Forthcoming, *Econometrica*, 2009.

Social Image and the 50-50 Norm:
A Theoretical and Experimental Analysis of Audience Effects*

James Andreoni, University of California, San Diego
B. Douglas Bernheim, Stanford University and NBER

April 6, 2009

Abstract

A norm of 50-50 division appears to have considerable force in a wide range of economic environments, both in the real world and in the laboratory. Even in settings where one party unilaterally determines the allocation of a prize (the dictator game), many subjects voluntarily cede exactly half to another individual. The hypothesis that people care about fairness does not by itself account for key experimental patterns. We consider an alternative explanation, which adds the hypothesis that people like to be *perceived* as fair. The properties of equilibria for the resulting signaling game correspond closely to laboratory observations. The theory has additional testable implications, the validity of which we confirm through new experiments.

*We are indebted to the following people for helpful comments: Iris Bohnet, Colin Camerer, Navin Kartik, Antonio Rangel, three anonymous referees, and seminar participants at the California Institute of Technology, NYU, and Stanford University's SITE workshop in Psychology and Economics. We acknowledge financial support from the National Science Foundation through grant numbers SES-0551296 (Andreoni), and SES-0452300 (Bernheim).

1 Introduction

Equal division of monetary rewards and/or costs is a widely observed behavioral norm. Fifty-fifty sharing is common in the context of joint ventures among corporations (e.g. Veuglers and Kesteloot [1996], Dasgupta and Tao [1998], and Hauswald and Hege [2003]),[1] share tenancy in agriculture (e.g. De Weaver and Roumasset [2002], Agrawal [2002]), and bequests to children (e.g. Wilhelm [1996], Menchik [1980, 1988]). "Splitting the difference" is a frequent outcome of negotiation and conventional arbitration (Bloom [1986]). Business partners often divide the earnings from joint projects equally, friends split restaurant tabs equally, and the U.S. government splits the nominal burden of the payroll tax equally between employers and employees. Compliance with a 50-50 norm has also been duplicated in the laboratory. Even when one party has all the bargaining power (the dictator game), typically 20 to 30 percent of subjects voluntarily cede half of a fixed payoff to another individual (Camerer [1997]).[2]

Our object is to develop a theory that accounts for the 50-50 norm in the dictator game, one we hope will prove applicable more generally.[3] Experimental evidence shows that a significant fraction of the population elects precisely 50-50 division even when it is possible to give slightly less or slightly more,[4] that subjects rarely cede more than 50% of the aggregate payoff, and that there is frequently a trough in the distribution of fractions ceded just below 50% (see, e.g., Forsythe et al. [1994]). In addition, choices depend on observability: greater anonymity for the dictator leads him to behave more selfishly and weakens the norm,[5] as do treatments that obscure the dictator's role in determining the outcome, or that enable him to obscure that role.[6] A good theory of behavior in the dictator game must account for all

[1] Where issues of control are critical, one also commonly sees a norm of 50-plus-one-share.

[2] The frequency of equal division is considerably higher in ultimatum games; see Camerer [2003].

[3] Our theory is not necessarily a good explanation for all 50-50 norms. For example, Bernheim and Severinov [2003] propose an explanation for equal division of bequests that involves a different mechanism.

[4] For example, according to Andreoni and Miller [2002], a significant fraction of subjects (15 to 30%) adhered to equal division regardless of the sacrifice to themselves.

[5] In double-blind trials, subjects cede smaller amounts, and significantly fewer adhere to the 50-50 norm (e.g. Hoffman et al. [1996]). However, when dictators and recipients face each other, adherence to the norm is far more common (Bohnet and Frey [1999]). Andreoni and Petrie [2004] and Rege and Telle [2004] also find a greater tendency to equalize payoffs when there is an audience. More generally, studies of field data confirm that an audience increases charitable giving (Soetevent [2005]). Indeed, charities can influence contributions by adjusting the coarseness of the information provided to the audience (Harbaugh [1998]).

[6] See Dana, Cain, and Dawes [2007], Dana, Weber, and Kuang [2007], and Broberg, Ellingsen, Johannesson

these robust patterns.

The leading theories of behavior in the dictator game invoke altruism or concerns for fairness (e.g., Fehr and Schmidt [1999], Bolton and Ockefels [2000]). One can reconcile those hypotheses with the observed distribution of choices, but only by making awkward assumptions – for example, that the utility function is fortuitously kinked, that the underlying distribution of preferences contains gaps and atoms, or that dictators are boundedly rational. Indeed, with a differentiable utility function, the fairness hypothesis cannot explain why *anyone* would choose equal division (see Section 2, below). Moreover, neither altruism nor a preference for fairness explains why observability, and hence audiences, play such an important role in determining the norm's strength.

This paper explores the implications of supplementing the fairness hypothesis with an additional plausible assumption: people like to be perceived as fair. We incorporate that desire directly into the utility function; alternatively, one could depict the dictator's preference as arising from concerns about subsequent interactions.[7] Our model gives rise to a signaling game wherein the dictator's choice affects others' inferences about his taste for fairness. Due to an intrinsic failure of the single-crossing property, the equilibrium distribution of transfers replicates the choice patterns listed above: there is a pool at precisely equal division, and no one gives either more or slightly less than half of the prize. In addition, consistent with experimental findings, the size of the equal division pool depends on the observability of the dictator's choice. Thus, while our theory does leave some experimental results unexplained (see, e.g., Oberholzer-Gee and Eichenberger [2008], or our discussion of Cherry et al. [2002] in Section 2), it nevertheless has considerable explanatory power.

We also examine an extended version of the dictator game in which (a) nature sometimes

[2007]. Various papers have made a similar point in the context of the ultimatum game (Kagel, Kim, and Moser [1996], Güth, Huck, and Ockenfels [1996], and Mitzkewitz and Nagel [1993]) and the hold-up problem (Ellingsen and Johannesson [2005]). However, when the recipient is sufficiently removed from the dictator, the recipient's potential inferences about the dictator's motives have a small effect on choices (Koch and Normann [2008]).

[7] For example, experimental evidence reveals that the typical person treats others better when he believes they have good intentions; see Blount [1995], Andreoni, Brown and Vesterlund [2002], or Falk et al. [2008].

intervenes, choosing an unfavorable outcome for the recipient, and (b) the recipient cannot observe whether nature intervened. We demonstrate that the equilibrium distribution of voluntary choices includes two pools, one at equal division and one at the transfer that nature sometimes imposes. An analysis of comparative statics identifies testable implications concerning the effects of two parameters. First, a change in the transfer that nature sometimes imposes changes the location of the lower pool. Second, an increase in the probability that nature intervenes reduces the size of the equal division pool and increases the size of the lower pool. We conduct new experiments designed to test those implications. Subjects exhibit the predicted behavior to a striking degree.

The most closely related paper in the existing theoretical literature is Levine [1998]. In Levine's model, the typical individual acts generously to signal his altruism so that others will act more altruistically toward him. Though Levine's analysis of the ultimatum game involves some obvious parallels with our work, he focuses on a different behavioral puzzle.[8] Most importantly, his analysis does not account for the 50-50 norm.[9] He explicitly addresses only one feature of the behavioral patterns discussed above – the absence of transfers exceeding 50 percent of the prize – and his explanation depends on restrictive assumptions.[10] As a general matter, a desire to signal altruism (rather than fairness) accords no special status to equal division, and those who care a great deal about others' inferences will potentially make even larger transfers.

One can view this paper as providing possible microfoundations for theories of warm-glow giving (Andreoni [1989, 1990]). It also contributes to the literature that explores the

[8] With respect to the ultimatum game, Levine's main point is that, with altruism alone, it is impossible to reconcile the relatively low frequency of selfish offers with the relatively high frequency of rejections.

[9] None of the equilibria Levine describes involves pooling at equal division. He exhibits a separating equilibrium in which only a single type divides the prize equally, as well as pooling equilibria in which no type chooses equal division. He also explicitly rules out the existence of a pure pooling equilibrium in which all types choose equal division.

[10] In Levine's model, the respondent's inferences matter to the proposer only because they affect the probability of acceptance. Given his parametric assumptions, an offer of 50 percent is accepted irrespective of inferences, so there is no benefit to a higher offer. If one assumes instead that a more favorable social image always has positive incremental value, then those who are sufficiently concerned with signaling altruism will end up transferring more than 50 percent. Rotemberg [2007] extends Levine's analysis and applies it to the dictator game, but imposes a maximum transfer of 50 percent by assumption.

behavioral implications of concerns for social image (e.g., Bernheim [1994], Ireland [1994], Bagwell and Bernheim [1996], Glazer and Konrad [1996]). Recent contributions in that general area include Ellingsen and Johannesson [2008], Tadelis [2007], and Manning [2007]. Our study is also related to the theoretical literature on *psychological games*, in which players have preferences over the beliefs of others (as in Geanakoplos, Pearce, and Stacchetti [1989]).

With respect to the experimental literature, our work is most closely related to a small collection of papers (cited in footnote 6) that study the effects of obscuring either a subject's role in dividing a prize or his intended division. By comparing obscured and transparent treatments, those experiments have established that subjects act more selfishly when the outcomes that follow from selfish choices have alternative explanations. We build on that literature by focusing on a class of games for which it is possible to derive robust comparative static implications from an explicit theory of audience effects; moreover, instead of studying one obscured treatment, we test the specific implications of our theory by varying two key parameters across a collection of obscured treatments.

More broadly, the experimental literature has tended to treat audience effects as unfortunate confounds that obscure "real" motives. Yet casual observation and honest introspection strongly suggest that people care deeply about how others perceive them, and that those concerns influence a wide range of decisions. Our analysis underscores both the importance and feasibility of studying audience effects with theoretical and empirical precision.

The paper proceeds as follows: Section 2 describes the model, Sections 3 and 4 provide theoretical results, Section 5 describes our experiment, and Section 6 concludes. Proofs of theorems appear in the Appendix. Other referenced appendices are available on-line.

2 The model

Two players, a dictator (D) and a receiver (R), split a prize normalized to have unit value. Let $x \in [0, 1]$ denote the transfer R receives; D consumes $c = 1 - x$. With probability $1-p$, D chooses the transfer, and with probability p, nature sets it equal to some fixed value,

x_0; then the game ends. The parameters p and x_0 are common knowledge, but R cannot observe whether nature intervened. For the standard dictator game, $p = 0$.

Potential dictators are differentiated by a parameter t, which indicates the importance placed on fairness; its value is D's private information. The distribution of t is atomless and has full support on the interval $[0, \bar{t}]$; H denotes the CDF.[11] We define H_s as the CDF obtained from H conditioning on $t \geq s$. D cares about his own prize, c, and his social image, m, as perceived by some audience A, which includes R (and possibly others, such as the experimenter). Preferences over c and m correspond to a utility function, $F(c, m)$, that is unbounded in both arguments, twice continuously differentiable, strictly increasing (with, for some $f > 0$, $F_1(c, m) > f$ for all $c \in [0, 1]$ and $m \in \mathbb{R}_+$), and strictly concave in c. D also cares about fairness, judged by the extent to which the outcome departs from the most fair alternative, x^F. Thus, we write D's total payoff as

$$U(x, m, t) = F(1 - x, m) + tG\left(x - x^F\right) \ .$$

We assume G is twice continuously differentiable, strictly concave, and reaches a maximum at zero. We follow Fehr and Schmidt [1999] and Bolton and Ockenfels [2000] in assuming the players see themselves as equally meritorious in the standard dictator game, so that $x^F = \frac{1}{2}$. Experiments by Cherry et al. [2002] suggest that a different standard may apply when dictators allocate earned wealth. While our theory does not explain the apparent variation in x^F across contexts, it can in principle accommodate that variation.[12]

Note that the dictator's preferences over x and m violate the single crossing property. Picture his indifference curves in the x, m-plane. As t increases, the slope of the indifference

[11] Some experiments appear to produce an atom in the choice distribution at 0, though the evidence for this pattern is mixed (see e.g., Camerer [2003]). Our model does not produce that pattern (for $p = 0$ or $x_0 > 0$) unless we assume that there is an atom in the distribution of types at $t = 0$. Because the type space is truncated below at 0, it may be reasonable to allow for that possibility. One could also generate a choice atom at zero with $p = 0$ by assuming that some individuals do not care about social image (in which case the analysis would be more similar to the case of $p > 0$ and $x_0 = 0$). In experiments, it is also possible that a choice atom at zero results from the discreteness of the choice set and/or approximate optimization.

[12] If the players are asymmetric with respect to publicly observed indicia of merit, the fairness of an outcome might depend on the extent to which it departs from some other benchmark, such as $x^F = 0.4$. Provided the players agree on x^F, similar results would follow, except that the behavioral norm would correspond to the alternate benchmark. However, if players have different views of x^F, matters are more complex.

curve through any point (x, m) declines if $x < \frac{1}{2}$, but rises if $x > \frac{1}{2}$. Intuitively, comparing any two dictators, if $x < \frac{1}{2}$ the one who is more fair-minded incurs a smaller utility penalty when increasing the transfer, because inequality falls; however, if $x > \frac{1}{2}$ that same dictator incurs a *larger* utility penalty when increasing the transfer, because inequality rises.

Social image m depends on A's perception of D's fairness. We normalize m so that, if A is certain D's type is $\widehat{t}$, then D's social image is $\widehat{t}$. We use Φ to denote the CDF representing A's beliefs about D's type, and $B(\Phi)$ to denote the associated social image.

Assumption 1: (1) B is continuous (where the set of CDFs is endowed with the weak topology). (2) $\min supp(\Phi) \leq B(\Phi) \leq \max supp(\Phi)$, with strict inequalities when the support of Φ is non-degenerate. (3) If Φ' is "higher" than Φ'' in the sense of first-order stochastic dominance, then $B(\Phi') > B(\Phi'')$.

As an example, B might calculate the mean of t given Φ. For some purposes, we impose a modest additional requirement (also satisfied by the mean):

Assumption 2: Consider the CDFs J, K, and L, such that $J(t) = \lambda K(t) + (1 - \lambda)L(t)$. If $\max supp(L) \leq B(J)$, then $B(J) \leq B(K)$, where the second inequality is strict if the first is strict or if the support of L is nondegenerate.[13]

The audience A forms an inference Φ about t after observing x. Even though D does not observe that inference directly, he knows A will judge him based on x, and therefore accounts for the effect of his decision on A's inference. Thus, the game involves signaling. We will confine attention throughout to pure strategy equilibria. A signaling equilibrium consists of a mapping Q from types (t) to transfers (x), and a mapping P from transfers (x) to inferences (Φ). We will write the image of x under P as P_x (rather than $P(x)$), and use $P_x(t)$ to denote the inferred probability that D's type is no greater than t upon observing x.

[13] It is perhaps more natural to assume that if $\max supp(L) \leq B(K)$, then $B(J) \leq B(K)$, where the second inequality is strict if the first is strict or if the support of L is nondegenerate. That alternative assumption, in combination with Assumption 1, implies Assumption 2 (see Lemma 5 in Andreoni and Bernheim [2007]).

Equilibrium transfers must be optimal given the inference mapping P (for all $t \in [0, \bar{t}]$, $Q(t)$ solves $\max_{x \in [0,1]} U(x, P_x, t)$), and inferences must be consistent with the transfer mapping Q (for all $x \in Q([0, \bar{t}])$ and $t \in [0, \bar{t}]$, $P_x(t) = \text{prob}(t' \leq t \mid Q(t') = x)$).

We will say that Q is an *equilibrium action function* is there exists P such that (Q, P) is a signaling equilibrium. Like most signaling models, ours has many equilibria, with many distinct equilibrium action functions. Our analysis will focus on equilibria for which the action function Q falls within a specific set, $\mathcal{Q}_1$ for the standard dictator game ($p = 0$) and $\mathcal{Q}_2$ for the extended dictator game ($p > 0$), both defined below. We will ultimately justify those restrictions by invoking a standard refinement for signaling games, the D1 criterion (due to Cho and Kreps [1987]), which insists that the audience attribute any action not chosen in equilibrium to the type that would choose it for the widest range of conceivable inferences.[14] Formally, let $U^*(t)$ denote the payoff to type t in a candidate equilibrium (Q, P), and for each $(x, t) \in [0, 1] \times [0, \bar{t}]$ define $m_x(t)$ as the value of m that satisfies $U(x, m, t) = U^*(t)$. Let $M_x = \arg\min_{t \in [0, \bar{t}]} m_x(t)$ if $\min_{t \in [0, \bar{t}]} m_x(t) \leq \bar{t}$, and $M_x = [0, \bar{t}]$ otherwise. The D1 criterion requires that, for all $x \in [0, 1] \backslash Q([0, \bar{t}])$, P_x places probability only on the set M_x.

If the dictator's type were observable, the model would not reproduce observed behavior: every type would choose a transfer strictly less than $\frac{1}{2}$ and there would be no gaps or atoms in the distribution of voluntary choices, apart from an atom at $x = 0$ (see Andreoni and Bernheim [2007]). Henceforth, we will use $x^*(t)$ to denote the optimal transfer for type t when type is observable (i.e., the value of x that maximizes $U(x, t, t)$).

3 Analysis of the standard dictator game

For the standard dictator game, we will focus on equilibria involving action functions belonging to a restricted set, $\mathcal{Q}_1$. To define that set, we must first describe differentiable

[14] We apply the D1 criteria once rather than iteratively. Similar results hold for other standard criteria (e.g. divinity). We acknowledge that experimental tests have called into question the general validity of equilibrium refinements for signaling games (see, e.g., Brandts and Holt [1992, 1993, 1995]). Our theory nevertheless performs well in this instance, possibly because the focalness of the 50-50 norm coordinates expectations.

action functions that achieve local separation of types. Consider a simpler game with $p = 0$ and types lying in some interval $[r, w] \subseteq [0, \bar{t}]$. In a separating equilibrium, for each type $t \in [r, w]$, t's choice, denoted $S(t)$, must be the value of x that maximizes the function $U(x, S^{-1}(x), t)$ over $x \in S([r, w])$. Assuming S is differentiable, the solution satisfies the first order condition $\frac{dU}{dx} = 0$. Substituting $x = S(t)$ into the first order condition, we obtain:

$$S'(t) = -\frac{F_2(1 - S(t), t)}{tG'\left(S(t) - \frac{1}{2}\right) - F_1(1 - S(t), t)} . \tag{1}$$

The preceding expression is a non-linear first order differential equation. We will be concerned with solutions with initial conditions of the form (r, z) (a choice z for type r) such that $z \geq x^*(r)$. For any such initial condition, (1) has a unique solution, denoted $S_{r,z}(t)$.[15] In the Appendix (Lemma 3), we prove that, for all r and z with $z \geq x^*(r)$, $S_{r,z}(t)$ is strictly increasing in t for $t \geq r$, and there exists a unique type $t^*_{r,z} > r$ (possibly exceeding $\bar{t}$) to which $S_{r,z}(t)$ assigns equal division (i.e., $S_{r,z}(t^*_{r,z}) = \frac{1}{2}$).

Now we define $\mathcal{Q}_1$. The action function Q belongs to $\mathcal{Q}_1$ if and only if it falls into one of the following three categories:

Efficient differentiable separating action function: $Q(t) = S_{0,0}(t)$ for all $t \in [0, \bar{t}]$, where $S_{0,0}(\bar{t}) \leq \frac{1}{2}$.

Central pooling action function: $Q(t) = \frac{1}{2}$ for all $t \in [0, \bar{t}]$.

Blended action functions: There is some $t_0 \in (0, \bar{t})$ with $S_{0,0}(t_0) < \frac{1}{2}$ such that, for $t \in [0, t_0]$, we have $Q(t) = S_{0,0}(t)$, and for $t \in (t_0, \bar{t}]$, we have $Q(t) = \frac{1}{2}$.

We will refer to equilibria that employ these types of action functions as, respectively, *efficient differentiable separating equilibria*, *central pooling equilibria*, and *blended equilibria*. A central pooling equilibrium requires $U(0, 0, 0) \leq U\left(\frac{1}{2}, B(H), 0\right)$, so that the lowest type weakly prefers to be in the pool rather than choose his first-best action and receive the worst

[15] If $z = x^*(r)$, then $S'(r)$ is undefined, but the uniqueness of the solution is still guaranteed; one simply works with the inverse separating function (see Proposition 5 of Mailath [1987]).

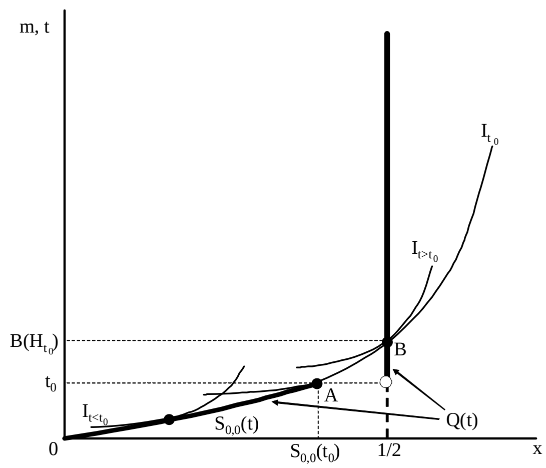

Figure 1: A Blended Equilibrium

possible inference. A blended equilibrium requires $U(S_{0,0}(t_0), t_0, t_0) = U\left(\frac{1}{2}, B\left(H_{t_0}\right), t_0\right)$, so that the highest type that separates is indifferent between separating and joining the pool.[16]

Figure 1 illustrates a blended equilibrium. Types separate up to t_0, and higher types choose equal division. An indifference curve for type t_0 (I_{t_0}) passes through both point A, the separating choice for t_0, and point B, the outcome for the pool. The indifference curve for any type $t > t_0$ through point B ($I_{t>t_0}$) is flatter than I_{t_0} to the left of B and steeper to the right. Therefore, all such types strictly prefer the pool to any point on $S_{0,0}(t)$ below t_0.

The following result establishes the existence and uniqueness of equilibria within $\mathcal{Q}_1$ and justifies our focus on that set.

Theorem 1: Assume $p = 0$ and that Assumption 1 holds. Restricting attention to $\mathcal{Q}_1$, there exists a unique equilibrium action function, Q^E. It is an efficient differentiable separating function iff $\bar{t} \leq t^*_{0,0}$.[17] Moreover, there exists an inference mapping P^E such that (Q^E, P^E) satisfies the D1 criterion, and for any other equilibrium (Q, P) satisfying

[16] Remember that H_{t_0} is defined as the CDF obtained starting from H (the population distribution) and conditioning on $t \geq t_0$. Because of t_0's indifference, there is an essentially identical equilibrium (differing from this one only on a set of measure zero) where t_0 resolves its indifference in favor of $\frac{1}{2}$ (that is, it joins the pool).

[17] According to the general definition given above, $t^*_{0,0}$ is defined by the equation $S_{0,0}(t^*_{0,0}) = \frac{1}{2}$.

that criterion, Q and Q^E coincide on a set of full measure.

Thus, our model of behavior gives rise to a pool at equal division in the standard dictator game if and only if the population contains sufficiently fair-minded people ($\bar{t} > t^*_{0,0}$). To appreciate why, consider the manner in which the single-crossing property fails: a larger transfer permits a dictator who cares more about fairness to distinguish himself from one who cares less about fairness if and only if $x < \frac{1}{2}$. Thus, $x = \frac{1}{2}$ serves as something of a natural boundary on chosen signals. In standard signaling environments (with single crossing), the D1 criterion isolates either separating equilibria or, if the range of potential choice is sufficiently limited, equilibria with pools at the upper boundary of the action set (Cho and Sobel [1990]). In our model, $\frac{1}{2}$ is not literally a boundary, and indeed there are equilibria in which some dictators transfer more than $\frac{1}{2}$. However, there is only limited scope in equilibrium for transfers exceeding $\frac{1}{2}$ (see Lemma 2 in the Appendix) and those possibilities do not survive the application of the D1 criterion. Accordingly, when $\bar{t}$ is sufficiently large, dictators who seek to distinguish themselves from those with lower values of t by giving more "run out of space, " and must therefore join a pool at $x = \frac{1}{2}$.[18]

Note that our theory accounts for the behavioral patterns listed in the introduction. First, provided that the some people are sufficiently fair-minded, there is a spike in the distribution of choices *precisely* at equal division, even if the prize is perfectly divisible. Second, no one transfers more than half the prize. Third, no one transfer slightly less than half the prize (recall that $S_{0,0}(t_0) < \frac{1}{2}$ for blended equilibria). Intuitively, if a dictator intends to divide the pie unequally, it makes no sense to divide it only *slightly* unequally, since negative inferences about his motives will overwhelm the tiny consumption gain.

Our theory also explains why greater anonymity for the dictator leads him to behave

[18] Despite some surface similarities, the mechanism producing a central pool in this model differs from those explored in Bernheim [1994] and Bernheim and Severinov [2003]. In those papers, the direction of imitation reverses when type passes some threshold; types in the middle are unable to adjust their choices to simultaneously deter imitation from the left and from the right. Here, higher types always try to deter imitation by lower types, but are simply unable to do that once x reaches $\frac{1}{2}$. The main result here is also cleaner in the following sense: in Bernheim [1994] and Bernheim and Severinov [2003], there is a range of possible equilibrium norms; here, equal division is the only possible equilibrium norm.

more selfishly and weakens the 50-50 norm. Presumably, treatments with less anonymity cause dictators to attach greater importance to social image. Formally, we say that $\widetilde{U}$ *attaches more importance to social image* than U if $\widetilde{U}(x,m,t) = U(x,m,t) + \phi(m)$, where ϕ is differentiable, and $\phi'(m)$ is strictly positive and bounded away from zero. The addition of the separable term $\phi(m)$ allows us to vary the importance of social image without altering the trade-off between consumption and equity.

The following result tells us that an increase in the importance attached to social image increases the extent to which dictators conform to the 50-50 norm:

Theorem 2: Assume $p = 0$, and that Assumption 1 holds. Suppose $\widetilde{U}$ attaches more importance to social image than U. Let $\widetilde{\pi}$ and π denote the measures of types choosing $x = \frac{1}{2}$ for $\widetilde{U}$ and U, respectively (based on the equilibrium action functions $\widetilde{Q}^E, Q^E \in \mathcal{Q}_1$). Then $\widetilde{\pi} \geq \pi$, with strict inequality when $\pi \in (0,1)$.

4 Analysis of the extended dictator game

Next we explore the theory's implications for our extended version of the dictator game. With $p > 0$ and x_0 close to zero, the distribution of *voluntary* choices has mass not only at $\frac{1}{2}$ (if $\bar{t}$ is sufficiently large), but also at x_0. Intuitively, the potential for nature to choose x_0 regardless of the dictator's type reduces the stigma associated with voluntarily choosing x_0. Moreover, as p increases, more and more dictator types are tempted to "hide" their selfishness behind nature's choice. That response mitigates the threat of imitation, thereby allowing higher types to reduce their gifts as well. Accordingly, the measure of types voluntarily choosing x_0 grows, while the measure of types choosing $\frac{1}{2}$ shrinks.

We will focus on equilibria involving action functions belonging to a restricted set, $\mathcal{Q}_2$. To simplify notation, we define $S^t \equiv S_{t,\max\{x_0,x^*(t)\}}$. The action function Q belongs to $\mathcal{Q}_2$ if and only if it falls into one of the following three categories:

Blended double-pool action function: There is some $t_0 \in (0,\bar{t})$ and $t_1 \in (t_0,\bar{t})$ with

$S^{t_0}(t_1) < \frac{1}{2}$ such that, for $t \in [0, t_0]$, we have $Q(t) = x_0$; for $t \in (t_0, t_1]$, we have $Q(t) = S^{t_0}(t)$; and for $t \in (t_1, \bar{t}]$, we have $Q(t) = \frac{1}{2}$.

Blended single-pool action function: There is some $t_0 \in (0, \bar{t})$ with $x^*(t_0) \geq x_0$ and $S^{t_0}(\bar{t}) < \frac{1}{2}$ such that, for $t \in [0, t_0]$, we have $Q(t) = x_0$, and for $t \in (t_0, \bar{t}]$, we have $Q(t) = S^{t_0}(t)$.

Double-pool action function: There is some $t_0 \in (0, \bar{t})$ such that, for $t \in [0, t_0]$, we have $Q(t) = x_0$, and for $t \in (t_0, \bar{t}]$, we have $Q(t) = \frac{1}{2}$.

We will refer to equilibria that employ such action functions as, respectively, *blended double-pool equilibria*, *blended single-pool equilibria*, and *double-pool equilibria*. In a blended double-pool equilibrium, type t_0 must be indifferent between pooling at x_0 and separating:

$$U\left(x_0, B\left(\widehat{H}_{t_0}^p\right), t_0\right) = U\left(\max\{x_0, x^*(t_0)\}, t_0, t_0\right) \ , \tag{2}$$

where $\widehat{H}_{t_0}^p$ is the CDF for types transferring x_0.[19] Also, type t_1 must be indifferent between separating and joining the pool choosing $\frac{1}{2}$:

$$U\left(\frac{1}{2}, B\left(H_{t_1}\right), t_1\right) = U\left(S^{t_0}(t_1), t_1, t_1\right) \ . \tag{3}$$

Finally, if $x_0 > 0$, type 0 must weakly prefer the lower pool to his first-best action combined with the worst possible inference:

$$U(0,0,0) \leq U(x_0, B\left(\widehat{H}_{t_0}^p\right), 0) \ . \tag{4}$$

In a blended single pool equilibrium, (2) and (4) must hold. Finally, in a double-pool equilibrium, expression (4) must hold; also, type t_0 must be indifferent between pooling at $\frac{1}{2}$ and pooling at x_0, and must weakly prefer both to all $x \in (x_0, \frac{1}{2})$ with revelation of its type:

$$U\left(x_0, B\left(\widehat{H}_{t_0}^p\right), t_0\right) = U\left(\frac{1}{2}, B\left(H_{t_0}\right), t_0\right) \geq U\left(\max\{x_0, x^*(t_0)\}, t_0, t_0\right) \ . \tag{5}$$

[19] Specifically, $\widehat{H}_t^p(t') \equiv \left(\frac{p}{p+(1-p)H(t)}\right) H(t') + \left(\frac{1-p}{p+(1-p)H(t)}\right) H\left(\min\{t, t'\}\right)$. Note that if $\max\{x_0, x^*(t_0)\} = x_0$, then $S^{t_0}(t_0) = x_0$. In that case, condition (2) simply requires $B\left(\widehat{H}_{t_0}^p\right) = t_0$, so that the outcome for t_0 is the same as separation.

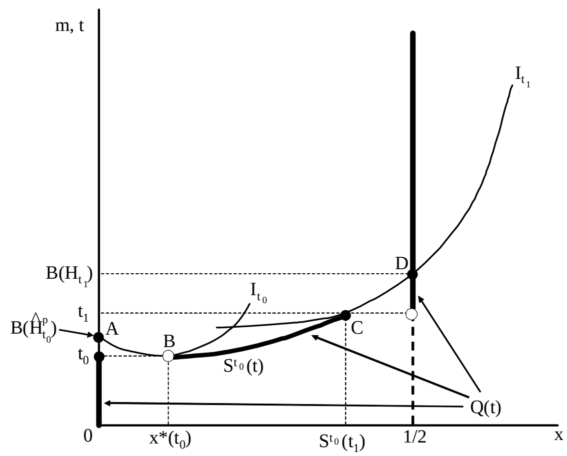

Figure 2: A Blended Double-Pool Equilibrium

Figure 2 illustrates a blended double-pool equilibrium for $x_0 = 0$. The indifference curve I_{t_0} indicates that type t_0 is indifferent between the lower pool (point A) and separating with its first best choice, $x^*(t_0)$ (point B). All types between t_0 and t_1 choose a point on the separating function generated using point B as the initial condition. The indifference curve I_{t_1} indicates that type t_1 is indifferent between separating (point C) and the upper pool at $x = \frac{1}{2}$ (point D). A blended single-pool equilibrium omits the pool at $\frac{1}{2}$, and a double-pool equilibrium omits the interval with separation of types.

Because (4) is required for all three types of equilibria described above, we will impose a condition on x_0 and p that guarantees it:

$$U(0,0,0) \leq U(x_0, \min_{t\in[0,\bar{t}]} B\left(\widehat{H}_t^p\right), 0) \tag{6}$$

One can show that $B\left(\widehat{H}_t^p\right)$ is continuous in t, so the minimization is well-defined; moreover, for $p > 0$, $\min_{t\in[0,\bar{t}]} B\left(\widehat{H}_t^p\right) > 0$. Therefore, for all $p > 0$, (6) is satisfied as long as x_0 is not too large. One can also show that, for any x_0 such that $U(x_0, B(H), 0) > U(0,0,0)$, (6) is satisfied for p sufficiently large.

The following theorem establishes the existence and uniqueness of equilibria within $\mathcal{Q}_2$ and justifies our focus on that set.

Theorem 3: Assume $p > 0$, that Assumptions 1 and 2 hold, and that (6) is satisfied.[20] Restricting attention to $\mathcal{Q}_2$, there exists a unique equilibrium action function, Q^E. If $\bar{t}$ is sufficiently large, Q^E is either a either a double-pool or blended double-pool action function. Moreover, there exists an inference mapping P^E such that (Q^E, P^E) satisfies the D1 criterion, and for any other equilibrium (Q, P) satisfying that criterion, Q and Q^E coincide on a set of full measure.

The unique equilibrium action function in $\mathcal{Q}_2$ has several notable properties. For *voluntary* choices, there is always mass at x_0. Nature's exogenous choice of x_0 induces players to "hide" their selfishness by mimicking that choice. There is never positive mass at any other choice except $\frac{1}{2}$. As before, there is a gap in the distribution of choices just below $\frac{1}{2}$.[21] In addition, one can show that both t_0 and t_1 are monotonically increasing in p. Consequently, as p increases, the mass at x_0 grows, and the mass at $x = \frac{1}{2}$ shrinks. Formally:

Theorem 4: Assume $p > 0$, that Assumptions 1 and 2 hold, and that (6) is satisfied. Let π_0 and π_1 denote the measures of types choosing $x = x_0$ and $x = \frac{1}{2}$, respectively (based on the equilibrium action function $Q^E \in \mathcal{Q}_2$). Then π_0 is strictly increasing in p, and π_1 is decreasing (strictly if positive) in p.[22]

After circulating an earlier draft of this paper, we became aware of work by Dana et al. [2006] and Broberg et al. [2007], which shows that many dictators are willing to sacrifice part of the total prize to opt out of the game, provided that the decision is not revealed to recipients. Though we did not develop our theory with those experiments in mind, it

[20] With some additional arguments, our analysis extends to arbitrary p and x_0. The possible equilibrium configurations are similar to those described in the text, except that there may also be an interval of separation involving types with t near zero who chose transfers below x_0 along $S_{0,0}$. For some parameter values, existence may be problematic unless one slightly modifies the game, e.g., by allowing the dictator to reveal his responsibility for the transfer.

[21] One can also show that a gap just above x_0 definitely forms for p sufficiently close to unity, and definitely does not form for p sufficiently close to zero. However, since we do not attempt to test those implications, we omit a formal demonstration for the sake of brevity.

[22] One can also show that the measure of types choosing $x = x_0$ converges to zero as p approaches zero; see Andreoni and Bernheim [2007].

provides an immediate explanation. Opting out permits the dictator to avoid negative inferences while acting selfishly. In that sense, opting out is similar (but not identical) to choosing an action that could be attributable to nature. Not surprisingly, a positive mass of dictator types take that option in equilibrium. For details, see Appendix A (on-line).

5 Experimental Evidence

We designed a new experiment to test the theory's most direct implications: increasing p should increase the mass of dictators who choose any given x_0 (close to zero) and reduce the mass who split the payoff equally. Thus, we examine the effects of varying both p and x_0.

5.1 Overview of the experiment

We divide subjects into pairs, with partners and roles assigned randomly. Each pair splits a \$20 prize. To facilitate interpretation, we renormalize x, measuring it on a scale of 0 to 20. Thus, equal division corresponds to $x = 10$ rather than $x = 0.5$. Dictators, recipients, and outcomes are publicly identified at the conclusion of the experiment to heighten the effects of social image. For our purposes, there is no need to distinguish between intrinsic concern for an audience's reaction and concern arising from subsequent social interaction.[23]

We examine choices for four values of p (0, 0.25, 0.5, and 0.75) and two values of x_0 (0 and 1). Identifying the distribution of voluntary choices for eight parameter combinations requires a great deal of data.[24] One possible approach is to use the strategy method: ask each dictator to identify binding choices for several games, in each case conditional on nature not intervening, and then choose one game at random to determine the outcome. Unfortunately, that approach raises two serious concerns. First, in piloting the study, we discovered that subjects tend to focus on *ex ante* fairness – that is, the equality of expected payoffs

[23] A similar statement applies to concerns involving experimenter demand effects in dictator games (see, e.g., List [2007]). Our experiment creates demand effects that mirror those present in actual social situations. Because they are the objects of our study, we do not regard them as confounds.

[24] Suppose, for example, that we wish to have 30 observations of voluntary choices for each parameter combination. If each pair of subjects played one game, the experiment would require 1,000 subjects and \$15,000 in subject payments.

before nature's move. If a dictator knows that nature's intervention will favor him, he may compensate by choosing a strategy that favors the recipient when nature does not intervene. While that phenomenon raises some interesting questions concerning *ex ante* versus *ex post* fairness, concerns for *ex ante* fairness are properly viewed as confounds in the context of our current investigation. Second, the strategy method potentially introduces unintended and confounding audience effects. If a subject views the experimenter as part of the audience, the possibility that the experimenter will make inferences about the subject's character from his *strategy* rather than from the outcome may influence his choices. Our theory assumes the relevant audience lacks that information.

We address those concerns through the following measures. (1) We use the strategy method only to elicit choices for different games, and not to elicit the subject's complete strategy for a game. For each game, the dictator is only asked to make a choice if he has been informed that his choice will govern the outcome. Thus, within each game, each decision is made *ex post* rather than *ex ante*, so there is no risk that the experimenter will draw inferences from portions of strategies that are never executed. (2) We modify the extended dictator game by making nature's choice symmetric: nature intervenes with probability p, transferring x_0 and $20 - x_0$ with equal probabilities ($p/2$). That symmetry neutralizes the tendency among dictators to compensate for any *ex ante* asymmetry in nature's choice. Notably, this modification does not alter the theoretical results described in Section 4.[25] (3) Our procedures guarantee that no one can associate any dictator with his or her *strategy*. We make that point evident to subjects. (4) Subjects' instructions emphasize that everyone present in the lab will observe the *outcome* associated with each dictator. We thereby focus the subjects' attention on the revelation of particular information to a particular audience. See Appendix B (on-line) for details concerning our experimental protocol, and Appendix D

[25] For the purpose of constructing an equilibrium, the mass at $20 - x_0$ can be ignored. It is straightforward to demonstrate that all types will prefer their equilibrium choices to that alternative, given it will be associated with the social image $B(H)$. They prefer their equilibrium choices to the action chosen by $\bar{t}$, and must prefer that choice to $20 - x_0$, because it provides more consumption, less inequality, and a better social image.

(on-line) for the subjects' instructions.

We examine two experimental conditions, one with $x_0 = 0$ ("condition 0") and one with $x_0 = 1$ ("condition 1"). Each pair of subjects is assigned to a single condition, and each dictator makes choices for all four values of p. Thus, we identify the effects of x_0 from variation between subjects, and the effects of p from variation within subjects. When $p = 0$ we should observe the same distribution of choices for both conditions, including a spike at $x = 10$, a 50-50 split. For $p = 0.25$, a second spike should appear, located at $x = 0$ for condition 0 and at $x = 1$ for condition 1. As we increase p to 0.50 and 0.75 the spikes at 10 should shrink and the spikes at x_0 should grow.

The subjects were 120 volunteers from undergraduates economic courses at the University of Wisconsin–Madison in March and April 2006. We divided them into 30 pairs for each condition; unexpected attrition left 29 pairs for condition 1. Each subject maintained the same role (dictator or recipient) throughout.

The closest existing parallel to our experiment is the "plausible deniability" treatment of Dana et al. [2007], which differs from ours in the following ways: (a) the probability that nature intervenes depends on the dictator's response time, (b) only two choices are available, and nature chooses both with equal probability, so that no choice is unambiguously attributable to the dictator, and (c) the effects of variations in the likelihood of intervention and the distribution of nature's choice are not examined.

5.2 Main Findings

Figures 3 and 4 show the distributions of dictators' voluntary choices in condition 0 ($x_0 = 0$) and condition 1 ($x_0 = 1$), respectively. For ease of presentation, we group values of x into five categories: $x = 0$, $x = 1$, $2 \leq x \leq 9$, $x = 10$, and $x > 10$.[26] In both conditions, as in previous experiments, transfers exceeding half the prize are rare.[27]

These figures provide striking confirmation of our theory's predictions. Look first at

[26] Although subjects were permitted to choose any division of the $20 prize, and although they were provided with hypothetical examples in which dictators chose allocations that involved fractional dollars, all

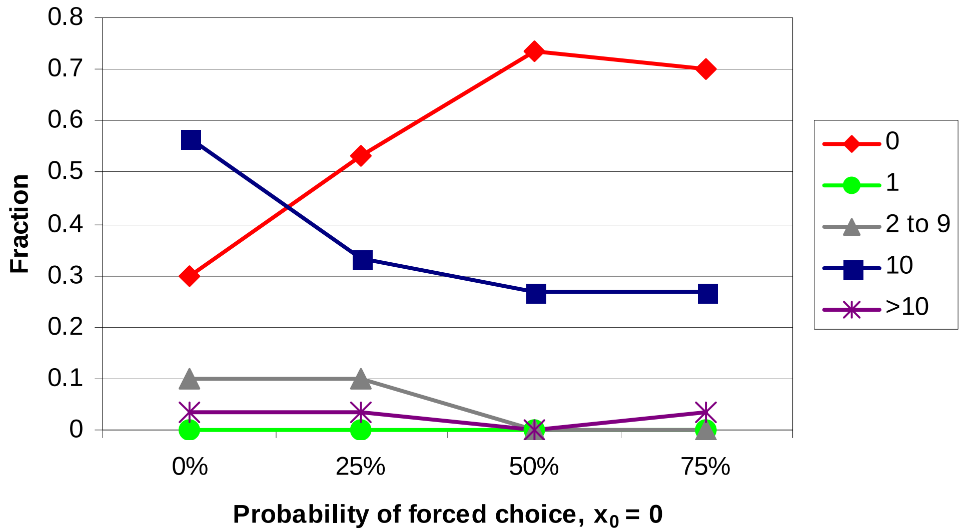

Figure 3: Distribution of Amounts Allocated to Partners, Condition 0

Figure 3 (condition 0). For $p = 0$ we expect a spike at $x = 10$. Indeed, 57 percent of dictators divided the prize equally. Consistent with results obtained from previous dictator experiments, a substantial fraction of subjects (30 percent) chose $x = 0$.[28] As we increase p we expect the spike at $x = 10$ to shrink and the spike at $x = 0$ to grow. That is precisely what happens. Note also that no subject chose $x = 1$ for any value of p.

Look next at Figure 4 (condition 1). Again, for $p = 0$ we expect a spike at $x = 10$.

chosen allocations involved whole dollars.

[27] For condition 0, there were three violations of this prediction (involving two subjects) out of 139 total choices. One subject gave away $15 when $p = 0$. A second subject gave away $15 in one of two instances with $p = 0.25$ (but gave away $10 in the other instance), and gave away $11 when $p = 0.75$. For condition 1, there were only two violations (involving just one subject) out of 134 total choices. That subject chose $x = 19$ with $p = 0.5$ and 0.75. When asked to explain her choices on the post-experiment questionnaire, she indicated that she alternated between giving $1 and $19 in order to "give me and my partner equal opportunities to make the same $." Thus, despite our precautions, she was clearly concerned with *ex ante* fairness. The total numbers of observations reported here exceeds the numbers reported in Tables 1 and 2 because here we do not average duplicative choices for $p = 0.25$.

[28] For instance, the fraction of dictators who kept the entire prize was 35 percent in Forsythe et al. (1994) and 33 percent in Bohnet and Frey (1999). In contrast to our experiment, however, no dictators kept the entire prize in Bohnet and Frey's "two-way identification" condition. One potentially important difference is that Bohnet and Frey's subjects were all students in the same course, whereas our subjects were drawn from all undergraduates enrolled in economics courses at the University of Wisconsin, Madison.

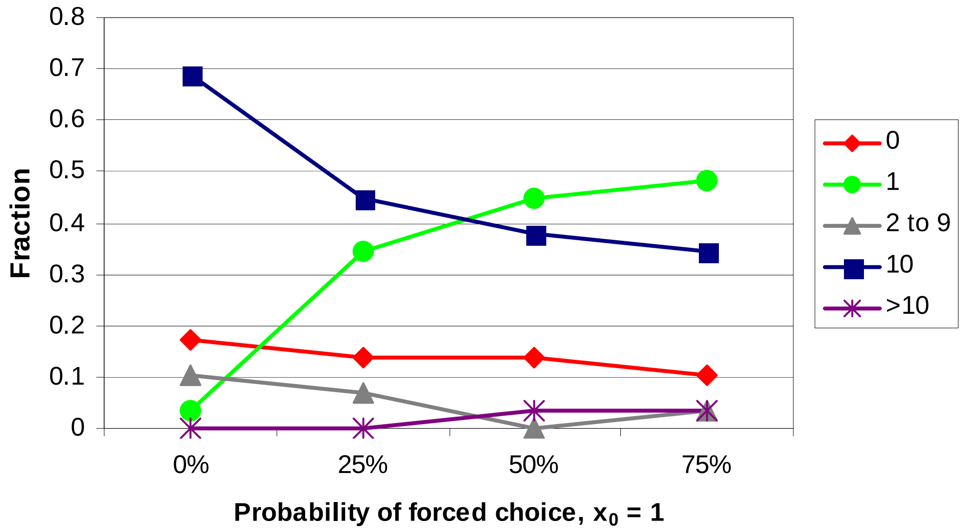

Figure 4: Distribution of Amounts Allocated to Partners, Condition 1

Indeed, 69 percent of dictators divide the prize equally, while 17 percent keep the entire prize ($x = 0$), and only 3% (one subject) chose $x = 1$. As we increase p the spike at $x = 10$ once again shrinks. In this case, however, a new spike emerges at $x = 1$. As p increases to 0.75, the fraction of dictators choosing $x = 1$ rises steadily from 3 percent to 48 percent, while the fraction choosing $x = 10$ falls steadily from 69 percent to 34 percent. Notably, the fraction choosing $x = 0$ *falls* in this case from 17 percent to 10 percent. Once again, the effect of variations in p on the distribution of choices is dramatic, and exactly as predicted.

Table 1 addresses the statistical significance of these effects by reporting estimates of two random-effects probit models. The specifications in columns (1)-(2) describe the probability of selecting $x = x_0$; those in column (3)-(4) describe the probability of selecting $x = 10$, equal division. The explanatory variables include indicators for $p \geq 0.25$, $p \geq 0.5$, $p = 0.75$, and $x_0 = 1$ (with $p \geq 0$ and $x_0 = 0$ omitted). In all cases, we report marginal effects at mean values, including the mean of the unobserved individual heterogeneity. We pool data from both conditions; similar results hold for each condition separately.

TABLE 1
Random effects probit models:
Marginal effects for regressions describing
(1)-(2) the probability of choosing $x = x_0$ and
(3)-(4) the probability choosing $x = 10$ (equal division).†

	(1)	(2)	(3)	(4)
	$\Pr(x = x_0)$	$\Pr(x = x_0)$	$\Pr(x = 10)$	$\Pr(x = 10)$
$p \geq 0.25$	0.467***	0.467***	-0.532***	-0.532***
	(0.110)	(0.110)	(0.124)	(0.124)
$p \geq 0.50$	0.346***	0.345***	-0.175*	-0.196**
	(0.129)	(0.113)	(0.133)	(0.116)
$p = 0.75$	-0.002		-0.042	
	(0.132)		(0.130)	
$x_0 = 1$	-0.524***	-0.524***	0.224	0.224
	(0.179)	(0.179)	(0.219)	(0.219)
Observations	236	236	236	236

† Standard errors in parentheses.
Significance: *** $\alpha < 0.01$, ** $\alpha < 0.05$, * $\alpha < 0.1$, one-sided tests

The coefficients in column (1) imply that there is a statistically significant increase in pooling at $x = x_0$ when p rises from 0 to 0.25 and from 0.25 to 0.5 ($\alpha < .01$, one tailed t-test), but not when p rises from 0.5 to 0.75. The significant negative coefficient for $x_0 = 1$ may reflect the choices of a subset of subjects who are unconcerned with social image, and who therefore transfer nothing. Dropping the insignificant $p = 0.75$ indicator has little effect on the other coefficients (column (2)). The coefficients in column (3) imply that there is a statistically significant decline in pooling at $x = 10$ when p rises from 0 to 0.25 ($\alpha < .01$, one tailed t-test) and from 0.25 to 0.5 ($\alpha < .1$, one tailed t-test), but not when p rises from 0.5 to 0.75. As shown in column (4), the effect of an increase in p from 0.25 to 0.5 on pooling at $x = 10$ becomes even more statistically significant when we drop the insignificant $p = 0.75$ indicator ($\alpha < .05$, one tailed t-test).

As an additional check on the model's predictions, we compare choices across the two conditions for $p = 0$. As predicted, we find no significant difference between the two distributions (Mann-Whitney $z = 0.670, \alpha < 0.50$, Kolmogorov-Smirnov $k = 0.13, \alpha < 0.95$). The higher fraction of subjects choosing $x = 0$ in condition 0 (30 percent versus 17 percent) and the higher fraction choosing $x = 1$ in condition 1 (3 percent versus 0 percent) suggest a modest anchoring effect, but that pattern is also consistent with chance (comparing choices of $x = 0$, we find $t = 1.145, \alpha < 0.26$).

Our theory implies that, as p increases, a subject in condition 0 will not increase his gift, x. Five of 30 subjects violate that monotonicity prediction; for each, there is one violation. The same prediction holds for condition 1, with an important exception: an increase in p could induce a subject to switch from $x = 0$ to $x = 1$. We find four violations of monotonicity for condition 1, but two involve switches from $x = 0$ to $x = 1$. Thus, problematic violations of monotonicity are relatively uncommon (11.9 percent of subjects).

As a further check on the validity of our main assumptions concerning preferences, and to assess whether our model generates the right predictions for the right reasons, we also examined data on attitudes and motivations obtained from a questionnaire administered after subjects completed the experiment. Self-reported motivations correlated with their choices in precisely the manner our theory predicts. For details, see Appendix C (on-line).

6 Concluding Comments

We have proposed and tested a theory of behavior in the dictator game that is predicated on two critical assumptions: first, people are fair-minded to varying degrees; second, people like others to see them as fair. We have shown that this theory accounts for previously unexplained behavioral patterns. It also has sharp and testable ancillary implications which new experimental data confirm.

Narrowly interpreted, this study enriches our understanding of behavior in the dictator game. More generally, it provides a theoretical framework that potentially accounts for the prevalence of the equal division norm in real-world settings. Though our theory may not provide the best explanation for all 50-50 norms, it nevertheless deserves serious consideration in many cases. In addition, this study underscores both the importance and feasibility of studying audience effects, which potentially affect a wide range of real economic choices, with theoretical and empirical precision.

Appendix

Lemma 1: In equilibrium, $G\left(Q(t)-\frac{1}{2}\right)$ is weakly increasing in t.

Proof: Consider two types, t and t' with $t<t'$. Suppose type t chooses x earning image m, while t' chooses x' earning image m'. Let $f=F(1-x,m)$, $f'=F(1-x',m')$, $g=G\left(x-\frac{1}{2}\right)$, and $g'=G\left(x'-\frac{1}{2}\right)$. Mutual non-imitation requires $f'+t'g'\geq f+t'g$ and $f'+tg'\leq f+tg$; thus, $(g'-g)(t'-t)\geq 0$. Since $t'-t>0$, it follows that $g'-g\geq 0$. □

Lemma 2: Suppose $Q(t)>\frac{1}{2}$. Define $x'<\frac{1}{2}$ as the solution (if any) to $G\left(x'-\frac{1}{2}\right)=G\left(Q(t)-\frac{1}{2}\right)$. Then for all $t'>t$, $Q(t')\in\{x',Q(t)\}$ if $p=0$, and $Q(t')\in\{x',Q(t),x_0\}$ if $p>0$.[29]

Proof: According to Lemma 1, $G\left(Q(t')-\frac{1}{2}\right)\geq G\left(Q(t)-\frac{1}{2}\right)$. To prove this lemma, we show that the inequality cannot be strict unless $p>0$ and $Q(t')=x_0$. Suppose on the contrary that it is strict for some t', and either $p=0$, or $p>0$ and $Q(t')\neq x_0$. Let $t^0=\inf\{\tau\mid Q(\tau)=Q(t')\}$. It follows from Lemma 1 that, for all $t''>t^0$, $Q(t'')\neq Q(t)$. Thus, $B\left(P_{Q(t)}\right)\leq t^0\leq B\left(P_{Q(t')}\right)$. Since G is single-peaked, $Q(t')<Q(t)$. Thus, all types, including t, prefer $Q(t')$ to $Q(t)$, a contradiction. □

Lemma 3: Assume $z\geq x^*(r)$. (a) $S_{r,z}(t)>x^*(t)$ for $t>r$. (b) For all $t>r$, $S'_{r,z}(t)>0$. (c) If $S_{r,z}(t')\leq\frac{1}{2}$ and $S_{r,z}(t'')\leq\frac{1}{2}$, type $t'\geq 0$ strictly prefers $(x,m)=(S_{r,z}(t'),t')$ to $(S_{r,z}(t''),t'')$. (d) There exists $t^*_{r,z}>r$ such that $S_{r,z}(t^*)=\frac{1}{2}$. (e) $S_{r,z}(t)$ is increasing in z and continuous in r and z.

Proof: (a) First consider the case of $z>x^*(r)$. Suppose the claim is false. Then, since the solution to (1) must be continuous, there is some t' such that $S_{r,z}(t')=x^*(t')$ and $S_{r,z}(t)>x^*(t)$ for $r\leq t<t'$. As t approaches t' from below, $S'_{r,z}(t_k)$ increases without bound (see (1)). In contrast, given our assumptions about F and G, the derivative of $x^*(t)$ is bounded within any neighborhood of t'. But then $S_{r,z}(t)-x^*(t)$ must increase over some interval (t'',t') (with $t''<t'$), which contradicts $S_{r,z}(t')-x^*(t')=0$.

Now consider the case of $z=x^*(r)$. If $U_1(x^*(r),r,r)=0$, then $S'_{r,z}(r)$ is infinite (see (1))

[29] As a corollary, it follows that there is at most one value of x greater than $\frac{1}{2}$ chosen in any equilibrium.

while $\left.\frac{dx^*(t)}{dt}\right|_{t=r}$ is finite. If $U_1(x^*(r), r, r) < 0$ (which requires $x^*(r) = 0$), then $S'_{r,z}(r) > 0$ while $\left.\frac{dx^*(t)}{dt}\right|_{t=r} = 0$. In either case, $S_{r,z}(t) > x^*(t)$ for t slightly larger than r; one then applies the argument in the previous paragraph.

(b) Given (1), the claim follows directly from part (a).

(c) Consider t' and t'' with $S_{r,z}(t')$, $S_{r,z}(t'') \leq \frac{1}{2}$. Assume that $t' < t''$. Then

$$U\left(S_{r,z}(t''), t'', t'\right) - U\left(S_{r,z}(t'), t', t'\right) = \int_{t'}^{t''} \frac{dU\left(S_{r,z}(t), t, t'\right)}{dt} dt$$

$$< \int_{t'}^{t''} \left\{ \left[tG'\left(S_{r,z}(t) - \frac{1}{2}\right) - F_1\left(1 - S_{r,z}(t), t\right) \right] S'_{r,z}(t) + F_2\left(1 - S_{r,z}(t), t\right) \right\} dt = 0$$

where the inequality follows from $S_{r,z}(t) < \frac{1}{2}$, and where the final equality follows from (1). The argument for $t'' < t'$ is symmetric.

(d) Assume the claim is false. Because $S_{r,z}(t)$ is continuous, we have $S_{r,z}(t) \in (0, \frac{1}{2})$ for arbitrarily large t. Using the boundedness of F_1 (implied by the continuous differentiability of F) and the unboundedness of F in its second argument, we have $\lim_{t\to\infty} \left[U\left(S_{r,z}(t), t, r\right) - U(S_{r,z}(r), r, r)\right] > 0$, which contradicts part (c).

(e) If $z > z'$, then $S_{r,z}(r) > S_{r,z'}(r)$. Because the two trajectories are continuous and (for standard reasons) cannot intersect, we have $S_{r,z}(t) > S_{r,z'}(t)$ for all $t > r$. Continuity in r and z follows from standard properties of the solutions of differential equations. □

Proof of Theorem 1

Step 1A: If $\bar{t} > t^*_{0,0}$, there is at most one equilibrium action function in $\mathcal{Q}_1$, and it must be either a central pooling or blended equilibrium action function.

We can rule out the existence of an efficient separating equilibrium: part Lemma 3(b) implies that $G\left(S_{0,0}(t) - \frac{1}{2}\right)$ is strictly decreasing in t for $t > t^*_{0,0}$, so according to Lemma 1, $S_{0,0}$ cannot be an equilibrium action function.

For $t \in [0, t^*_{0,0}]$, define $\psi(t)$ as the solution to $U(S_{0,0}(t), t, t) = U\left(\frac{1}{2}, \psi(t), t\right)$. The existence and uniqueness of a solution are trivial given our assumptions; continuity of ψ follows from continuity of $S_{0,0}$ and U. In addition, $\psi'(t) = \left[G\left(S_{0,0}(t) - \frac{1}{2}\right) - G\left(0\right)\right] \left[F_2(\frac{1}{2}, \psi(t))\right]^{-1}$, which implies that $\psi(t)$ is strictly decreasing in t on $[0, t^*_{0,0})$. Note that we can rewrite the weak preference condition for a central pooling equilibrium as $\psi(0) \leq B(H)$, and the indifference condition for a blended equilibrium as $\psi(t_0) = B(H_{t_0})$ for $t_0 \in (0, t^*_{0,0})$.

First suppose $\psi(0) \leq B(H)$. $B(H_t)$ is plainly strictly increasing in t and $\psi(t)$ is strictly decreasing, so there is no $t_0 \in (0, t^*_{0,0})$ for which $\psi(t_0) = B(H_{t_0})$, and hence no blended equilibrium action function; if there is an equilibrium action function in $\mathcal{Q}_1$, it employs the unique central pooling action function. Next suppose $\psi(0) > B(H)$, so there is no central pooling equilibrium. Note that $\psi(t^*_{0,0}) = t^*_{0,0} < B(H_{t^*_{0,0}})$. The existence of a unique $t_0 \in (0, t^*_{0,0})$ with $\psi(t_0) = B(H_{t_0})$ follows from the continuity and monotonicity of $B(H_t)$ and $\psi(t)$ in t. Thus, there is at most one blended equilibrium action function.

Step 1B: If $\bar{t} \leq t^*_{0,0}$, there is at most one equilibrium action function in $\mathcal{Q}_1$, and it must be an efficient differentiable separating action function.

Notice that $B(H_{\bar{t}}) = \bar{t} \leq \psi(\bar{t})$. Given the monotonicity of $B(H_t)$ and $\psi(t)$ in t, we have $B(H_t) < \psi(t)$ for all $t \in [0, \bar{t})$, which rules out both blended equilibria and central pooling equilibria. There is at most one efficient differentiable separating equilibrium action function because the solution to (1) with initial condition $(r, z) = (0, 0)$ is unique. □

Step 1C: There exists an equilibrium action function $Q^E \in \mathcal{Q}_1$ and an inference function P^E such that (Q^E, P^E) satisfies the D1 criterion.

Suppose $\bar{t} \leq t^*_{0,0}$. Let $Q^E = S_{0,0}$. Choose any inference function P^E such that P^E_x places probability only on $S^{-1}_{0,0}(x)$ for $x \in [0, S_{0,0}(\bar{t})]$ (which guarantees consistency with Q^E), and only on M_x (defined at the outset of Section 3) for $x > S_{0,0}(\bar{t})$. Lemma 3(c) guarantees that, for each t, $Q^E(t)$ is optimal within the set $[0, S_{0,0}(\bar{t})]$. Since (i) t prefers its equilibrium outcome to $(S_{0,0}(\bar{t}), \bar{t})$, (ii) $S_{0,0}(\bar{t}) \geq S_{0,0}(t) > x^*(t)$ (Lemma 3(a) and (b)), and (iii) $B(P^E_x) \leq \bar{t}$ (Assumption 1, part (2)), we know that t also prefers its equilibrium outcome to all $(x, B(P^E_x))$ for $x > S_{0,0}(\bar{t})$. Thus, (Q^E, P^E) satsifies the D1 criterion.

Now suppose $\bar{t} > t^*_{0,0}$ and $\psi(0) \leq B(H)$. Let $Q^E(t) = \frac{1}{2}$ for all t. Consider the inference function P^E such that $P^E_{1/2} = H$ (which guarantees consistency with Q^E) and P^E_x places all weight on type $t = 0$ for each $x \neq \frac{1}{2}$. It is easy to verify that $0 \in M_x$ for all $x \neq \frac{1}{2}$, and that all types t prefer $(\frac{1}{2}, H)$ to $(x, 0)$. Thus, (Q^E, P^E) satisfies the D1 criterion.

Finally suppose $\bar{t} > t^*_{0,0}$ and $\psi(0) > B(H)$. Let t_0 satisfy $\psi(t_0) = B(H_{t_0})$ (step 1A showed that a solution exists within $(0, t^*_{0,0})$). For $t \in [0, t_0)$, let $Q^E(t) = S_{0,0}(t)$, and for $t \in [t_0, \bar{t}]$, let $Q^E(t) = \frac{1}{2}$. Choose any inference function P^E such that (i) P^E_x places probability only on $S^{-1}_{0,0}(x)$ for $x \in [0, S_{0,0}(t_0)]$, (ii) $P^E_{1/2} = H_{t_0}$, and (iii) P^E_x places probability only on $M_x \cap [0, t_0]$ for $x \in (S_0(t_0), \frac{1}{2}) \cup (\frac{1}{2}, 1]$. It is easy to verify that $M_x \cap [0, t_0]$ is non-empty for $x \in (S_{0,0}(t_0), \frac{1}{2}) \cup (\frac{1}{2}, 1]$ (because for $t > t_0$, $m_x(t) > m_x(t_0)$), so the existence of such an inference function is guaranteed. Parts (i) and (ii) guarantee that P^E is consistent with Q^E. It is easy to verify (based on Lemma 3(c) and a simple additional argument) that, for each t, $Q^E(t)$ is optimal within the set $[0, S_{0,0}(t_0)] \cup \{\frac{1}{2}\}$. For all $x \in (S_{0,0}(t_0), \frac{1}{2}) \cup (\frac{1}{2}, 1]$, we have $B(P^E_x) \leq t_0$, from which it follows (by another simple argument) that no type prefers $(x, B(P^E_x))$ to its equilibrium outcome. Thus, (Q^E, P^E) satisfies the D1 criterion.

Step 1D: If an equilibrium (Q, P) satisfies the D1 criterion, there is no pool at any action other than $\frac{1}{2}$.

Suppose there is a pool that selects an action $x' \neq \frac{1}{2}$. Select some t' from the pool such that $t' > B(P_{x'})$. We claim that, for any x'' sufficiently close to x' with $G\left(x'' - \frac{1}{2}\right) > G\left(x' - \frac{1}{2}\right)$, $B(P_{x''}) \geq t'$. Assuming x'' is chosen by some type in equilibrium, the claim follows from Lemma 1. Assuming x'' is not chosen by any type in equilibrium, it is easy to check that $m_{x''}(t'') > m_{x''}(t')$ for any $t'' < t'$; with x'' sufficiently close to x', we have $m_{x''}(t') < \bar{t}$, which then implies $t'' \notin M_{x''}$, and hence $B(P_{x''}) \geq t'$. The lemma follows from the claim, because t' would deviate at least slightly toward $\frac{1}{2}$.

Step 1E: If an equilibrium (Q, P) satisfies the D1 criterion, type $t = 0$ selects either $x = 0$ or $x = \frac{1}{2}$.

Suppose $Q(0) \notin \{0, \frac{1}{2}\}$. By step 1D, $P_{Q(0)}$ places probability one on type 0. But then $U(0, B(P_0), 0) \geq U(0, 0, 0) > U\left(Q(0), B\left(P_{Q(0)}\right), 0\right)$, which contradicts the premise that $Q(0)$ is optimal for type 0.

Step 1F: For any equilibrium (Q, P) satisfying the D1 criterion, Q and Q^E (the unique equilibrium action function within $\mathcal{Q}_1$) coincide on a set of full measure.

Lemma 2 and step 1D together imply $Q(t) \leq \frac{1}{2}$ for all $t \in [0, \bar{t})$. Let $t_0 = \sup\{t \in [0, \bar{t}] \mid Q(t) < \frac{1}{2}\}$ (if the set is empty, then $t_0 = 0$).

We claim that $Q(t) = S_{0,0}(t)$ for all $t \in [0, t_0)$. By Lemma 1, $Q(t)$ is weakly increasing on $t \in [0, \bar{t})$; hence $Q(t) < \frac{1}{2}$ for $t \in [0, t_0)$. By step 1D, $Q(t)$ fully separates all types in $[0, t_0)$, and is therefore strictly increasing on that set. Consider the restricted game in which the type space is $[0, t_0 - \varepsilon]$ and the dictator chooses $x \in [0, Q(t_0 - \varepsilon)]$ for small $\varepsilon > 0$. It is easy to construct another signaling model for which the type space is $[0, t_0 - \varepsilon]$, the dictator chooses $x \in \mathbb{R}$, preferences are the same as in the original game for $(x, m) \in [0, Q(t_0 - \varepsilon)] \times [0, t_0 - \varepsilon]$, and conditions (1)-(5) and (7) of Mailath [1987] are satisfied on the full domain $\mathbb{R} \times [0, t_0 - \varepsilon]$. Theorem 2 of Mailath [1987] therefore implies that $Q(t)$ (which we have shown achieves full separation on $[0, t_0)$) must satisfy (1) on $[0, t_0 - \varepsilon]$ for all $\varepsilon > 0$. The desired conclusion then follows from step 1E, which ties down the initial condition, $Q(0) = 0$.

There are now three cases to consider: (i) $t_0 = 0$, (ii) $t_0 \in (0, \bar{t})$, and (iii) $t_0 = \bar{t}$. In case (i), we know that $Q(t) = \frac{1}{2}$ for $t \in (0, \bar{t}]$ ($\bar{t}$ is included by Lemma 1). It is easy to check that if Q is an equilibrium, then so is $Q^*(t) = \frac{1}{2}$ for all $t \in [0, \bar{t}]$ (for the same inferences, if type 0 has an incentive to deviate from Q^*, then some type close to zero would have an incentive to deviate from Q). In case (ii), we know that $Q(t) = S_{0,0}(t)$ for $t \in [0, t_0)$ and $Q(t) = \frac{1}{2}$ for $t \in (t_0, \bar{t}]$. It is easy to check that if Q is an equilibrium, then so is $Q^*(t) = Q(t)$ for $t \neq t_0$ and $Q^*(t_0) = S_{0,0}(t_0)$. In case (iii), we know that $Q(t) = S_{0,0}(t)$ for $t \in [0, \bar{t})$. It is easy to check that if Q is an equilibrium, then so is $Q^*(t) = S_{0,0}(t)$ for $t \in [0, \bar{t}]$. In each case, $Q^* \in \mathcal{Q}_1$, and Q and Q^* coincide on a set of full measure. □

Proof of Theorem 2: First we claim that $S_{0,0}(t) < \widetilde{S}_{0,0}(t)$ for all t. It is easy to check that $S'_{0,0}(0) < \widetilde{S}'_{0,0}(0)$, so $S_{0,0}(t) < \widetilde{S}_{0,0}(t)$ for small t. If the claim is false, then since the separating functions are continuous, $t' = \min\left\{t > 0 \mid S_{0,0}(t) = \widetilde{S}_{0,0}(t)\right\}$ is well-defined. It is easy to check that $S'_{0,0}(t') < \widetilde{S}'_{0,0}(t')$; moreover, because the slopes of the separating functions vary continuously with t, there is some $t'' < t'$ such that $S'_{0,0}(t) < \widetilde{S}'_{0,0}(t)$ for all $t \in [t'', t']$.

But since $S_{0,0}(t'') < \widetilde{S}_{0,0}(t'')$, we must then have $S_{0,0}(t') < \widetilde{S}_{0,0}(t')$, a contradiction.

Define $\psi(t)$ as in step 1A of the proof of Theorem 1, and define $\widetilde{\psi}(t)$ for $\widetilde{U}$ analogously. Note that for $t \in (0, t^*_{0,0})$, $\widetilde{U}(\widetilde{S}_{0,0}(t), t, t) < U(S_{0,0}(t), t, t) + \phi(t) = U\left(\frac{1}{2}, \psi(t), t\right) + \phi(t) < \widetilde{U}\left(\frac{1}{2}, \psi(t), t\right)$. It follows that $\widetilde{\psi}(t) < \psi(t)$.

If $\pi = 1$, then $\psi(0) \leq B(H)$, so $\widetilde{\psi}(0) < B(H)$, which implies $\widetilde{\pi} = 1$ (see step 1A of the proof of Theorem 1). If $\pi \in (0,1)$, then $\psi(0) > B(0)$, and there is a unique blended equilibrium for which t_0 solves $B(H_{t_0}) = \psi(t_0)$. In that case, either $\widetilde{\psi}(0) \leq B(0)$, which implies $\widetilde{\pi} = 1 > \pi$, or $\widetilde{\psi}(0) > B(0)$ and $B(H_{t_0}) > \widetilde{\psi}(t_0)$, which implies (given the monotonicity properties of B and $\widetilde{\psi}$) $B(H_{\widetilde{t}_0}) = \widetilde{\psi}(\widetilde{t}_0)$ for $\widetilde{t}_0 < t_0$, and hence $\widetilde{\pi} > \pi$. □

Proof of Theorem 3

Step 3A: Equation (2) has a unique solution, $t^*_0 \in (0, \bar{t})$.

Define the function $\xi(t)$ as the solution to $F\left(1 - x_0, \xi(t)\right) + tG\left(x_0 - \frac{1}{2}\right) = F\left(1 - \max\{x_0, x^*(t)\}, t\right) + tG\left(\max\{x_0, x^*(t)\} - \frac{1}{2}\right)$. It is easy to check that for $t \in [0, \bar{t}]$, $\xi(t)$ exists and satisfies $\xi(t) \geq t$, with strict inequality if $x^*(t) > x_0$. Note that we can rewrite (2) as $\xi(t_0) = B\left(\widehat{H}^p_{t_0}\right)$. Also note that $\xi(0) = 0 < B(H) = B\left(\widehat{H}^p_0\right)$; furthermore, $\xi(\bar{t}) \geq \bar{t} > B(H) = B\left(\widehat{H}^p_{\bar{t}}\right)$. Thus, by continuity, there must exist at least one value of $t_0 \in (0, \bar{t})$ satisfying (2).

Now suppose, contrary to the claim, that there are two solutions to (2), t' and t'', with $t' > t''$. Define a CDF $L(t) \equiv \frac{H(\min\{t,t'\}) - H(\min\{t,t''\})}{H(t') - H(t'')}$. Note that max $supp(L) = t' \leq \xi(t') = B\left(\widehat{H}^p_{t'}\right)$. One can check that $\widehat{H}^p_{t'}(t) = \lambda \widehat{H}^p_{t''}(t) + (1-\lambda)L(t)$ where $\lambda = \frac{p + (1-p)H(t'')}{p + (1-p)H(t')} \in (0,1)$. By Assumption 2, $B\left(\widehat{H}^p_{t'}\right) \leq B\left(\widehat{H}^p_{t''}\right)$. Next note that $\xi'(t) = \left\{F_2\left(1 - \max\{x_0, x^*(t)\}, t\right) + \left[G\left(\max\{x_0, x^*(t)\} - \frac{1}{2}\right) - G\left(x_0 - \frac{1}{2}\right)\right]\right\}\left[F_2\left(1 - x_0, \xi(t)\right)\right]^{-1} > 0$.[30] Thus, $t' > t''$ implies $\xi(t') > \xi(t'')$. Putting these facts together, we have $\xi(t'') < \xi(t') = B\left(\widehat{H}^p_{t'}\right) \leq B\left(\widehat{H}^p_{t''}\right)$, which contradicts the supposition that t'' is a solution.

Step 3B: A solution to expression (5) exists iff $U\left(\max\{x_0, x^*(t^*_0)\}, t^*_0, t^*_0\right) \leq U\left(\frac{1}{2}, B\left(H_{t^*_0}\right), t^*_0\right)$. When it exists, it is unique, and $t_0 \in (0, t^*_0]$.

We define the function $\zeta(t)$ as follows: (1) if $U\left(x_0, 0, t\right) \geq U\left(\frac{1}{2}, B\left(H_t\right), t\right)$, then $\zeta(t) =$

[30]For t such that $x^*(t) \geq x_0$, the envelope theorem allows us to ignore terms involving $dx^*(t)/dt$. Thus, even when $x^*(t) = x_0$, the left and right derivatives are identical.

0; (2) if $U(x_0, 0, t) < U\left(\frac{1}{2}, B(H_t), t\right)$, then $\zeta(t)$ solves $U(x_0, \zeta(t), t) = U\left(\frac{1}{2}, B(H_t), t\right)$. Existence, uniqueness, and continuity of $\zeta(t)$ are easy to verify. Moreover, the equality in (5) is equivalent to the statement that $\zeta(t_0) = B\left(\widehat{H}^p_{t_0}\right)$. In step 3A, we showed that $U\left(x_0, B\left(\widehat{H}^p_t\right), t\right) - U(\max\{x_0, x^*(t)\}, t, t)$ exceeds zero for $t < t_0^*$, is less than zero for $t > t_0^*$, and equals zero at $t = t_0^*$. Consequently, the inequality in (5) holds iff $t_0 \leq t_0^*$. Therefore, (5) is equivalent to the statement that $\zeta(t_0) = B\left(\widehat{H}^p_{t_0}\right)$ for $t_0 \in [0, t_0^*]$.

We can rewrite the equation defining $\zeta(t)$ (when $U(x_0, 0, t) < U\left(\frac{1}{2}, B(H_t), t\right)$) as $F(1 - x_0, \zeta(t)) = t\left(G(0) - G\left(x_0 - \frac{1}{2}\right)\right) + F\left(\frac{1}{2}, B(H_t)\right)$. The right-hand side of this expression is strictly increasing in t, and the left-hand side is strictly increasing in ζ. Consequently, there exists $\widehat{t} \in [0, \overline{t}]$ such that $\zeta(t) = 0$ for $t \in [0, \widehat{t})$, and $\zeta(t)$ is strictly increasing in t for $t \geq \widehat{t}$.

Next note that $B\left(\widehat{H}^p_t\right)$ is weakly decreasing in t for $t \in [0, t_0^*]$. Consider any two values, t', $t'' \leq t_0^*$, with $t' > t''$. By the argument in step 3A, $t' \leq \xi(t') \leq B\left(\widehat{H}^p_{t'}\right)$. Defining $L(t)$ and λ exactly as in step 3A, we have $B\left(\widehat{H}^p_{t'}\right) \leq B\left(\widehat{H}^p_{t''}\right)$ by Assumption 2.

Now suppose $U(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*) \leq U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$. In that case, $U\left(x_0, B\left(\widehat{H}^p_{t_0^*}\right), t_0^*\right) = U(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*) \leq U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$, so $\zeta(t_0^*) \geq B\left(\widehat{H}^p_{t_0^*}\right) > 0$, which also implies $\widehat{t} < t_0^*$. Plainly, $\zeta(t) = 0 < B\left(\widehat{H}^p_t\right)$ for all $t < \widehat{t}$, so any solutions to (5) must lie in $[\widehat{t}, t_0^*]$. Because $\zeta(\widehat{t}) = 0 \leq B\left(\widehat{H}^p_{\widehat{t}}\right)$ and $\zeta(t_0^*) \geq B\left(\widehat{H}^p_{t_0^*}\right)$, continuity guarantees that a solution exists. Since $\zeta(t)$ is strictly increasing and $B\left(\widehat{H}^p_t\right)$ weakly decreasing in t on $[\widehat{t}, t_0^*]$, the solution is unique. Because $\zeta(0) = 0 < B(H)$, we can rule out $t_0 = \widehat{t} = 0$.

Finally suppose $U(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*) > U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$. In that case, $U\left(x_0, B\left(\widehat{H}^p_{t_0^*}\right), t_0^*\right) = U(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*) > U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$, so $\zeta(t_0^*) < B\left(\widehat{H}^p_{t_0^*}\right)$. Given the monotonicity of ζ and B, $\zeta(t) < B\left(\widehat{H}^p_t\right)$ for all $t < t_0^*$. Hence there exists no t_0 satisfying (5).

Step 3C: If $U(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*) \leq U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$, there is at most one equilibrium action function in $\mathcal{Q}_2$, and it must be a double-pool action function.

In a blended double-pool equilibrium, $U(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*) = U(S^{t_0^*}(t_0^*), t_0^*, t_0^*) > U(S^{t_0^*}(t_1), t_1, t_0^*) > U\left(\frac{1}{2}, B(H_{t_1}), t_0^*\right) > U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$ (where the first inequality follows from Lemma 3(c), the second from $t_0^* < t_1$, $S^{t_0^*}(t_1) < \frac{1}{2}$, and (3), and the third from $t_0^* < t_1$), contradicting the supposition. Now consider blended single-pool equilibria. Let x_m solve $\max_x U\left(x, \bar{t}, t_0^*\right)$. It is easy to check that $x_m \leq x^*(\bar{t})$. Note that $U\left(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*\right) = U(S^{t_0^*}(t_0^*), t_0^*, t_0^*) > U(S^{t_0^*}(\bar{t}), \bar{t}, t_0^*) > U\left(\frac{1}{2}, \bar{t}, t_0^*\right) > U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$ (where the first inequality follows from Lemma 3(c), the second from $x_m \leq x^*(\bar{t}) < S^{t_0^*}(\bar{t}) \leq \frac{1}{2}$, and the third from $\bar{t} > B(H_{t_0^*})$), contradicting the supposition. Finally, since the solution for (5) is unique (step 3B), there can be at most one double-pool equilibrium action function.

Step 3D: If $U\left(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*\right) > U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$, there is at most one equilibrium action function in $\mathcal{Q}_2$. If $S^{t_0^*}(\bar{t}) > \frac{1}{2}$, it must be a blended double-pool action function. If $S^{t_0^*}(\bar{t}) \leq \frac{1}{2}$, it must be a blended single-pool action function.

By step 3B, (5) has no solution, so double-pool equilibria do not exist. From step 3A, the value of t_0^* is uniquely determined. Analytically, ruling out blended single-pool equilibria (blended double-pool equilibria) when $S^{t_0^*}(\bar{t}) > \frac{1}{2}$ ($S^{t_0^*}(\bar{t}) \leq \frac{1}{2}$) in the extended dictator game is analogous to ruling out efficient separating equilibria (blended equilibria) when $S_{0,0}(\bar{t}) > \frac{1}{2}$ ($S_{0,0}(\bar{t}) \leq \frac{1}{2}$) in the standard dictator game; we omit the details to conserve space.

Step 3E: If (6) is satisfied, there exists an equilibrium action function $Q^E \in \mathcal{Q}_2$ and an inference function P^E such that (Q^E, P^E) satisifes the D1 criterion.

If $U\left(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*\right) \leq U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$, let Q^E be the double-pool action function for which the highest type in the pool at x_0 is the t_0 that solves (5); if $U\left(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*\right) > U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$ and $S^{t_0^*}(\bar{t}) \leq \frac{1}{2}$, let Q^E be a blended single-pool action function for which the highest type in the pool at x_0 is t_0^*; if $U\left(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*\right) > U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$ and $S^{t_0^*}(\bar{t}) > \frac{1}{2}$, let Q^E be a blended double-pool action function for which the highest type in the pool at x_0 is t_0^*, and the highest type in the separating region is the t_1 that solves (3). In each case, one can verify that, for all $t \in [0, \bar{t}]$, $Q^E(t)$ is type t's best choice within $Q^E([0, \bar{t}])$. For $x \in [0, x_0)$, it is easily shown (in each case) that $0 \in M_x$ and, given (6), every type t prefers its equilibrium outcome to $(x, 0)$; therefore, let P_x^E place all probability on $t = 0$. For any unchosen $x > x_0$, let x_L be the greatest chosen action less than x, t_x^H be the greatest type choosing x_L, and t_x^L be the infinum of types choosing x_L. For any unchosen $x > x_0$

with $Q^E(\bar{t}) > x$, one can show (in each case) that $t_x^H \in M_x$ and every type t prefers its equilibrium outcome to (x, t_x^H); therefore, let P_x place all probability on t_x^H. For any unchosen $x > x_0$ with $Q^E(\bar{t}) < x$, one can show (in each case) that $M_x \cap [0, t_x^L]$ is nonempty and every type t prefers its equilibrium outcome to (x, t_x^L); therefore, let P_x be any distribution over $M_x \cap [0, t_x^L]$. Then in each case (Q^E, P^E) is an equilibrium and satisfies the D1 criterion.

Step 3F: If (6) is satisfied, for any equilibrium (Q, P) satisfying the D1 criterion, Q and Q^E (the unique equilibrium action function within $\mathcal{Q}_2$) coincide on a set of full measure.

One can verify that any equilibrium satisfying the D1 criterion must have the following properties: (1) no type chooses $x > \frac{1}{2}$ (any type choosing $x > \frac{1}{2}$ would deviate to a slightly lower transfer in light of Lemma 2 and the inferences implied by the D1 criterion), (2) choices are weakly monotonic in type (follows from property (1) and Lemma 1), (3) there is no pool at any action other than x_0 and $\frac{1}{2}$ (the proof is similar to that of step 1D).

First we claim that $Q(t) \geq x_0$ $\forall t$. We will prove that $Q(0) \geq x_0$; the claim then follows from property (2). If $Q(0) < x_0$, then $B(P_{Q(0)}) = 0$ (property (3)), so $U(Q(0), B(P_{Q(0)}), 0) \leq U(0, 0, 0)$. Using property (2) and part (3) of Assumption 1, one can show that $B(P_{x_0}) \geq \min_{t \in [0, \bar{t}]} B(\widehat{H}_t^p)$, so by (6), $U(x_0, B(P_{x_0}), 0) > U(Q(0), B(P_{Q(0)}), 0)$, a contradiction.

Next we claim that $Q(t) = x_0$ for some $t > 0$. If not, then by property (2) we have $B(H) = B(P_{x_0})$, and for sufficiently small $t > 0$, $B(P_{Q(t)}) \leq B(H)$. Given that $Q(t) > x_0 > x^*(t)$ for small t, such t would prefer $(x_0, B(P_{x_0}))$ to $(Q(t), B(P_{Q(t)}))$, a contradiction.

Next we claim that $Q(t) > x_0$ for some $t < \bar{t}$. If not, then $B(P_{x_0}) = B(H)$ and (applying the D1 criterion) $B(P_x) = \bar{t}$ for x slightly greater than x_0, so all types could beneficially deviate to that x, a contradiction.

Property (2) and the last three claims imply that $\exists$ $t_0 \in (0, \bar{t})$ such that $Q(t) = x_0$ for $t \in [0, t_0)$, and $Q(t) > x_0$ for $t \in (t_0, \bar{t}]$. Now we claim that for all $t > t_0$, $Q(t) \in \{S^{t_0}(t), \frac{1}{2}\}$. The claim is obviously true if $Q(t) = \frac{1}{2}$ for all $t \in (t_0, \bar{t}]$. By properties (1) and (2), there is only one other possibility: $\exists$ $t_1 \in (t_0, \bar{t}]$ such that $Q(t) \in (x_0, \frac{1}{2})$ for $t \in (t_0, t_1)$, and $Q(t) = \frac{1}{2}$ for $t \in (t_1, \bar{t}]$. Arguing as in step 1F, we see that $\exists$ $z \geq x_0$ such that $Q(t) = S_{t_0,z}(t)$ for

$t \in (t_0, t_1)$. We must have $z \geq x^*(t_0)$: if not, then by equation (1), $Q'(t) = S'_{t_0,z}(t) < 0$ for t close to t_0, contrary to property (2). Thus, $z \geq \max\{x_0, x^*(t_0)\}$. Next we rule out $z > \max\{x_0, x^*(t_0)\}$: in that case, for sufficiently small $\varepsilon > 0$, $\max\{x_0, x^*(t_0)\} + \varepsilon$ is not chosen by any type, and it can be shown that the D1 criterion implies $B(P_{\max\{x_0,x^*(t_0)\}+\varepsilon}) \geq t_0$, so for small $\eta > 0$ type $t_0 + \eta$ strictly prefers $(\max\{x_0, x^*(t_0)\} + \varepsilon,\, t_0)$ to $(Q(t_0 + \eta),\, t_0 + \eta)$, a contradiction. Thus, $z = \max\{x_0, x^*(t_0)\}$, which establishes the claim.

Thus, Q must fall into one of three categories: (a) $\exists\, t_0 \in (0, \bar{t})$ such that $Q(t) = x_0$ for $t \in [0, t_0)$ and $Q(t) = \frac{1}{2}$ for $t \in (t_0, \bar{t}]$; (b) $\exists\, t_0 \in (0, \bar{t})$ with $S^{t_0}(\bar{t}) \leq \frac{1}{2}$ such that $Q(t) = x_0$ for $t \in [0, t_0)$ and $Q(t) = S^{t_0}(t)$ for $t \in (t_0, \bar{t}]$; or (c) $\exists\, t_0 \in (0, \bar{t})$ and $t_1 \in (t_0, \bar{t})$ with $S^{t_0}(t_1) \leq \frac{1}{2}$ such that $Q(t) = x_0$ for $t \in [0, t_0)$, $Q(t) = S^{t_0}(t)$ for $t \in (t_0, t_1)$, and $Q(t) = \frac{1}{2}$ for $t \in (t_0, \bar{t}]$. If Q falls into categories (a) or (b), let $Q^*(t) = Q(t)$ for $t \neq t_0$, and $Q^*(t_0) = x_0$; if Q falls into category (c), let $Q^*(t) = Q(t)$ for $t \notin \{t_0, t_1\}$, $Q^*(t_0) = x_0$, and $Q^*(t_1) = S^{t_0}(t_1)$. In each case, one can show that, because Q is an equilibrium action function, so is Q^*; also, $Q^* \in \mathcal{Q}_2$, and Q and Q^* coincide on a set of full measure. □

Proof of Theorem 4: To reflect the dependence of t_0^* (defined in step 3A) on p, we will use the notation $t_0^*(p)$. Let $\widehat{t_0}(p)$ equal $t_0^*(p)$ when either a blended single-pool or double-pool equilibrium exists, and the solution to the equality in (5) when a double-pool equilibrium exists. Let $t_1^*(p)$ equal the solution to (3) when a blended double-pool equilibrium exists, $\bar{t}$ when a blended single-pool equilibrium exists, and $\widehat{t_0}(p)$ when a double-pool equilibrium exists. Regardless of which type of equilibrium prevails, types $t \in [0, \widehat{t_0}(p)]$ choose $x = 0$ and types $t \in (t_1^*(p), \bar{t}]$ choose $x = \frac{1}{2}$. We demonstrate that $\widehat{t_0}(p)$ is strictly increasing in p, and $t_1^*(p)$ is increasing in p (strictly when $t_1^*(p) < \bar{t}$), which establishes the theorem.

Step 4A: $\widehat{t_0}(p)$ and $t_1^*(p)$ are continuous in p. Continuity of $t_0^*(p)$ follows from uniqueness and continuity of the functions in (2). For similar reasons, when a solution to the equality in (5) exists, it is continuous in p. Finally, it is easy to check that when $U\left(\max\{x_0, x^*(t_0^*(p))\}, t_0^*(p), t_0^*(p)\right) = U\left(\frac{1}{2}, B\left(H_{t_0^*(p)}\right), t_0^*(p)\right)$, the solutions to (2) and the equality in (5) coincide. Thus, $\widehat{t_0}(p)$ is continuous. Continuity of the solution to (3) (when it exists) follows from the following observations: (1) $t_0^*(p)$ is continuous in p, (2) $S^{t_0}(t)$ is continuous for all t_0 and t, and (3) the solution to (3), when it exists, is unique. When $U\left(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*\right) = U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$, we must have $S^{t_0^*(p)}(\bar{t}) > \frac{1}{2}$ (otherwise type $t_0^*(p)$ would gain by deviating to $Q(\bar{t}) = S^{t_0^*(p)}(\bar{t})$), and it is easy to check that $t_0^*(p)$ satisfies both (3) and (5). When $U\left(\max\{x_0, x^*(t_0^*)\}, t_0^*, t_0^*\right) > U\left(\frac{1}{2}, B\left(H_{t_0^*}\right), t_0^*\right)$ and $S^{t_0^*(p)}(\bar{t}) = \frac{1}{2}$, then $\bar{t}$ solves (3). Thus, $t_1^*(p)$ is continuous.

Step 4B: $t_0^*(p)$ is strictly increasing in p. From step 3A, $t_0^*(p)$ satisfies $\xi(t_0^*(p)) = B\left(\widehat{H}^p_{t_0^*(p)}\right)$. Consider p' and $p'' < p'$. One can verify that $\widehat{H}^{p''}_\tau(t) = \lambda \widehat{H}^{p'}_\tau(t) + (1-\lambda)L(t)$ where $\lambda = \left(\frac{p''}{p'}\right)\left(\frac{p'+(1-p')H(\tau)}{p''+(1-p'')H(\tau)}\right) \in (0,1)$ and $L(t) = \frac{H(\min\{\tau,t\})}{H(\tau)}$. For $\tau \leq t_0^*(p'')$, max $supp(L) = \tau \leq \xi(\tau) \leq B\left(\widehat{H}^{p''}_\tau\right)$. Since the support of L is nondegenerate, Assumption 2 implies $B\left(\widehat{H}^{p'}_\tau\right) > B\left(\widehat{H}^{p''}_\tau\right)$ for $\tau \leq t_0^*(p'')$. Note that $\xi(t)$, defined in step 3A, is independent of p. Therefore, $B\left(\widehat{H}^{p'}_\tau\right) > \xi(\tau)$ for $\tau \leq t_0^*(p'')$, so $t_0^*(p') > t_0^*(p'')$, as claimed.

Step 4C: If a double-pool equilibrium exists for p' and $p'' < p'$, then $\widehat{t}_0(p') > \widehat{t}_0(p'')$. Recall from step 3B that, in such cases, $\widehat{t}_0(p)$ satisfies $\zeta(\widehat{t}_0(p)) = B\left(\widehat{H}^p_{\widehat{t}_0(p)}\right)$, and that $\widehat{t}_0(p) \leq t_0^*(p)$. We have shown that $B\left(\widehat{H}^p_\tau\right)$ is weakly decreasing in τ for $\tau \leq t_0^*(p)$ (step 3B), and that $B\left(\widehat{H}^{p'}_\tau\right) > B\left(\widehat{H}^{p''}_\tau\right)$ for $\tau \leq t_0^*(p'')$ (step 4B). Note that $\zeta(t)$, as defined in step 3B, is independent of p. From these observations, it follows that $B\left(\widehat{H}^{p'}_\tau\right) > \zeta(\tau)$ for $\tau \leq \widehat{t}_0(p'')$, so $\widehat{t}_0(p') > \widehat{t}_0(p'')$, as desired.

Step 4D: If a blended double-pool equilibrium exists for p' and $p'' < p'$, then $t_1^*(p') > t_1^*(p'')$. We know $t_0^*(p') > t_0^*(p'')$. Since $S^{t_0^*(p'')}\left(t_0^*(p')\right) > \max\{x_0, x^*\left(t_0^*(p')\right)\} = S^{t_0^*(p')}\left(t_0^*(p')\right)$, we know $S^{t_0^*(p')}(t) < S^{t_0^*(p'')}(t)$ for all $t > t_0^*(p')$ (Lemma 3(e)). Analogously to step 1A, define $\psi^p(t)$ as the solution to $U\left(S^{t_0^*(p)}(t), t, t\right) = U\left(\frac{1}{2}, \psi^p(t), t\right)$. We can rewrite the solution for $t_1^*(p)$ (when a blended double-pool equilibrium exists) as $\psi^p(t) = B(H_t)$. Arguing as in step 1A, one can show that $\psi^p(t)$ is decreasing and continuous in t, while $B(H_t)$ is increasing and continuous in t (and independent of p). Moreover, since $S^{t_0^*(p')}(t) < S^{t_0^*(p'')}(t)$, we have $U\left(S^{t_0^*(p')}(t), t, t\right) < U\left(S^{t_0^*(p'')}(t), t, t\right)$, which means $\psi^{p'}(t) > \psi^{p''}(t)$. Thus, the value of t satisfying $\psi^p(t) = B(H_t)$ is larger for $p = p'$ than for $p = p''$. $\square$

References

[1] Agrawal, Pradeep. 2002. "Incentives, Risk, and Agency Costs in the Choice of Contractual Arrangements in Agriculture." *Review of Development Economics* 6, 460-77.

[2] Andreoni, J. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy* 97, 1447-1458.

[3] Andreoni, J. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving ." *Economic Journal* 100, 464-477.

[4] Andreoni, J. and B.D. Bernheim. 2007. "Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects." Working Paper.

[5] Andreoni, J., P. Brown, and L. Vesterlund. 2002. "What Makes an Allocation Fair? Some Experimental Evidence." *Games and Economic Behavior* 40, 1-24.

[6] Andreoni, J. and J.H. Miller. 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70, 737-753.

[7] Andreoni, J. and R. Petrie. 2004. "Public Goods Experiments Without Confidentiality: A Glimpse into Fund-raising." *Journal of Public Economics* 88, 1605-1623.

[8] Bagwell, L. S. and B. D. Bernheim. 1996. "Veblen Effects in a Theory of Conspicuous Consumption." *American Economic Review* 86(3), 349-373.

[9] Bernheim, B. D. 1994. "A Theory of Conformity." *Journal of Political Economy* 102(5), 841-877.

[10] Bernheim, B. D. and S. Severinov. 2003. "Bequests as Signals: An Explanation for the Equal Division Puzzle." *Journal of Political Economy* 111, 733-764.

[11] Bloom, D.E. 1986. "Empirical Models of Arbitrator Behavior under Conventional Arbitration." *Review of Economics and Statistics* 68, 578-85.

[12] Blount, S. 1995. “When Social Outcomes Aren’t Fair: The Effect of Causal Attributions on Preferences.” *Organ. Behav. Human Dec. Proc.* 63, 131-144.

[13] Bohnet, I. and B. S. Frey. 1999. “The Sound of Silence in Prisoner’s Dilemma and Dictator Games.” *Journal of Economic Behavior & Organization* 38, 43-57.

[14] Bolton, G. E. and A. Ockenfels. 2000. “ERC: A Theory of Equity, Reciprocity, and Competition.” *American Economic Review* 90(1), 166-193.

[15] Brandts, J. and C.A. Holt. 1992. “An Experimental Test of Equilibrium Dominance in Signaling Games.” *American Economic Review* 82(5), 1350-1365.

[16] Brandts, J. and C.A. Holt. 1993. “Adjustment patterns and equilibrium selection in experimental signaling games.” *International Journal of Game Theory* 22(3), 279-302.

[17] Brandts, J. and C.A. Holt. 1995. “Limitations of Dominance and Forward Induction: Experimental Evidence.” *Economics Letters* 49 (4), 391-95.

[18] Broberg, T., T. Ellingsen, and M. Johannesson. 2007. “Is Generosity Involuntary?” *Economic Letters* 94, 32-37.

[19] Camerer, C. F. 1997. “Progress in Behavioral Game Theory.” *Journal of Economic Perspectives* 11, 167-188.

[20] Camerer, C. F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton University Press, Princeton.

[21] Cherry, T.L., P. Frykblom, and J.F. Shogren. 2002. “Hardnose the Dictator.” *American Economic Review* 92, 1218-1221.

[22] Cho, I.-K. and D. Kreps. 1987. “Signaling Games and Stable Equilibria.” *Quarterly Journal of Economics* 102, 179-221.

[23] Cho, I.-K. and J. Sobel. 1990. “Strategic Stability and Uniqueness in Signaling Games.” *Journal of Economic Theory* 50, 381-413.

[24] Dana, J., D. M. Cain, and R. M. Dawes. 2006. “What You Don’t Know Won’t Hurt Me: Costly (But Quiet) Exit in Dictator Games.” *Organizational Behavior and Human Decision Processes* 100, 193-201.

[25] Dana, J., R.A. Weber, and X. Kuang. 2007. “Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness.” *Economic Theory* 33, 67-80

[26] Dasgupta, S. and Z. Tao. 1998. “Contractual Incompleteness and the Optimality of Equity Joint Ventures.” *Journal of Economic Behavior and Organization* 37, 391-413.

[27] De Weaver, M. and J. Roumasset. 2002. “Risk Aversion as Effort Incentive: A Correction and Prima Facie Test of the Moral Hazard Theory of Share Tenancy.” *Economics Bulletin* 15, 1-16.

[28] Ellingsen, T., and M. Johannesson. 2005. “Sunk Costs and Fairness in Incomplete Information Bargaining,” *Games and Economic Behavior* 50, 155-177.

[29] Ellingsen, T. and M. Johannesson. 2006. “Pride and Prejudice: The Human Side of Incentive Theory.” *American Economic Review* 98 (3), 990-1008, 2008.

[30] Falk, A., E. Fehr, and U. Fischbacher. 2008. “Testing Theories of Fairness—Intentions Matter.” *Games and Economic Behavior* 62, 287-303.

[31] Fehr, E. and K. Schmidt. 1999. “A Theory of Fairness, Competition, and Cooperation.” *Quarterly Journal of Economics* 114, 817-68.

[32] Forsythe, R., J. Horowitz, N. Savin, and M. Sefton. 1994. “The Statistical Analysis of Experiments with Simple Bargaining Games.” *Games and Economic Behavior* 6, 347-69.

[33] Geanakoplos, J., D. Pearce, and E. Stacchetti. 1989. "Psychological Games and Sequential Rationality." *Games and Economic Behavior* 1, 60-79.

[34] Glazer, Amihai, and Kai A. Konrad. 1996. "A Signaling Explanation for Charity." *American Economic Review* 86(4), 1019-28.

[35] Güth, Werner, S. Huck, and P. Ockenfels. 1996. "Two-Level Ultimatum Bargaining with Incomplete Information: An Experimental Study" *The Economic Journal* 106, 593-604.

[36] Harbaugh, W. 1998. "The Prestige Motive for Making Charitable Transfers." *American Economic Review* 88, 277– 282.

[37] Hauswald, R. and U. Hege. 2003. "Ownership and Control in Joint Ventures: Theory and Evidence." C.E.P.R. Discussion Paper 4056.

[38] Hoffman, E., K. McCabe, and V. Smith. 1996. "Social Distance and Other-Regarding Behavior in Dictator Games." *American Economic Review* 86, 653-60.

[39] Ireland, Norman. 1994. "On Limiting the Market for Status Signals." *Journal of Public Economics* 53(1), 91-110.

[40] Kagel, J.H., C. Kim, and D. Moser."Fairness in Ultimatum Games with Asymmetric Information and Asymmetric Payoffs." *Games and Economic Behavior* 13, 100-110.

[41] Koch, A.K., and H.T. Normann. 2008. "Giving in dictator games: Regard for others or regard by others?" *Southern Economic Journal*, forthcoming.

[42] List, John A. 2007. "On the Interpretation of Giving in Dictator Games." *Journal of Political Economy* 115(3), 482-493.

[43] Mailath, George J. 1987. "Incentive Compatibility in Signaling Games with a Continuum of Types." *Econometrica* 55(6), 1349-1365.

[44] Manning, A. 2007. "Respect." Mimeo, London School of Economics.

[45] Menchik, P.L. 1980. "Primogeniture, Equal Sharing„ and the U.S. Distribution of Wealth." *Quarterly Journal of Economics* 94, 299-316.

[46] Menchik, P.L. 1988. "Unequal Estate Division: Is It Altruism, Reverse Bequests, or Simply Noise?" In *Modelling the Accumulation and Distribution of Wealth.* D. Kessler and A. Masson (eds.). New York: Oxford Univ. Press.

[47] Mitzkewitz, M., and R. Nagel. 1993. "Experimental Results on Ultimatum Games with Incomplete Information." *International Journal of Game Theory* 22, 171-198.

[48] Oberholzer-Gee, F, and R. Eichenberger. 2008. "Fairness in Extended Dictator Game Experiments," unpublished manuscript.

[49] Rege, M. and K. Telle. 2004. "The Impact of Social Approval and Framing on Cooperation in Public Good Situations" *Journal of Public Economics* 88, 1625-1644.

[50] Rotemberg, Julio. 2007. "Minimally Acceptable Altruism and the Ultimatum Game." *Journal of Economic Behavior and Organization.* Forthcoming.

[51] Soetevent, A.R. 2005. "Anonymity in Giving in a Natural Context—A Field Experiment in 30 Churches." *Journal of Public Economics,* 89, 2301-2323.

[52] Tadelis, S. 2007. "The Power of Shame and the Rationality of Trust." Mimeo, Haas School of Business, U. C. Berkeley.

[53] Veugelers, R. and K. Kesteloot. 1996. "Bargained Shares in Joint Ventures among Asymmetric Partners: Is the Matthew Effect Catalyzing?" *Journal of Economics* 64, 23-51.

[54] Wilhelm, M.O. 1996. "Bequest Behavior and the Effect of Heirs' Earnings: Testing the Altruistic Model of Bequests." *American Economic Review* 86, 874-92.