediction." *Journal of Business and Psychology*, *4*, 329–41.

Sweeney, P. D., McFarlin, D. B., & Inderrieden, E. J. (1990). "Using Relative Deprivation Theory to Explain Satisfaction With Income and Pay Level: A Multistudy Examination." *Academy of Management Journal*, *33*, 423–36.

Thomas, K. A., DeScioli, P., Haque, O. S., & Pinker, S. (2014). "The Psychology of Coordination and Common Knowledge." *Journal of Personality and Social Psychology*, *107*, 657–676.

Tybur, J. M., Lieberman, D., Kurzban, R., & DeScioli, P. (2013). "Disgust: Evolved Function and Structure." *Psychological Review*, *120*, 65–84.

Tyler, T. R., & Blader, S. L. (2003). "Procedural Justice, Social Identity, and Cooperative Behavior." *Personality and Social Psychology Review*, *7*, 349–61.

Tyler, T. R., Casper, J. D., & Fisher, B. "Maintaining Allegiance Toward Political Authorities: The Role of Prior Attitudes and the Use of Fair Procedures." *American Journal of Political Science*, *33*, 629-652.

Van Vugt, M. (2006). "Evolutionary Origins of Leadership and Followership." *Personality and Social Psychology Review*, *10*, 354-371.

Supplementary Results

Comparing payment rules with nonparametric tests

     We compared the payment rules with post hoc nonparametric tests (Wilcoxon rank-sum) for disadvantaged and advantaged participants (holding constant the distribution of payoffs) (Table A1). Most important, the results are the same as the t-tests for the disadvantaged participants, who are the main focus of the experiment. Among the advantaged participants, there is one exception: the difference between door rule and party rule is no longer significant (reduced to marginal significance) with the non-parametric test. Importantly, the significance of this case either way does not substantively alter our interpretation in the main text, because it was already clear that advantaged partisans were more accepting of unfairness toward their partisan opponents than we initially expected, based on the comparison to those advantaged by the eyes rule. This is now further reinforced by the fact that even the difference between the party rule and door rule is small and unstable for a nonparametric test.

Table A1

Wilcoxon Rank-Sum Tests for Payment
Rules by Relative Pay

Comparison	z	p
Disadvantaged		
Door vs. Eyes	3.83	0.000
Door vs. Party	4.47	0.000
Eyes vs. Party	1.01	0.311
Advantaged		
Door vs. Eyes	4.64	0.000
Door vs. Party	1.85	0.065
Eyes vs. Party	2.95	0.003

Regression analysis of fairness judgments with covariates for age and sex.

     We conducted post hoc regression analyses for how the payment rules affected fairness with added covariates for age and sex, separately for disadvantaged and advantaged participants (Table A2). Age and sex were not significant in either regression. Among disadvantaged participants, the differences are the same as for the t-tests in the main text: the party rule and eyes rule are less fair than the door rule, while the party rule and eyes rule did not differ from each other (planned comparison of the coefficients for the party rule and the eye rule, $p = 0.66$). Among advantaged participants, we see the same exception as for the non-parametric tests above: the party rule did not differ from the door rule. This further underscores that this

difference was tenuous, which more generally fits with the unexpected finding that partisans were relatively accepting of low pay for their partisan rivals.

Table A2

Fairness Ratings Predicted by Payment Rule

	Disadvantaged	Advantaged
Party Condition	-1.15 (0.22)***	-0.61 (0.35)
Eyes Condition	-1.06 (0.22)***	-1.69 (0.36)***
Female	-0.34 (0.18)	-0.35 (0.29)
Age	-0.00 (0.01)	-0.01 (0.01)
Constant	-1.10 (0.31)***	0.46 (0.51)
Observations	181	176
R-squared	0.18	0.13

Note: OLS estimates with standard errors in parentheses.
The door condition is the omitted category.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Descriptive Statistics by Partisan Strength

　　　　Table A3 reports descriptive statistics for fairness judgments by strength of partisanship.

Table A3

Fairness Ratings for Payment Rule by Strength of Partisanship

	Moderate/Strong		Weak	
	n	*M* (*SD*)	*n*	*M* (*SD*)
Door				
Disadvantaged	37	-1.22 (1.69)	24	-2.31 (1.47)
Advantaged	22	-0.36 (2.11)	38	0.11 (2.06)
Equal	31	0.16 (1.81)	29	0.66 (2.13)
Eyes				
Disadvantaged	27	-2.15 (1.79)	34	-2.52 (1.95)
Advantaged	30	-1.80 (1.45)	27	-1.77 (1.70)
Equal	20	0.65 (1.79)	42	0.60 (1.95)
Party				
Disadvantaged	38	-2.58 (0.89)	22	-2.31 (0.95)
Advantaged	30	-0.80 (2.06)	29	-0.69 (2.09)
Equal	29	0.86 (1.88)	32	1.25 (1.83)

Note. Fairness scale from *very unfair* = -3 to *very fair* = +3.

Experimental Protocol

      Participants were recruited from Amazon Mechanical Turk and they completed the survey in Qualtrics. Participants were assigned to a payment rule (door, eyes, or party) and a level of relative pay (disadvantaged, advantaged, or equal).

Survey Materials

<div align="center">Page 1</div>

In this HIT, you will complete a task and receive a bonus payment for the task based on your effort.

You will earn 50 cents for completing this HIT, and you can make additional money from the tasks you complete. Any additional money that you earn will be paid to you as an MTurk bonus. The exact amount that you earn depends on your choices for the tasks. <u>All of the instructions that you read beyond this point refer to money that you can earn above and beyond the 50 cents you will automatically earn from completing the HIT.</u>

<div align="center"><u>Page 2</u></div>

1. What is your sex?
   - Male
   - Female
2. What is your age? (open response)
3. What is your race? (open response)
4. What is your eye color?
   - Blue/Green
   - Brown/Other
5. How important is this eye color to you?
   - Not Important
   - Somewhat Important
   - Very Important
6. Generally speaking, which of the following do you consider yourself to be? (If none of these, then please choose the closest option.)
   - Very Liberal
   - Moderately Liberal
   - Slightly Liberal
   - Slightly Conservative
   - Moderately Conservative
   - Very Conservative
7. How important is this ideological preference to you?
   - Not Important
   - Somewhat Important
   - Very Important

8. Generally speaking, which of the following do you consider yourself to be? (If none of these, then please choose the closest option.)
   - Strong Democrat
   - Moderate Democrat
   - Weak Democrat
   - Weak Republican
   - Moderate Republican
   - Strong Republican
9. Which of these four types of music do you most prefer?
   - Rock
   - Classical/Jazz
   - Country
   - Hip Hop/Rap
10. How important is this music preference to you?
    - Not Important
    - Somewhat Important
    - Important
11. You will now pick between two doors. This choice might affect what happens later in the study.

Author's Note: Items 5 and 7 were not included in the party condition.

Page 3

Effort: Transcription Task
You will earn additional money by completing the following transcription task. You might earn more, less, or the same amount as other participants for completing the task. The exact amount and the Payment Rule used to determine your pay will be explained after you complete the task.

Page 4

Transcription Task: Below, you will find lines of scanned text in image form and your goal is to accurately type the sentences into the textbox underneath the image. If you transcribe the paragraphs accurately, then you will receive payment for the task (determined by the Payment Rule)

Eggs are a very popular breakfast dish. They're most often scrambled, best served hot, and often accompanied by sausage, bacon, potatoes, or any number of familiar breakfast items. Some people like to eat them with ketchup, Tabasco sauce, or just plain old salt and pepper. Another popular way to make eggs (especially if you need a snack for the road) is to boil them.

(space to transcribe paragraph)

Eggs contain a lot of cholesterol so it's best to limit yourself to one or two a day, depending on your body size and dietary needs. It's for this reason that some people restrict themselves to egg whites, which are made by separating the yolk from the rest of the contents of the egg. However, yolks also contain all of the vitamin A, D, and E in the egg, so this leaves out some key nutrients.

(space to transcribe paragraph)

(Page 5)

Payment

We will use the following Payment Rule to determine your pay from the transcription task.

(Eye Color/Disadvantaged Condition)

Payment Rule: Participants who stated they have (Blue/Green colored eyes OR Brown/Other Colored Eyes) receive 30 cents from completing the transcription task. Participants who stated they have (Brown/Other colored eyes OR Blue/Green colored eyes) receive 60 cents for completing the transcription task.

You previously chose: (Blue/Green OR Brown/Other). Your payment is: 30 cents.

(Eye Color/Advantaged Condition)

Payment Rule: Participants who stated they have (Blue/Green colored eyes OR Brown/Other Colored Eyes) receive 60 cents from completing the transcription task. Participants who stated they have (Brown/Other colored eyes OR Blue/Green colored eyes) receive 30 cents for completing the transcription task.

You previously chose: (Blue/Green OR Brown/Other). Your payment is: 60 cents.

(Fair)

Payment Rule: All participants receive the same payment of 30 cents for completing the transcription task.

(Door/Disadvantaged Condition)

Payment Rule: Participants who stated they have (Door 1 OR Door 2)) receive 30 cents from completing the transcription task. Participants who stated they have (Door 2 OR Door 1) receive 60 cents for completing the transcription task.

You previously chose: (Door 1 OR Door 2). Your payment is: 30 cents.

(Door/Advantaged Condition)

Payment Rule: Participants who stated they have (Door 1 OR Door 2) receive 60 cents from completing the transcription task. Participants who stated they have (Door 2 or Door 1)) receive 30 cents for completing the transcription task.

You previously chose: (Door 1 OR Door 2). Your payment is: 60 cents.

(Party/Disadvantaged Condition**)**

Payment Rule: Participants who stated they have (Democrat OR Republican)) receive 30 cents from completing the transcription task. Participants who stated they have (Republican OR Democrat) receive 60 cents for completing the transcription task.

You previously chose: (Democrat OR Republican). Your payment is: 30 cents.

(Party/Advantaged Condition)

Payment Rule: Participants who stated they have (Democrat OR Republican) receive 60 cents from completing the transcription task. Participants who stated they have ((Republican OR Democrat) receive 30 cents for completing the transcription task.

You previously chose: (Democrat OR Republican). Your payment is: 60 cents.

(Page 6)

In your opinion, how fair was the Payment Rule used for the transcription task?
- Very Unfair
- Somewhat Unfair
- Slightly Unfair
- Neither Fair nor Unfair
- Slightly Fair
- Somewhat Fair
- Very Fair

(Page 7)

1. If you have any additional thoughts about the fairness of the Payment Rule used for the task, then please share those thoughts here: (open response)
2. If you have any other thoughts about the study, please share those here: (open response)

We appreciate your participation in the study. The purpose of this project is to study the effects of fair and unfair procedures. For participants who experienced unfair procedures, we want you to know that this was critical for addressing an important research question about the effects of unfairness. We appreciate your patience and understanding. Your participation is an important contribution to this area of research. Once again, we thank you for your participation in the study.

Please make sure to click on the arrow below to get to the next page and receive your Mturk validation code. Without this code, we cannot verify that you have completed the HIT and therefore will not be able to pay you.