Social Identity and Cooperation

by

Roy W. Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2012

Doctoral Committee:

Professor Yan Chen, Chair
Professor Tilman M. Borgers
Assistant Professor Erin L. Krupka
Assistant Professor Yusuf Can Masatlioglu

This dissertation is dedicated to my parents Jim and Miranda, and my sister Anne.

ACKNOWLEDGEMENTS

I would like to thank my coauthors for this dissertation, Yan Chen, Qiaozhu Mei and Yang Liu. I would also like to thank my other coauthors Jacob Goeree, Angelo Polydoro, Jan Boone and Suzy Salib.

I would like to thank Yan Chen, Daniel Ackerberg, Tilman Börgers, Colin Camerer, David Cooper, Dan Friedman, Jacob Goeree, Jeremy Fox, Benedikt Hermann, Nancy Kotzian, Erin Krupka, Stephen Leider, Sherry Xin Li, Yusufcan Masatlioglu, Rosemarie Nagel, Neslihan Uler, Roberto Weber, Daniel Zizzo, members of the BEE & ICD lab group, and seminar participants at CERGE-EI, the University of Michigan, Virginia Commonwealth, Simon Fraser University, the National University of Singapore, the 2008 International Meetings of the Economic Science Association (Pasadena, CA), the 2011 North American Meetings of the Economic Science Association (Tucson, AZ), the Third Maastricht Behavioral and Experimental Economics Symposium, and the 2010 Econometric Society World Congress (Shanghai, China) for helpful discussions and comments, and Ashlee Stratakis, Tyler Fisher and Benjamin Spulber for excellent research assistance. The financial support from the National Science Foundation through grant no. SES-0720943 and from Rackham Graduate School through the Rackham Graduate Student Research Grant is gratefully acknowledged.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

<u>**Table**</u>

LIST OF APPENDICES

Appendix

ABSTRACT

Social Identity and Cooperation

by

Roy W. Chen

Chair: Yan Chen

This dissertation examines the effects of social identity in an economic context.

The first chapter studies when a common group identity improves efficiency in coordination games. To address this, we propose a group-contingent social preference model and derive conditions under which social identity changes equilibrium selection. We test our predictions in the minimum-effort game in the laboratory under parameter configurations which lead to an inefficient low-effort equilibrium for subjects with no group identity. For those with a salient group identity, consistent with our theory, we find that learning leads to ingroup coordination to the efficient high-effort equilibrium. Additionally, our theoretical framework reconciles findings from a number of coordination game experiments.

In the second chapter, we investigate the effects of lending team competition on pro-social lending on Kiva.org, the first peer-to-peer microlending site which matches citizen lenders with entrepreneurs in developing countries. We find that lenders are more similar to their team members than to those outside the team. When choosing a team, lenders pay more attention to team rank than to member characteristics. Furthermore, joining a lending team increases lending, but mostly for those who are

not already engaged in the site. Lenders who join teams make 1.8 more loans per month than those who do not, which translates to at least $45 per month. Using a random sample of team forum data, we find that teams who post many links to specific loans increase their lending.

The third chapter studies the endogenous choice of groups, which is an important but underexplored aspect of social identity. While group membership is frequently voluntary in real-world situations, most studies of group effects limit their analyses to either pre-determined or induced groups. In this study, I separately examine how endogenous group choice and intergroup comparison affect pro-social behavior in a laboratory setting. Allowing subjects some choice over their group membership increases pro-social behavior, while intergroup comparison does so only when subjects care about their groups. Finally, I calibrate a learning model incorporating social preferences, yielding estimates of the subjects' group-contingent other-regarding preferences.

CHAPTER I

The Potential of Social Identity for Equilibrium Selection

1.1 Introduction

Today's workplace is comprised of increasingly diverse social categories, including various racial, ethnic, religious and linguistic groups. Within this environment, many organizations face competition among employees in different departments, as well as conflicts between permanent employees and contingent workers (temporary, part-time, seasonal and contracted employees). While a diverse workforce contains a variety of abilities, experiences and cultures which can lead to innovation and creativity, diversity may also be costly and counterproductive if members of work teams find it difficult to integrate their diverse backgrounds and work together (Heap and Zizzo, 2009). This issue of integrating and motivating a diverse workforce is thus an important consideration for organizations. One method to achieve such integration is to develop a common identity. In practice, common identities have often been used to create common goals and values. To create a common identity and to teach individuals to work together towards a common purpose, companies have attempted various creative team-building exercises, such as simulated space missions where the crew works together to overcome malfunctions, perform research and keep life support systems operational while navigating through space (Ball, 1999), and rowing competitions where "each person in the boat is totally reliant on other team members and therefore must learn to trust and respect the unique skills and personalities of the whole team."[1] Given the importance of building a common identity, social identity research offers insight into the potential value of creating a common ingroup identity to override potentially fragmenting identities.

The large body of empirical work on social identity throughout the social sciences has established several robust findings regarding the development of a group identity and its effects. Most fundamentally, the research shows that group identity affects individual behavior. For example, Tajfel et al. (1971) find that group membership creates ingroup enhancement in ways that favor the ingroup at the expense of the outgroup. Additionally, many experiments in social psychology identify factors which enhance or mitigate ingroup favoritism. Furthermore, as a person derives self-esteem from the group she identifies with, salient group identity induces people to conform to stereotypes (Shih et al., 1999).

Since the seminal work of Akerlof and Kranton (2000), there has been increased interest in social identity research in economics, yielding new insights into phenomena which standard economic analysis on individual-level incentives proves unable to explain. Social identity models have been applied to the analyses of gender discrimination, the economics of poverty and social exclusion, the household division of labor (Akerlof and Kranton, 2000), contract theory (Akerlof and Kranton, 2005), economic development (Basu, 2006), and public goods provision (e.g., Croson et al. (2008), Eckel and Grossman (2005)), summarized in Akerlof and Kranton (2010).

In this paper, we systematically induce groups and social preferences in the laboratory, and associate this experimental manipulation with forming group identities. We model social identity as part of an individual's group-contingent social preference.

[1] Olympic rower Bo Hanson (Horswill, Amanda. 2007. "Putting Mateship to Work." *The Courier-Mail*, October. http://www.team8.com.au/PDFs/mateshiptowork_download.pdf).

We are aware of three such extensions of social preference models. First, Basu (2006) uses an altruism model where the weight on the other person's payoff is independent of payoff distributions to derive conditions for cooperation in the prisoner's dilemma game. In comparison, McLeish and Oxoby (2007) and Chen and Li (2009) both incorporate social identity as part of an individual's difference-averse social preference, extending the piece-wise linear models of Fehr and Schmidt (1999) and Charness and Rabin (2002). In this paper, we apply the group-contingent social preference model to the class of potential games with multiple Pareto-ranked equilibria.

This class of games is a challenging domain for economic models of social identity, as "predicting which of the many equilibria will be selected is perhaps the most difficult problem in game theory" (Camerer, 2003). Using a group-contingent social preference model, we derive the conditions under which social identity changes equilibrium selection in the class of potential games with multiple Pareto-ranked equilibria, which includes the minimum-effort games of Van Huyck et al. (1990). We then use laboratory experiments to verify the theoretical predictions. The results show that, under parameter configurations where learning would result in convergence to the inefficient, low-effort equilibrium (Goeree and Holt, 2005), an induced salient group identity can lead to ingroup coordination to the efficient high-effort equilibrium. Furthermore, we show that, at least for the class of potential games, social identity changes equilibrium behavior by changing the potential function.

Our findings contribute to the experimental economics literature, where the fact that social norms, group identity or group competition can lead to a more efficient equilibrium has been demonstrated in the context of the minimum-effort game (e.g., Weber (2006), Gary Bornstein and Nagel (2002)), the provision point mechanism (Croson et al., 2008) and the Battle of the Sexes (Charness et al., 2007). Our theoretical model provides a unifying framework for understanding these experimental results (Appendix F).

The rest of the paper is organized as follows. Section 1.2 reviews the main experimental and theoretical results on minimum-effort games. In Section 1.3, we present the theory of potential games, incorporate social identity into the potential function, and derive theoretical predictions. In Section 1.4, we present our experimental design. Section 1.5 presents our hypotheses. Section 1.6 presents the analysis and results. Section 1.7 concludes.

1.2 The Minimum-Effort Coordination Game

The minimum-effort game is one the most well known coordination games. Rather than exhaustively reviewing the large experimental economics literature on coordination games,[2] we summarize the main findings for the minimum-effort games, leaving a more thorough discussion of the literature on the effects of social identity and group competition on equilibrium selection to Appendix F.

The general form of the payoff function for a player i in an n-person minimum-effort game is as follows:

$$\pi_i(x_1, \ldots, x_n) = a \cdot \min\{x_1, \ldots, x_n\} - c \cdot x_i + b, \tag{1.1}$$

where a, c and b are real, nonnegative constants, and $x_i \geq 0$ is the effort provided by player i. This game has multiple Pareto-ranked pure-strategy Nash equilibria. Specifically, any situation where every player provides the same effort level is a Nash equilibrium, and any equilibrium where the chosen effort is higher Pareto-dominates any equilibrium where the chosen effort is lower.

The most widely-cited paper in coordination games is the experimental test of the minimum-effort game by Van Huyck et al. (1990), frequently shortened to VHBB. They conduct three treatments, all of which use the parameters $a = 0.2$ and $b = 0.6$.

[2] We refer the reader to chapter 7 of Camerer (2003) for an overview of the literature.

In the first treatment, $c = 0.1$ and the number of players in each game, n, ranges from 14 to 16. Subjects can choose any integer effort level from 1 to 7. After 10 rounds of this game, the subjects mostly converge to providing the lowest effort level of 1. In the second treatment, when n is reduced to 2, VHBB find that subjects converge to providing the highest effort level of 7. In a third treatment, n again ranges from 14 to 16, but the cost of providing effort is reduced to zero ($c = 0$). In this case, where offering the highest effort is a weakly dominant strategy for each subject, VHBB find that the subjects again converge to providing the highest effort level. These results suggest that whether group members exert high effort is sensitive to group size (n), the marginal benefit of the public good (a), and the individual marginal cost of effort (c).

Two streams of theoretical work explore the observed equilibria from the order-statistic coordination experiments, with the minimum-effort game as a special case. In the first, Crawford and coauthors use learning dynamics, including evolutionary dynamics (Crawford, 1991) and history-dependent adaptive learning models (Crawford 1995, Crawford and Broseta 1998) to track behavior in the experimental data. In comparison, Monderer and Shapley (1996) note that the minimum-effort game is a potential game,[3] and that the empirical regularities from VHBB are consistent with maximization of the potential function. Intuitively, the potential-maximizing equilibrium has the largest basin of attraction under adaptive learning dynamics. Thus, both streams of theoretical work use learning dynamics to predict which equilibrium will be selected empirically.

While maximization of the standard potential yields a Nash equilibrium, experimental data are often noisy and better explained by statistical equilibrium concepts such as the quantal response equilibrium (McKelvey and Palfrey, 1995). Motivated by this consideration, Anderson et al. (2001) derive the logit equilibrium prediction for

[3]We introduce potential games in Section 1.3.

the minimum-effort game and show that the logit equilibrium maximizes the stochastic potential of the game. To test the theoretical predictions of the logit equilibrium, Goeree and Holt (2005) design a version of the minimum-effort game with a continuous strategy space, where the subjects can choose any real effort level from 110 to 170. They use the parameters $a = 1, b = 0, n = 2$, i.e.,

$$\pi_i(x_i, x_j) = \min\{x_i, x_j\} - c \cdot x_i. \tag{1.2}$$

With these parameter values, the authors show that, consistent with the logit equilibrium prediction, when $c = 0.25$ subjects converge to an effort level close to 170, and when $c = 0.75$ subjects converge to an effort level close to 110. Our experimental design, described in Section 1.4, follows Goeree and Holt's, with the addition of induced group identities to test the effect of group identity on equilibrium selection.

1.3 Potential Games

Both theoretical and experimental studies of coordination games point to the importance of learning dynamics in equilibrium selection. When incorporating dynamic learning models, it is useful to examine the potential function of the game, as described by Monderer and Shapley (1996) and defined below. As Monderer and Shapley note, the minimum-effort game is a potential game, in that it yields a potential function. One interesting property of potential games is that several learning algorithms converge to the argmax set of the potential, including a log-linear strategy revision process (Blume, 1993), myopic learning based on a one-sided better reply dynamic and fictitious play (Monderer and Shapley, 1996). Under these learning dynamics, the potential-maximizing equilibrium has the largest basin of attraction. It is for this reason that we study the potential function of the minimum-effort game.

Monderer and Shapley (1996) formally define *potential games* as games that admit a potential function P such that:

$$\pi_i(x_i, x_{-i}) \geq \pi_i(x_i^{'}, x_{-i}) \quad \Leftrightarrow \quad P(x_i, x_{-i}) \geq P(x_i^{'}, x_{-i}), \quad \forall i, x_i, x_i^{'}, x_{-i}. \tag{1.3}$$

A potential function is a global function defined on the space of pure strategy profiles such that the change in any player's payoffs from a unilateral deviation is exactly matched by the change in the potential P. To determine whether a game has a potential function, Ui (2000) notes that every potential game has a symmetric structure. The Cournot oligopoly game with a linear inverse demand function is a well-known example of a potential game, where each player's payoff depends on a symmetric market aggregate of all players' outputs (the inverse demand function), and also on her own output (the cost of production). Similarly, the minimum-effort game defined by Equation (1.1) has a symmetric interaction term, $a \cdot \min\{x_1, \ldots, x_n\}$, and a term depending only on a player's own strategy, $c \cdot x_i$.

When the payoff functions are twice continuously differentiable, Monderer and Shapley (1996) present a convenient characterization of potential games. That is, a game is a potential game if and only if the cross partial derivatives of the utility functions for any two players are the same, i.e.,

$$\frac{\partial^2 \pi_i(x_i, x_{-i})}{\partial x_i \partial x_j} = \frac{\partial^2 \pi_j(x_j, x_{-j})}{\partial x_i \partial x_j} = \frac{\partial^2 P(x_i, x_{-i})}{\partial x_i \partial x_j}, \quad \forall i, j \in N. \tag{1.4}$$

Equation (1.4) can be used to identify potential games. If (1.4) holds, the potential function P can be calculated by integrating (1.4). Similar conditions hold for nondifferentiable payoff functions by replacing "differentials" with "differences" (Monderer and Shapley, 1996).

As noted by Monderer and Shapley (1996), the minimum-effort game with a payoff function defined by Equation (1.1) is a potential game with the potential function:

$$P(x_1, \ldots, x_n) = a \cdot \min\{x_1, \ldots, x_n\} - c\sum_{i=1}^{n} x_i. \tag{1.5}$$

In most previous experiments using the minimum-effort game, subjects converge or begin to converge towards the equilibrium that maximizes the potential function.[4] Let the threshold marginal cost be $c^* = a/n$. When $c > c^*$, subjects converge to the least efficient equilibrium. Examples of this convergence include the VHBB treatment with parameters $a = 0.2$, $c = 0.1$, and $14 \leq n \leq 16$, and the $c = 0.75$ treatment in Goeree and Holt (2005). When $c < c^*$, subjects converge to the Pareto-dominant equilibrium. Examples of this convergence include the VHBB treatment with $c = 0$, and the $c = 0.25$ treatment in Goeree and Holt (2005).

We next incorporate social identity into players' social preferences to demonstrate how identity can change equilibrium selection by changing the potential function. Let $g \in \{I, O, N\}$ be an indicator variable denoting whether the other players' group membership are ingroup, outgroup or group-neutral.

We use a group-contingent social preference model similar to those of Basu (2006), McLeish and Oxoby (2007) and Chen and Li (2009), where an agent maximizes a weighted sum of her own and others' payoffs, with weighting dependent on the group categories of the other players.[5] In the n-player case, player i's utility function is a convex combination of her own payoff and the average payoff of the other players,[6]

$$u_i(x) = \alpha_i^g \cdot \bar{\pi}_{-i} + (1 - \alpha_i^g) \cdot \pi_i(x) = \min\{x_1, \cdots, x_n\} - c \cdot [\alpha_i^g \cdot \bar{x}_{-i} + (1 - \alpha_i^g) \cdot x_i], \tag{1.6}$$

[4]Exceptions, such as Gary Bornstein and Nagel (2002), use intergroup competition to promote higher effort levels, which is consistent with our theoretical framework (Appendix F).

[5]See Eaton et al. (2011) for an evolutionary explanation for this type of preferences.

[6]Key social preference models include Rabin (1993), Levine (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Falk and Fischbacher (2006), and Cox et al. (2007), etc. See Sobel (2005) for a review of these models. Chen and Li (2009) extend the linear model of Charness and Rabin (2002) to incorporate social identity. We use a linear model here for simplicity.

where $\alpha_i^g \in [-1, 1]$ is player i's group-contingent other-regarding parameter, $\bar{\pi}_{-i} = \sum_{j \neq i} \pi_j(x)/(n-1)$ is the average payoff of the other players, and $\bar{x}_{-i} = \sum_{j \neq i} x_j/(n-1)$ is the average effort of the other players. Based on estimations of α_i^g from Chen and Li (2009), we expect that $\alpha_i^I > \alpha_i^N > \alpha_i^O$.

The transformed game with a utility function defined by Equation (1.6) is a potential game, which admits the following potential function,

$$P(x_1, \cdots, x_n) = \min\{x_1, \cdots, x_n\} - c\sum_{i=1}^{n}(1 - \alpha_i^g)x_i. \qquad (1.7)$$

Note that the Nash equilibria for the transformed game defined by (1.6) remain the same as those in the original minimum-effort game in Goeree and Holt (2005), as long as $c < \frac{1}{1-\alpha_i^g}$, for all i. We now use this formulation to derive a set of comparative statics results, which underscore the effects of group identity on equilibrium selection and form the basis for our experimental design. In what follows, ingroup (outgroup) matching refers to the treatment when only members of the same group (different groups) play the minimum-effort game with each other. We present the propositions in this section and relegate all proofs to Appendix A.

Proposition 1. *Ingroup matching increases the threshold marginal cost, c^*, compared to outgroup or group-neutral matching. Furthermore, a more salient group identity increases c^*.*

Proposition 1 implies that, under parameter configurations where the theory predicts convergence to a low-effort equilibrium when players have no defined group identity, an induced or enhanced group identity can raise the threshold marginal cost level and thus lead to the selection of a high-effort equilibrium. In our experimental design, we use the parameter configurations in Goeree and Holt (2005) where the marginal cost of effort is above the threshold, i.e., $c > c^*(n, \{\alpha_i^N\}_{i=1}^n)$, so that play converges to the low-effort equilibrium, and investigate whether induced group identity can lead to convergence to the high-effort equilibrium.

As experimental data are often noisy and better explained by statistical equilibrium concepts, Anderson et al. (2001) derive the logit equilibrium prediction for the minimum-effort game and show that the predicted average efforts are remarkably close to the data averages in the final periods.

We now derive the logit equilibrium predictions for the transformed minimum-effort game with a group-dependent other-regarding utility function as defined by Equation (1.6). Based on the standard assumption of the logit model that payoffs are subject to unobserved shocks from a double-exponential distribution, player i's probability density is an exponential function of the expected utility, $u_i^e(x)$,

$$f_i(x) = \frac{\exp(\lambda u_i^e(x))}{\int_{\underline{x}}^{\bar{x}} \exp(\lambda u_i^e(s))ds}, \qquad i = 1, \cdots, n, \tag{1.8}$$

where $\lambda > 0$ is the inverse noise parameter and higher values correspond to less noise. As $\lambda \to +\infty$, the probability of choosing an action with the highest expected utility goes to 1. As $\lambda \to 0$, the density function becomes uniform over its support and behavior becomes random.

The logit equilibrium is a probability density over effort levels. As the characterization of the logit equilibrium for the transformed minimum-effort game follows from Anderson et al. (2001), we summarize its properties in the following proposition without presenting the proof.

Proposition 2. *There exists a logit equilibrium for the extended minimum-effort game with social identity. Furthermore, the logit equilibrium is unique and symmetric across players.*

Using symmetry and further assuming $\alpha_i = \alpha$ for all i, we first derive the equilibrium distribution of efforts.

Proposition 3. *The equilibrium effort distribution for the logit equilibrium is characterized by the following first-order differential equation:*

$$f(x) = f(\underline{x}) + \frac{\lambda}{n}[1 - (1 - F(x))^n] - c(1 - \alpha)\lambda F(x). \tag{1.9}$$

Equation (1.9) plays a key role in both our comparative statics results and our data analysis. We compute the logit equilibrium effort distribution in Section 1.4 and Appendix A as a benchmark for the final-rounds analysis in Section 1.6 and Appendix E. Anderson et al. (2001) prove that increases in the marginal cost, c, or the number of players, n, result in lower equilibrium effort in the sense of first-order stochastic dominance. Similarly, using (1.9), we next characterize the effect of group-contingent social preference on equilibrium selection.

Proposition 4. *Increases in the group-contingent social preference parameter, α, result in higher equilibrium effort (in the sense of first-order stochastic dominance).*

If players are more altruistic towards their ingroup members than towards out-group members, i.e., $\alpha^I > \alpha^N > \alpha^O$, Proposition 4 implies that the distribution of effort under ingroup matching first-order stochastically dominates the distribution under group-neutral matching, which, in turn, first-order stochastically dominates the distribution under out-group matching, i.e., $F^I(x) \leq F^N(x) \leq F^O(x)$. Consequently, the average equilibrium effort is the highest with ingroup matching, followed by group-neutral and then outgroup matching.

Lastly, as a limit result, we note that the equilibrium density converges to a point mass as the noise goes to zero, which coincides with the predictions of potential maximization.

Proposition 5. *When the inverse of the noise parameter, λ, goes to infinity, the equilibrium density converges to a point mass at the maximum effort $\bar{x}$ if $c < c^*$, at $(\bar{x} - \underline{x})/n$ if $c = c^*$, and at the minimum effort $\underline{x}$ if $c > c^*$, where $c^* = 1/[n(1 - \alpha)]$.*

Together, Propositions 1, 3, 4 and 5 form the basis for our experimental design and hypotheses, which we present in the next two sections.

1.4 Experimental Design

We design our experiments to determine the effects of group identity on equilibrium selection, to test the comparative statics results from Section 1.3, and to investigate the interactions of group identity and learning. In our experiments, we focus on two-person matches in the minimum-effort game. We now present the economic environments and our experimental procedure.

1.4.1 Economic Environments

To study equilibrium selection, we use the same payoff parameters as those of the two-person treatment in Goeree and Holt (2005). However, since our main interest is to investigate the effects of group identity on equilibrium selection, we induce group identities in the lab before the subjects play the minimum-effort game. Furthermore, we run longer repetitions to study the effects of learning dynamics.

Within our experiments, the payoff function, in tokens, for a subject i matched with another subject j is the following: $\pi_i(x_i, x_j) = \min\{x_i, x_j\} - 0.75 \cdot x_i$, where x_i and x_j denote the effort levels chosen by subjects i and j, respectively; each can be any number from 110 to 170, with a resolution of 0.01. By Equation (1.5), the threshold marginal cost of effort, c^*, is equal to 0.5. Therefore, absent of group identities, we expect subjects to converge close to the lowest effort level, 110, which is confirmed by Goeree and Holt (2005).

With group-contingent social preferences, however, the potential function for this game becomes $P(x_i, x_j) = \min\{x_i, x_j\} - 0.75 \cdot [(1 - \alpha_i^g)x_i + (1 - \alpha_j^g)x_j]$, where α_i^g is the weight that subject i places on her match's payoff. Proposition 5 implies that, in the limit with no noise, this potential function is maximized at the most efficient equilibrium if $\alpha^g > \frac{1}{3}$, and at the least efficient equilibrium if $\alpha^g < \frac{1}{3}$. Proposition 4 implies that, with sufficiently salient group identities, ingroup matching leads to a higher average equilibrium effort than either outgroup matching or control (nongroup) matching.

1.4.2 Experimental Procedure

A key design choice for our experiment is whether to use participants' natural identities, such as race and gender, or to induce their identities in the laboratory. Both approaches have been used in lab settings. However, because of the multi-dimensionality of natural identities which might lead to ambiguous effects in the laboratory, we induce identity, which gives the experimenter greater control over the participant's guiding identity.

Our experiment follows a 2×3 between-subject design. In one dimension, we vary the strength of group identity, with near-minimal and enhanced treatments. Our near-minimal treatment is so named because it implements groups in a way that is nearly minimal. The criteria for *minimal groups* (Tajfel and Turner, 1979b) are as follows:

1. Subjects are randomly assigned to groups.
2. Subjects do not interact.
3. Group membership is anonymous.
4. Subjects' choices do not affect their own payoffs.

Our near-minimal treatments achieve the first three of these four criteria, as subjects are assigned to groups based on the random choice of an envelope with a certain colored card inside, and are not allowed to speak to one another or open their envelopes in public. The fourth criterion cannot be realistically achieved in most economics experiments, including ours, since subjects' monetary payoffs are usually tied to their choices. Since this criterion is not met, we refer to these treatments as *near minimal*.

Our enhanced treatment is designed to increase the salience of group identity by incorporating a group problem-solving stage, where salience refers to the relative importance or prominence of group membership. In our model, salience can be captured by the group-contingent other-regarding parameter, $\alpha_i^I - \alpha_i^N$, i.e., the difference between how altruistic player i feels towards an ingroup match when group identity is induced or primed relative to when it is not induced, such as in the control condition.[7] To implement the enhanced treatment, after being randomly assigned to groups, subjects are asked to solve a problem about a pair of paintings. They can use an online communication program to discuss the problem with other members of their group. This problem-solving stage is designed to enhance group identity.

To minimize experimenter demand effects, we use a between-subject design. For treatment sessions, each subject is in either an ingroup session where she is always matched with a member of her own group, or an outgroup session where she is always matched with a member of the other group. To control for the time between group assignment and the minimum-effort games, we use two different controls, one for the near-minimal treatments, and one for the enhanced treatments.[8] In the former, subjects play the minimum-effort game without being assigned to groups. In the latter, each subject is asked to solve the same painting problem on their own, without the online communication program.

Our experimental process is summarized as follows:

1. Random assignment to groups: Every session has twelve subjects. In the treat-

[7] Alternative formulations of salience, e.g., the difference of group-contingent altruism parameters between ingroup and outgroup members, $\alpha_i^I - \alpha_i^O$, are not quite as general. Our formulation can incorporate situations where one dimension of own group identity is primed without necessarily activating an outgroup, such as in Shih et al. (1999).

[8] Chen and Li (2009) note that group effect induced by categorization deteriorates over time in their experiment. Therefore, it is important to control for the time between categorization and the minimum-effort game in the treatment and the corresponding control.

ment sessions, each subject randomly chooses an envelope which contains either a red or a green index card with a subject ID number on it. The subject is assigned to the Red or the Green group based on this index card; each group has six members. In the control sessions, there is no assignment into different groups. Instead, each subject randomly chooses an envelope which contains a white index card with a subject ID number on it.

2. Problem solving: In the enhanced treatments and their corresponding control sessions, the subjects are asked to solve a problem. First, subjects are given five minutes to review five pairs of paintings, each of which contains one painting by Paul Klee and one painting by Wassily Kandinsky. The subjects are also given a key indicating which of the two artists painted each of the ten paintings.[9] Next, subjects are shown two final paintings and are told that each of them was painted by either Klee or Kandinsky, and that they both could have been painted by the same artist. The subjects are then asked to determine, within ten minutes, which artist painted each of these final two paintings.[10] In the treatment sessions, each subject is allowed to use an online communication program to discuss the problem with other members of her own group. A subject is not required to give answers that conform to any decision reached by her group, and she is not required to contribute to the discussion. In comparison, subjects in the corresponding control sessions are given the same amount of time to solve the painting problem on their own, without the online communication option. For each correct answer, a subject earns 350 tokens (the equivalent of $1), though she is not told what the correct responses are until the end of

[9] The five pairs of paintings are: 1A *Gebirgsbildung*, 1924, by Klee; 1B *Subdued Glow*, 1928, by Kandinsky; 2A *Dreamy Improvisation*, 1913, by Kandinsky; 2B *Warning of the Ships*, 1917, by Klee; 3A *Dry-Cool Garden*, 1921, by Klee; 3B *Landscape with Red Splashes I*, 1913, by Kandinsky; 4A *Gentle Ascent*, 1934, by Kandinsky; 4B A *Hoffmannesque Tale*, 1921, by Klee; 5A *Development in Brown*, 1933, by Kandinsky; 5B *The Vase*, 1938, by Klee.

[10] Painting #6 is *Monument in Fertile Country*, 1929, by Klee, and Painting #7 is *Start*, 1928, by Kandinsky.

the experiment, after the minimum-effort game has been played. Note that the near-minimal treatments and the corresponding control sessions do not contain this stage.

3. Minimum-effort game: Each subject plays the minimum-effort game 50 times. For each round, each subject is randomly rematched with one other subject in the same session. In the ingroup treatment sessions, subjects are matched only with members of their own group. In outgroup treatment sessions, subjects are matched only with members of the other group. In the control sessions, there are no groups, so subjects can be matched with any other person in the same session.[11]

4. Survey: At the end of each experimental session, subjects fill out a postexperimental survey which contains questions about demographics, past giving behavior, strategies used during the experiment, group affiliation, and prior knowledge about the artists and paintings.

Past experimental research finds that the extent to which induced identity affects behavior depends on the salience of the social identity. For example, Eckel and Grossman (2005) use induced team identity to study the effects of identity strength on cooperative behavior in a repeated VCM game. They find that "just being identified with a team is, alone, insufficient to overcome self-interest." However, actions designed to enhance team identity, such as group problem solving, contribute to higher levels of team cooperation. Similar findings on the effect of group salience are reported in Charness et al. (2007). Based on previous findings, we expect that group effects will be stronger in our enhanced treatments than our near-minimal treatments.

[11]This matching protocol introduces a potential confound, as subjects interact with 5 other (ingroup), or 6 other (outgroup), or 11 other (control) players. It is possible that interacting with a smaller number of players could increase the weight on one's match. In particular, this could be a reason for the lack of difference in effort levels between the outgroup and control sessions in both treatments. We thank an anonymous referee for pointing this out.

Table 1.1: Features of Experimental Sessions

Treatment		Number of Subjects	Group Assignment	Problem Solving
Near-Minimal	Control	3×12	None	None
	Ingroup	3×12	Random	None
	Outgroup	3×12	Random	None
Enhanced	Control	3×12	None	Self
	Ingroup	3×12	Random	Chat
	Outgroup	3×12	Random	Chat

Table 1.1 summarizes the features of the experimental sessions. In each of the four treatments and two corresponding controls, we run three independent sessions, each with 12 subjects. Overall, 18 independent computerized sessions were conducted in the Robert B. Zajonc Laboratory at the University of Michigan between October 2007 and May 2008, yielding a total of 216 subjects. All sessions were programmed in z-Tree (Fischbacher, 2007). Nearly all of our subjects were drawn from the student body of the University of Michigan.[12] Subjects were allowed to participate in only one session. Each enhanced session lasted approximately one hour, whereas each near-minimal session lasted about forty minutes. The exchange rate was set to 350 tokens for $1. In addition, each participant was paid a $5 show-up fee. Average earnings per participant were $10.82 for those in the near-minimal sessions and $11.69 for those in the enhanced sessions. The experimental instructions are included in Appendix B, while the survey and response statistics are included in Appendix C. Data are available from the authors upon request.

1.5 Hypotheses

In this section, we present our hypotheses regarding subject effort in the minimum-effort game as related to group identity. Our general null hypothesis is that behavior

[12]One subject was from Eastern Michigan University, and one subject was not affiliated with a school.

does not differ between any pair of treatments.

Hypothesis 1 (Effect of Groups on Effort Choices: Ingroup vs. Control)**.** *The average effort level in the ingroup treatment is greater than that in the control sessions:* $\bar{x}^I > \bar{x}^N$.

Hypothesis 2 (Effect of Groups on Effort Choices: Ingroup vs. Outgroup)**.** *The average effort level in the ingroup treatment is greater than that in the outgroup treatment:* $\bar{x}^I > \bar{x}^O$.

Hypothesis 3 (Effect of Groups on Effort Choices: Control vs. Outgroup)**.** *The average effort level in the control sessions is greater than that in the outgroup treatment:* $\bar{x}^N > \bar{x}^O$.

These hypotheses are based on Proposition 4. As α^g increases, the stochastic choice function shifts the probability weight from lower effort to higher effort. Since we expect $\alpha^I > \alpha^N > \alpha^O$, we expect subjects in the ingroup sessions to choose higher effort than those in control sessions, and subjects in the control sessions to choose higher effort than those in the outgroup sessions.

Furthermore, when we enhance the groups, we expect the effect on α^g to be more extreme, so $\alpha^{EI} > \alpha^{MI}$ and $\alpha^{EO} < \alpha^{MO}$, where EI (MI) stands for "enhanced (near-minimal) ingroup" and EO (MO) stands for "enhanced (near-minimal) outgroup." Thus, we obtain the following hypotheses on the effect of identity salience.

Hypothesis 4 (Effect of Identity Salience on Effort Choices: Ingroup)**.** *The average effort level in the enhanced ingroup treatment is greater than that in the near-minimal ingroup treatment:* $\bar{x}^{EI} > \bar{x}^{MI}$.

Hypothesis 5 (Effect of Identity Salience on Effort Choices: Outgroup)**.** *The average effort level in the enhanced outgroup treatment is less than that in the near-minimal outgroup treatment:* $\bar{x}^{EO} < \bar{x}^{MO}$.

We would also like to examine which aspects of the problem-solving stage have an effect on effort. We do this by examining the communication logs from the problem-solving stage. We identify components of these communications and examine how they affect effort. Our belief is that subjects who contribute more to the communication process feel more closely connected to their groups, and therefore have a higher value of α^I and a lower value of α^O.

Hypothesis 6 (Effect of Communication on Effort Choices: Ingroup)**.** *The average effort level in the enhanced ingroup treatment is higher when a subject submits more lines, is more engaged, and gives more analysis during the problem-solving stage.*

Hypothesis 7 (Effect of Communication on Effort Choices: Outgroup)**.** *The average effort level in the enhanced outgroup treatment is lower when a subject submits more lines, is more engaged, and gives more analysis during the problem-solving stage.*

An additional measure of interest in our experiment is efficiency. We define a normalized efficiency measure following the convention in experimental economics:

$$\text{Efficiency} = \frac{\text{Total Payoff - Minimal Payoff}}{\text{Maximal Payoff - Minimal Payoff}},$$

where Total Payoff is the total amount earned by two subjects in a match; Minimal Payoff (10) is the minimum possible total amount that can be earned between two subjects in a match, achieved if one subject chooses an effort of 110, and the other chooses an effort of 170; and Maximal Payoff (85) is the maximum possible total amount that can be earned between two subjects in a match, achieved if both subjects choose an effort of 170. With this definition, efficiency can be any value from 0 to 1, with 0 denoting the case where subjects earn the minimum possible total payoff, and with 1 denoting the case where subjects earn the maximum possible total profit. As theoretical benchmarks, we use the equilibrium distribution described in Equation

(1.9) to compute the expected effort and efficiency for different values of α. These computation results are included in Appendix A.

1.6 Results

In this section, we first present our main results for the effects of group identity on equilibrium selection. We then present our analysis of the interaction of learning and group identity.

Several common features apply throughout our analysis and discussion. First, standard errors in the regressions are clustered at the session level to control for the potential dependency of decisions across individuals within a session. Second, we use a 5 percent statistical significance level as our threshold (unless stated otherwise) to establish the significance of an effect.

1.6.1 Group Identity and Effort

In this experiment, we are interested in whether social identity increases chosen effort. Figure 1.1 presents the median (top row) and minimum (bottom row) efforts in the near-minimal (left column) and enhanced (right column) group treatments.

Our first observation is that the time-series effort levels in the control sessions move towards the lowest effort, with a fairly widespread distribution in round 50. This is consistent with the prediction of the stochastic potential theory and replicates the findings from the two-person, high-cost treatment in Goeree and Holt (2005).[13] However, when group identity is induced, 8 out of 12 sessions show convergence towards the highest effort. In particular, all 3 sessions of the enhanced-ingroup treatment converge towards the highest effort. Group identity also seems to increase the effort level in the near-minimal treatments, but the effects are not as strong. We next use

[13]Using Kolmogorov-Smirnov tests of the equality of distributions for last round choices, we find that the distribution of choices in our control sessions is not significantly different from that in the corresponding treatment in Goeree and Holt (2005) ($p = 0.170$, two-sided).

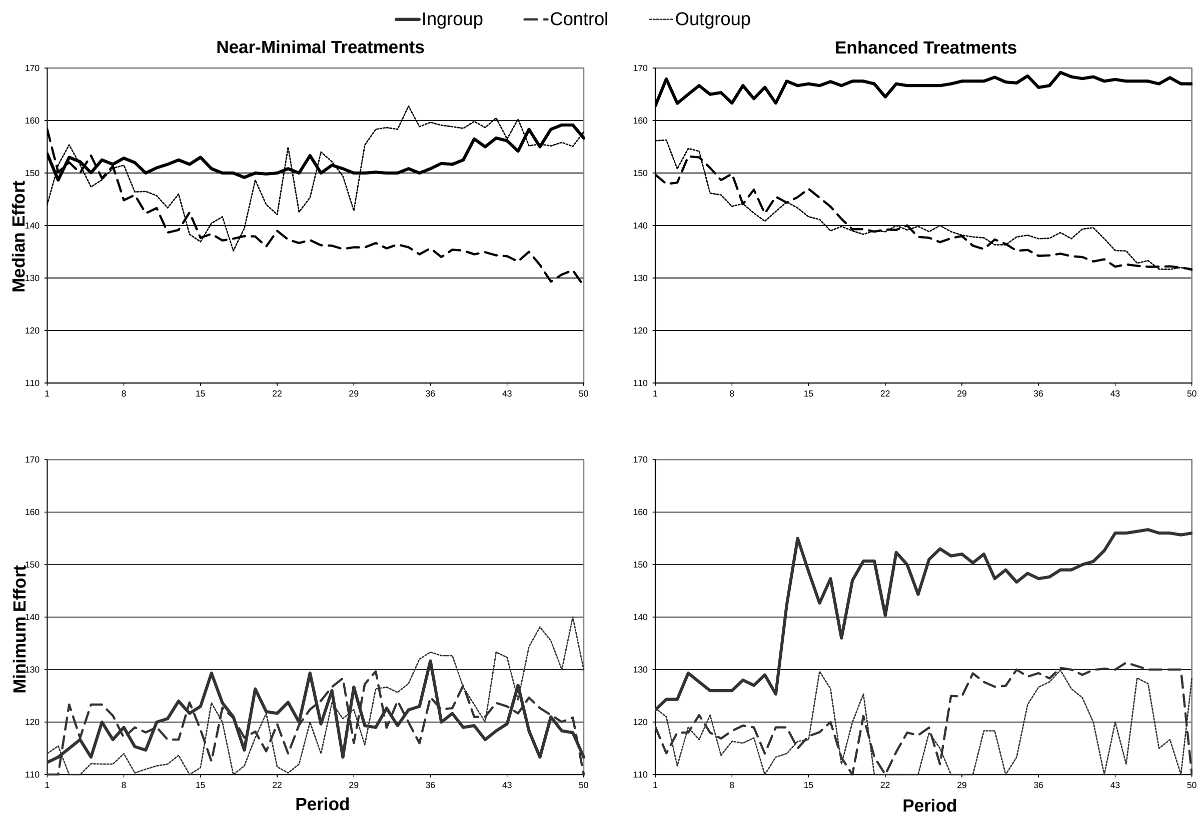

Figure 1.1: Median (Top Row) and Minimum (Bottom Row) Effort in the Near-Minimal (Left Column) and Enhanced (Right Column) Treatments

random-effects regressions to investigate the significance of the observed patterns.

Table 1.2: Group Identity and Effort Choice: Random Effects
(Effort = $\beta_0 + \beta_1$*Ingrp+β_2*Outgrp+β_3*Ingrp*Enh+β_4*Outgrp*Enh+β_5*X+u_{it})

	Dependent Variable: Effort	
	(1)	(2)
Ingroup	8.82	6.22
	(7.15)	(7.04)
Outgroup	10.76	7.80
	(7.67)	(7.65)
Ingroup*Enhanced	15.38***	15.25***
	(4.57)	(4.49)
Outgroup*Enhanced	-10.41	-9.84
	(11.58)	(10.95)
Female		-3.41*
		(1.99)
Asian		2.35
		(2.64)
Black		-2.75
		(3.52)
Hispanic		0.79
		(4.57)
Married		-2.43
		(7.63)
Constant	139.13***	141.28***
	(5.73)	(5.74)
Observations	10800	10200
R^2	0.1691	0.1961

Notes: Standard errors are adjusted for clustering at the session level.
Significant at the: * 10 percent level; *** 1 percent level.

In Table 1.2, we present two random-effects regressions, one with and one without demographic variables included, with clustering at the session level. The dependent variable for these two regressions is the effort level chosen, while the independent variables for all regressions include dummy variables describing whether the subject participated in an ingroup or an outgroup session, with the control as the omitted group. Two other independent variables included in both regressions are the interaction terms between the matching scheme and a dummy variable for whether the session was an enhanced session. These two independent variables allow us to test the effect of group salience on effort level. For these regressions, we treat the two controls in our design (one for the near-minimal and one for the enhanced sessions) as the same group of sessions. The demographic variables include the following dummy variables (with omitted variables in parentheses): age (over 23), gender (male), race (Caucasian), marital status (single), employment status (unemployed), number of siblings (zero siblings), expenses (self), voting history (not a voter), and volunteer status (not a volunteer). The "expenses" variable captures the response to the question of who in the subject's household is responsible for the finances of the household (see Appendix C). In Table 1.2, we omit some demographic variables, but none that are significant. We summarize the results from Table 1.2 below.

Result 1 (Group effect on effort in near-minimal treatments)**.** *In the near-minimal sessions, participants in the different treatments do not choose significantly different effort levels.*

Support. *In Table 1.2, the coefficients for the ingroup dummies ($p = 0.217$ for (1) and $p = 0.377$ for (2)) and for the outgroup dummies ($p = 0.160$ for (1) and $p = 0.308$ for (2)) are not significant. A test of equality of the ingroup and outgroup dummies yields $p = 0.771$ for (1), and $p = 0.821$ for (2).*

Result 1 indicates that, in the near-minimal treatments, subjects in different sessions make roughly the same effort choices throughout the experiment. While the subjects in ingroup sessions provide a slightly higher level of effort than subjects in the control sessions (by 8.82 and 6.22 units of effort in (1) and (2), respectively), this amount is not significant. In fact, subjects in sessions where they are paired with people not in their own group provide an amount of effort that is even higher than that of subjects in the control sessions (10.76 and 7.80 more units, respectively). However, this difference is insignificant. Thus, this result fails to reject the null in favor of Hypotheses 1, 2, and 3 for the near-minimal treatments.

Result 2 (Group effect on effort in enhanced treatments)**.** *In the enhanced sessions, participants in the ingroup sessions choose significantly higher effort levels than those in the control and outgroup sessions, while participants in the control and outgroup sessions do not choose significantly different effort levels.*

Support. *A test that the sum of the coefficients on the ingroup dummy and ingroup-enhanced interaction term is equal to 0 yields* $p < 0.0001$ *for (1) and* $p = 0.0005$ *for (2), while a test that the previous sum is equal to the corresponding outgroup sum yields* $p = 0.023$ *for (1) and* $p = 0.008$ *for (2). A test that this outgroup sum is equal to 0 yields* $p = 0.976$ *for (1) and* $p = 0.846$ *for (2).*

Result 2 indicates that, in the enhanced treatments, subjects in the ingroup sessions provide significantly higher effort than subjects in the other sessions (by 24.20 in (1) and 21.47 units of effort in (2) compared to the control sessions, obtained by summing the coefficients on the ingroup dummy and the ingroup-enhanced interaction term). Subjects in the outgroup sessions provide approximately the same amount of effort compared to subjects in the control sessions (0.35 units more in (1) and 2.04 units fewer in (2)). By Result 2, we reject the null in favor of Hypotheses 1 and 2, but we fail to reject the null in favor of Hypothesis 3 for the enhanced treatments. Both of these results are consistent with those outlined in Brewer's (1999) survey of social psychology experiments relating to social identity. Brewer (1999) notes that ingroup favoritism does not have to be mirrored by outgroup discrimination. Here, we see a significant ingroup favoritism effect with no corresponding outgroup discrimination effect. The lack of outgroup discrimination in our experimental setup is not surprising, as the outgroup and control sessions do not differ except for the categorization of groups, whereas in other environments negative behavior towards outsiders can and do happen (Deaux, 1996).

In several social identity experiments, such as Eckel and Grossman (2005) and Charness et al. (2007), identity salience is crucial in changing behavior. We observe a similar effect in our experiment.

Result 3 (Effect of group salience on effort)**.** *When groups are more salient, participants in the ingroup sessions choose significantly higher effort levels.*

Support. *In Table 1.2, the coefficients on the interaction terms between the ingroup dummy and the enhanced dummy are highly significant ($p = 0.001$ for both (1) and (2)), while the coefficients on the interaction terms between the outgroup dummy and the enhanced dummy are not significant ($p = 0.369$ for both (1) and (2)).*

Result 3 shows that subjects matched with salient ingroup members are more likely to exhibit a high effort than those matched with less-salient ingroup members (by 15.38 and 15.25 units of effort in (1) and (2), respectively). Also, subjects matched with salient outgroup members do not exhibit significantly less effort than subjects matched with less-salient outgroup members (they exhibit 10.41 and 9.84 fewer units of effort in (1) and (2), respectively). Therefore, we reject the null in favor of Hypothesis 4, but we do not reject the null for Hypothesis 5.

Overall, the effect of placing people into groups and then having them solve a problem with each other is to increase their group-contingent other-regarding parameter, α_i^g. In the control sessions, α_i^g is at its base level. In the ingroup sessions, we expect this value to increase; if the increase is great enough, then the potential-maximizing effort choice changes from the minimum effort to the maximum effort. In our experiments, the near-minimal ingroup sessions possibly increase α_i^g, but not enough to change the potential-maximizing effort. In addition, the purpose of the enhanced sessions is to further increase subjects' group-contigent other-regarding parameters. The results show that such a process increases α_i^g enough to also substantially increase the effort level chosen by the participants. In Subsection 1.6.3, we estimate the parameter α_i^g together with other parameters of an adaptive learning model.

We next investigate the factors in the problem-solving stage that affect the amount of effort given in the minimum-effort game. We concentrate on the enhanced sessions, which included 3 control, 3 ingroup, and 3 outgroup sessions. In the control sessions, the subjects guessed the artists by themselves. In the ingroup and outgroup sessions, the subjects were allowed to communicate with other members of their own group via the chat feature in z-Tree.

In order to examine the components of communication, we take the communication logs and code them. Our coding procedures follow the standard practice in content analysis (Krippendorff, 2003). There are 6 sessions with communication, with 2 sets of logs for each session. We have 4 independent coders read through each communication log and identify various aspects of the communication. These coders are asked to examine the communication logs on 3 different levels: the line level, the subject level, and the group level. Details of the coding procedure and instructions can be found in Appendix D.

Among the enhanced sessions, we include these coded variables and other variables in a random-effects regression. First, we examine the inter-rater reliability for each coding category. The interclass correlation (ICC) value for each category is displayed in Table E.1 in Appendix E. As is standard when examining coded communication logs, we drop all variables that do not have an ICC of at least 2/3. This means that we only keep the painting analysis (whether a line shows painting analysis), question (whether a line is a question about the paintings), and (group-level) agreement variables, as well as the subject engagement variable. For these variables, we include in the regression the number of times a subject had a line that was coded in the respective category by 3 out of 4 of the coders.

We also include each subject's line count, painting responses, and demographics in the random-effects regression. The line count is simply the number of times a subject clicks "submit" during the communication process. This variable is a measure of a subject's level of contribution to the communication, since speaking more during this process helps everyone else in the group and costs the speaker a small amount of effort. The painting responses are dummy variables, one for each of paintings 6 and 7, indicating whether that subject correctly identified the paintings' artists. So, a subject received a 1 for the painting 6 (7) dummy variable if that subject submitted the answer "Klee" ("Kandinsky") for painting 6 (7) and a 0 otherwise. While we expect the line count to have some effect on the amount of contributed effort, we do not expect the painting variables to have an effect since the subjects are not told who the actual painters are for paintings 6 and 7 until after the minimum-effort game is played. Finally, we include the same demographics that were included in the original regressions.

Result 4 (Effect of communication on effort)**.** *Subjects give more effort to ingroup members in the minimum-effort game if they ask more questions during the problem-solving stage.*

Support. *In Table 1.3, the coefficient on the Ingroup*Questions variable is significant (p =* 0.020*), while the coefficient on the Outgroup*Questions variable is marginally significant (p =* 0.079*). The coefficients for the other coded variables, the line count, and the painting responses are not significant.*

Table 1.3 shows the results of the regression, not including the demographic variables. Result 4 shows that only the act of asking questions during the communication stage has any significant effect on effort. When a subject asks more questions to members of her own group, she gives more to members of her own group and less to members of the other group. This result refutes both Hypotheses 6 and 7. Also, as predicted, answering the painting problems correctly does not affect the amount of effort given later in the experiment. Even though subjects perform better in the problem-solving task after having communicated with their group members,[14] this

[14]In the enhanced treatment sessions, 83.3 percent of the participants provided correct answers to both paintings, 9.7 percent provided one correct answer, and 6.9 percent provided zero correct answers. In the enhanced control sessions, 66.7 percent of the participants provided correct answers

does not affect the amount of effort they give in the minimum-effort game. Result 4 suggests that an increase in group salience takes place during the communication stage, seemingly through generalized reciprocity (Yamagishi and Kiyonari, 2000). When a subject asks a question regarding the paintings, it is answered by another subject 94 percent of the time. By asking more questions and therefore receiving more help from their group members, subjects seem to feel obligated to give more effort to their group members in the minimum-effort game.

1.6.2 Equilibrium Play and Efficiency

In addition to examining the relation between group identity and effort, we also examine the degree of coordination subjects exhibit in the various treatments. Figure 1.2 shows the frequencies of "wasted" efforts exhibited by each match for the first 10 (left column) and last 10 (right column) periods in each session. The top rows show the near-minimal treatments while the bottom rows show the enhanced treatments. Here, "wasted" effort is defined as the difference in the maximum effort chosen in a match and the minimum effort chosen in that match. Since subjects are paid only the minimum effort chosen in a match, if a subject provides more than the minimum effort, then that subject pays more but receives no extra benefit. This figure shows the degree of coordination that the matches exhibit. In Figure 1.2, matches with no wasted effort indicate subjects are in a Nash equilibrium.

Several results can be observed from this figure. First, for the first 10 periods in the near-minimal treatments, there is not much difference between the control, ingroup, and outgroup sessions in terms of the amount of wasted effort. Furthermore, wasted effort seems to be uniformly distributed among the allowed values. In the enhanced treatments, the first 10 periods show that there is a much higher frequency of little to

to both paintings, 19.4 percent provided one correct answer, and 13.9 percent provided zero correct answers. The average number of correct answers is significantly higher in the enhanced treatment than in the enhanced control sessions ($p = 0.048$, one-tailed t-test).

Table 1.3: Communication Characteristics and Effort Choice: Random Effects

	Dependent Variable: Effort
Ingroup	33.90***
	(11.608)
Outgroup	-0.19
	(11.388)
Lines	0.37
	(0.645)
Painting 6 Correct	-1.74
	(5.606)
Painting 7 Correct	-0.32
	(8.396)
Ingroup*Analysis	-0.50
	(0.509)
Outgroup*Analysis	-0.19
	(1.091)
Ingroup*Question	3.56**
	(1.528)
Outgroup*Question	-4.02*
	(2.287)
Ingroup*Agreement	0.06
	(1.415)
Outgroup*Agreement	1.51
	(1.692)
Ingroup*Engagement	-4.41
	(4.032)
Outgroup*Engagement	0.33
	(4.532)
Constant	139.84***
	(13.441)
Observations	5400
R^2	0.3549

Notes: Standard errors are adjusted for clustering at the session level.
Significant at the: * 10 percent level; ** 5 percent level; *** 1 percent level.

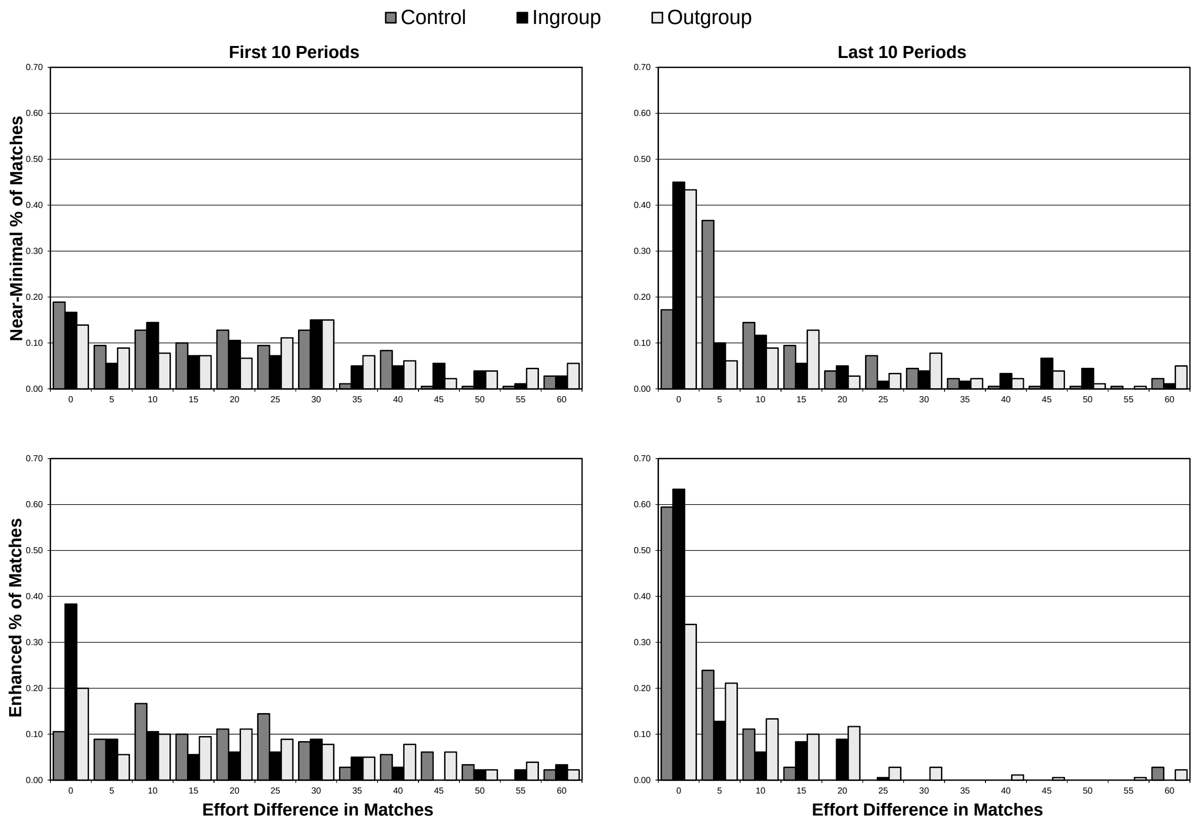

Figure 1.2: Wasted Effort in each Match for the First 10 Periods (Left Column) and the Last 10 Periods (Right Column), Separated by Near-Minimal (Top) and Enhanced (Bottom) Sessions

no waste in enhanced ingroup sessions, indicating a higher degree of equilibrium play than in the near-minimal treatments, the outgroup, or the control sessions. However, as we move to the last 10 periods, several changes occur. First, in all treatments, the fraction of matches that have little to no wasted effort increases greatly. As the game is repeated 50 times, subjects learn to coordinate with their matches, and are more successful in doing so than in the first 10 periods. Furthermore, the frequency of no waste is higher in the enhanced than the near-minimal treatments.

We now use a probit regression to investigate the significance of the observed patterns. In Table E.2 in Appendix E, we present the results of this regression, reporting the marginal effects. The dependent variable is a dummy variable indicating whether each pair is in an equilibrium (i.e. whether the subjects in each pair choose the same level of effort). The independent variables are an ingroup dummy, an outgroup dummy, an ingroup-enhanced interaction term, and an outgroup-enhanced interaction term. The definitions of the independent variables are the same as described above for the effort choice regressions. We summarize the results below.

Result 5 (Group effect on coordination)**.** *In the near-minimal sessions, matches in the ingroup, outgroup, and control sessions coordinate to an equilibrium at about the same rate. In the enhanced sessions, matches in the ingroup sessions coordinate to an equilibrium significantly more often than subjects in the control or outgroup sessions while subjects in the outgroup sessions do so at about the same rate as those in the control sessions. Increased group salience significantly increases the rate of coordination in the ingroup treatment, but not in the outgroup treatment.*

Support. *In the regression, neither the coefficient for the ingroup dummy (*$p = 0.186$*) nor that for the outgroup dummy (*$p = 0.821$*) are significant. A test of equality of the ingroup and outgroup dummies yields* $p = 0.270$*. The coefficient on the interaction term between the ingroup dummy and the enhanced dummy is significant (*$p = 0.023$*), while the coefficient on the interaction term between the outgroup dummy and the* *enhanced dummy is not significant ($p = 0.918$). A test that the sum of the coefficients of the ingroup dummy and the ingroup-enhanced interaction term is equal to 0 yields* $p = 0.0005$*, while a test that this sum is equal to the corresponding outgroup sum yields* $p = 0.0162$*. Finally, a test that this outgroup sum is equal to 0 yields* $p = 0.775$*.*

Result 5 indicates that pairs in different near-minimal treatments choose the same effort level at about the same rate. Both the near-minimal ingroup and near-minimal outgroup sessions produce slightly higher probabilities of matching effort (by 14 and 2 percent for the ingroup and outgroup sessions, respectively), but neither increase is statistically significant. The result also shows that pairs of salient ingroup members are significantly more likely to give equal efforts than pairs of less-salient ingroup members (by 21 percent). Also, pairs of salient outgroup members are equally likely to give equal efforts when compared to pairs of less-salient outgroup members (a 1 percent increase in effort matching). Finally, the result indicates that, if we examine only the enhanced treatments, subjects in the ingroup sessions choose the same effort more often than subjects in either the outgroup or control sessions. While subjects in the ingroup sessions choose the highest effort level of 170 nearly exclusively by the end of 50 periods, making the probability of obtaining an equilibrium result more likely, subjects in the outgroup and control sessions seem unable to decide whether to choose the lowest effort level of 110 or the highest effort level of 170 even after 50 periods. The minimum effort in each pair is 110 as often as it is 170. This result generally supports the predictions of the theoretical model.

Next, we examine efficiency in each treatment, as defined in Section 1.5. The average efficiency in each session and the overall efficiency in each treatment are presented in Table E.3 in Appendix E. To evaluate the statistical significance of the treatment effects on efficiency, we present a random-effects regression in Table E.4 in Appendix E. The dependent variable is the efficiency of each pair. The independent variables of the regression are the ingroup and outgroup dummy variables, and the ingroup-enhanced and outgroup-enhanced interaction terms.

Consistent with the treatment effects on individual behavior, we find that, in the near-minimal sessions, there is no significant difference in efficiency across the ingroup, outgroup and control treatments ($p > 0.10$ for all pairwise comparisons). Furthermore, in the enhanced sessions, efficiency in the ingroup treatment is significantly higher than that in the control and outgroup treatments ($p < 0.01$). Therefore, efficiency increases when subjects are matched with members of their own group, but only when groups are more salient. This finding is consistent with the predictions of the model in Table A.1 (column 4) in Appendix A. However, the predicted efficiency in Table A.1, e.g., when α approaches 1, is generally lower compared to the actual achieved efficiency, e.g., in the enhanced ingroup treatment (in Table E.3 in Appendix E). This is because most of our subjects are able to coordinate on integer values while the computation reported in Table A.1 assumes a continuous strategy space.[15] Subjects in the enhanced ingroup sessions, by coordinating on the highest effort level, are able to achieve much greater efficiencies than subjects in either the control or outgroup sessions. Coordination on the lowest effort level occurs in both the control and outgroup sessions, causing them to be fairly similar in terms of efficiency.

1.6.3 Learning Dynamics and Group Identity

While our reduced-form regression analysis establishes the significance of the effect of enhanced group identity on effort, equilibrium selection and efficiency, it does not provide an explanation for the learning dynamics observed in both Figures 1.1 and 1.2. In this subsection, we estimate a structural learning model and thus demonstrate the interaction between group identity and learning. In what follows, we first examine initial round choices and learning dynamics. We then estimate the parameters of the structural model and use these estimates to run a simulation. Finally, we compare

[15] Of all choices, 92.8-percent are integer choices.

choices in the final rounds with the predictions of our logit equilibrium model with calibrated parameters.

We first examine whether any significant behavioral differences exist in the initial round choices. Using Kolmogorov-Smirnov tests of the equality of distributions for first-round effort choices, we find that, within the near-minimal and enhanced treatments, only one of the pairwise comparisons is significantly different: near-minimal outgroup $\neq$ control ($p = 0.043$, two-sided). Likewise, comparing the near-minimal treatments with the corresponding enhanced treatments, only one of the pairwise comparisons is significantly different, and that only weakly: NM outgroup $\neq$ E outgroup ($p = 0.083$, two-sided). Thus, we observe none of the significant treatment effects in the first round.

We now use a structural learning model to explain the effects of group identity on the dynamics and convergence to various equilibria of the minimum-effort game. To do so, we look for a learning algorithm which incorporates key features of the adaptive learning models in the theoretical derivations (Monderer and Shapley, 1996). A model which meets this criterion is the stochastic fictitious play model with discounting (Cheung and Friedman (1997), Fudenberg and Levine (1998)). Unlike the deterministic fictitious play used for the theoretical analysis in Monderer and Shapley (1996), the stochastic version allows decision randomization and thus better captures the human learning process. It also more closely follows our theoretical model, which uses decision randomization.

In our stochastic fictitious play model, player i holds a belief regarding her match's effort level x_j in every period t based on history. We calculate this belief using a weight function $w_i^t(x_j)$. This weight function assigns to each of her match's possible effort levels a number which is positively correlated with the number of times she has seen her match give that level of effort in the past. She believes that the more times her match has given a particular effort level, the more likely it is that her match will give that effort level again. Note that for this analysis, we use a discrete strategy space. The initial value of this weight function, $w_i^1(x_j)$, is left unspecified by the model. This function is then updated using the following rule:

$$w_i^{t+1}(x_j) = \delta \cdot w_i^t(x_j) + \begin{cases} 1 & \text{if } x_j = x^t \\ 0 & \text{otherwise,} \end{cases} \tag{1.10}$$

where x^t is the effort level exhibited by player i's match in period t, and δ is the discount factor which gives distant experience less weight than recent ones. Player i's beliefs in period t are then calculated as follows:

$$\mu_i^t(x_j) = \frac{w_i^t(x_j)}{\sum_{x_k} w_i^t(x_k)}. \tag{1.11}$$

Equation (1.11) captures player i's beliefs about the likelihood that her match will use each strategy in the upcoming period. These beliefs are then used to calculate player i's expected utility for playing a strategy x_i:

$$\overline{u}_i^t(x_i) = \frac{1}{(\overline{x} - \underline{x})} \sum_{x_j} [u_i(x_i, x_j) \cdot \mu_i^t(x_j)], \tag{1.12}$$

where $u_i(x_i, x_j)$ is as defined in Equation (1.6). We assume that all subjects in a given session have the same group-contingent other-regarding parameter, so $\alpha_i^g = \alpha^g \; \forall i$ in the same session. Our incorporation of group-contingent social preference into a learning model follows the recent literature (Cooper and Stockman (2002), Arifovic and Ledyard (2009)) which merges the social preference and learning models to explain behavioral regularities in public goods experiments that cannot be satisfactorily explained by either social preference or learning alone.

Using this expected utility, player i determines which strategies are the best for her, choosing strategies with higher expected payoffs more frequently. Specifically, she randomly chooses an effort level x_i with a distribution defined by the following:

$$f_i^t(x_i) = \frac{\exp[\lambda \cdot \overline{u}_i^t(x_i)]}{\sum_{x_k} \exp[\lambda \cdot \overline{u}_i^t(x_k)]}, \tag{1.13}$$

where λ is the inverse noise level that describes how much randomization a player will employ. With this specification, as $\lambda \to 0$, the player uses full randomization, and as $\lambda \to \infty$, she plays her best response to her belief of what her match will play with probability 1. This model has three parameters: the sensitivity parameter λ, the discount factor δ, and the other-regarding parameter α^g.

We next compare the observations from our experiment to the predictions of the above model. Performing a grid search over the three parameters, we calculate a score using the quadratic scoring rule described in Selten (1998) for each subject and round. In any given round, let $f_{ij} = (f_{i1}, \ldots, f_{iK})$ be the predicted probability distribution over player i's strategies, where K is the number of strategies available to the players, and $a_{ij} = (a_{i1}, \ldots, a_{iK})$ be the observed relative frequency distribution over player i's strategies, where $a_{ij} = 1$ if player i chooses action j, and zero otherwise. This score, $S_i(f)$, is calculated by $S_i(f) = 1 - \sum_{j=1}^{K} (a_{ij} - f_{ij})^2$. Our estimates for the parameters are the values of λ, δ, and α^g that give the highest summed score in each session (over all subjects and rounds).

We perform the calibration in 2 steps. First, we allow λ to vary from 0 to 7 in increments of 0.1, δ to vary from 0 to 1 in increments of 0.1, and α^g to vary from -1 to 1 in increments of 0.1.[16] We perform this analysis over all sessions at once. Next, we fix λ and δ at the calibrated values, then recalibrate α^g, allowing the parameter to vary from -1 to 1 in increments of 0.01. This part was performed on the session level. Each calibration consists of the following steps. First, we set all subjects' initial beliefs regarding their matches' first-period efforts to the empirical distribution of first

[16] The upper bound for λ is based on an initial exploration where we tested fewer values of λ over a larger range.

period effort levels in the subjects' sessions. For each subsequent period, we update a subject's beliefs based on the history of effort levels that the subject has observed from her matches according to Equation (1.10), and calculate the probability distribution of effort levels that the subject is predicted to play according to Equation (1.13). We then use the quadratic scoring rule to calculate a score for the particular combination of λ, δ, and α^g for each period the subject plays, and sum the period scores in order to obtain a score for that subject. After we have completed this process for every combination of λ, δ, and α^g, we find the parameters that give the highest score (or, in the recalibration, the α^g that gives the highest score in each session).

The results for the analysis are reported in Table 1.4 in the rows labeled "Near-Minimal" and "Enhanced," with treatment averages reported in the rows labeled "Average." The globally calibrated inverse noise and discount parameters are $\lambda = 3.0$ and $\delta = 0.7$, respectively. For our purposes, the most important parameter is α^g, which measures the level of group-contingent social preference. As expected, we see that the enhanced ingroup treatment obtains the highest average α^g, consistent with our effort and efficiency results. Also, every session of the enhanced ingroup treatment achieves a higher α^g than any session of the near-minimal control. Using a permutation test, this comparison (enhanced ingroup > near-minimal control) is significant ($p = 0.05$). The other comparisons are not significant since every other treatment has one session in which the subjects converge to the efficient equilibrium.

To connect our learning model with the logit equilibrium model discussed earlier, we use the calibrated values of α^g from the learning model to compute theoretical distribution functions of effort choices in the logit equilibrium, i.e., Equation (1.9). We then compare the means and standard deviations of these theoretical distributions with the actual means and standard deviations of the effort choices in the last 5 rounds. We perform this analysis on a treatment level. These values are reported in Table E.5 in Appendix E. The means for the theoretical and actual distributions

Table 1.4: α^g Calibration of the Stochastic Fictitious Play Model ($\lambda = 3.0$, $\delta = 0.7$)

Treatments	Sessions	Control	Ingroup	Outgroup
Near-minimal	1	0.26	0.45	0.32
	2	0.28	0.23	0.82
	3	0.68	0.84	0.81
	Average	0.41	0.51	0.65
Enhanced	1	0.07	0.80	0.94
	2	0.87	1.00	0.09
	3	0.00	0.70	0.22
	Average	0.31	0.83	0.42

all fall within one standard deviation of each other, with the highest and lowest actual average efforts mirrored in the highest and lowest theoretical average efforts, respectively. We take this as a sign that the theoretical model performs well in describing the data in the final rounds.

1.7 Conclusion

In this paper, we study the effects of social identity on one of the most important and yet unresolved problems in game theory, the problem of equilibrium selection in games with multiple Nash equilibria. By incorporating group-contingent social preferences into Monderer and Shapley's theory of potential games, we make theoretical predictions on how and when salient group identities can influence equilibrium selection. We also provide a unifying framework for a number of previous experimental studies performed on coordination games in Appendix F.

To further test the ability of this model to predict behavior in an experimental setting, we design an experiment that uses induced group identity to increase group-contingent other-regarding preferences in the minimum-effort games. In our near-minimal treatments, we show that, while matching subjects with ingroup or outgroup members when playing the minimum-effort game has some effect on the effort levels chosen, they are not statistically distinguishable from the control, where no groups are induced. On the other hand, when we enhance the groups by allowing them to communicate with group members in solving a simple task before playing the minimum-effort game, we find that matching subjects with ingroup members has a statistically significant positive effect on subject effort. When inducing groups, we find that it is only after the groups are made more salient that we see an effect on the provided effort. These findings are consistent with the predictions of our model.

In order to understand the mechanism through which this result is achieved, we incorporate group-contingent social preferences into a learning model of stochastic fictitious play. This enables us to specify the effect that creating groups and increasing their salience has on subjects' other-regarding preferences. The calibrated model also does well in predicting the empirical actions used by the subjects.

Our paper contributes to the theoretical foundations of social identity by demonstrating that, by using a simple group-contingent social preference model, we can derive the comparative statics result that stronger group identification leads to higher effort in equilibrium (Proposition 4), whereas a higher (lower) effort level is assumed to be the exogenous behavioral norm of a worker who identifies (does not identify) with an organization in Akerlof and Kranton (2005). Although our framework captures only a piece of what Akerlof and Kranton call identity, it can explain experimental findings in a number of coordination games.

Beyond the fundamental problem of understanding and modeling identity on economic behavior, our results have practical implications for organizational design. As the world becomes more integrated, organizations are more frequently encountering the issue of integrating a diverse workforce, and motivating members from different backgrounds to work towards a common goal. Our paper demonstrates that creating a deep sense of common identity can motivate people to exert more effort to reach a more efficient outcome.

A successful application of this idea comes from Kiva (http://www.kiva.org/), a person-to-person microfinance lending site, which organizes loans to entrepreneurs around the globe. In August 2008, Kiva launched its lending teams program, which organizes lenders into identity-based teams. Any lender can join a team based on her school, religion, geographic location, sports, or other group affiliation. As of November 2010, the top five most successful teams are "Atheists, Agnostics, Skeptics, Freethinkers, Secular Humanists and the Non-Religious," "Kiva Christians," "Team Europe," "Poverty2Prosperity.org - Poverty-Escape," and "Team Obama." The lending teams program substantially increases the amount of funds raised.

There are several directions for future research. A possible next step in this line of research would be to extend this result to other coordination games, such as the provision point mechanism. Our model predicts that successful coordination to higher levels of public goods can be achieved systematically even with a very weak method of increasing other-regarding preferences. Another possible direction would be to evaluate the effect of identity-based teams in the field through natural field experiments in fundraising or online communities.

CHAPTER II

Social Identity in Online Microfinance

2.1 Introduction

Currently about 10 million households around the world are served by microfinance programs, which help very poor households meet basic needs, improve household economic welfare, empower women, and promote entrepreneurship.[1] While most research on microfinance focuses on the demand (i.e., borrower) side, we investigate the supply (i.e., lender) side. Specifically, we analyze incentives to increase lender participation and lending activities on Kiva (`http://www.kiva.org/`), the first peer-to-peer microlending site which matches citizen lenders with low-income entrepreneurs in developing countries and small businesses in selected cities in the United States. Through Kiva's platform, anyone can make an interest-free loan of $25 or more to support an entrepreneur. In February 2012, the total value of all loans made through Kiva reached $286 million.

In August 2008, Kiva launched its lending teams program, where lenders can organize or join teams of other lenders. Along with the establishment of lending teams, a leaderboard prominently features the top teams sorted by total amount loaned, and more recently, the number of new members, to encourage team competition. Since August 2008, more than 20,000 teams have been created, many of which are organized

[1] cgap.org

based on lender group affiliations, such as school, organization, religious affiliation, geographic location, or sports. In particular, many of the highly ranked teams are identity-based. For example, as of today, the top two most successful teams in terms of total amount loaned are "the Atheists," followed by "Kiva Christians."

Because of the prevalence of identity-based teams, we evaluate the effects of the lending teams program on pro-social lending from the lens of social identity theory. Social identity is commonly defined as a person's sense of self derived from perceived membership in social groups. The large body of empirical work on social identity throughout the social sciences has established several robust findings regarding the development of a group identity and its effects on behavior. Most fundamentally, the research shows that group identity affects individual behavior. For example, Tajfel et al. (1971) and Tajfel and Turner (1979a) find that group membership creates ingroup enhancement in ways that favor the ingroup at the expense of the outgroup. In economics, experimental research demonstrates that intergroup competition can increase public goods provision (Eckel and Grossman, 2005) and improve coordination (Gary Bornstein and Nagel, 2002; Croson et al., 2008; Chen and Chen, 2011). Pro-social lending has elements of public goods provision and coordination.

Of particular note is the work of Akerlof and Kranton (2000) that brought social identity to the economics literature. In their proposed model, teams have existing social norms. Someone who joins a team and adopts the team's identity will want to adhere to the team's norms. That person receives disutility if she deviates from those norms. In the model, that person also obtains utility for being a member of the team. Applied to Kiva's lending teams program, a question arises: does belonging to a team increase lending? Due to their existence on Kiva, a social norm of lending should exist within all teams. We explore whether this can affect lending behavior.

While joining a lending team is associated with more lending (Liu et al., 2012), we disaggregate this correlation into two separate effects. First, we examine the decision to join teams, i.e., who is more likely to join a team? Second, we examine the effects of team competition on lending behavior for those who belong to a team. To perform this analysis, we download individual lender and lending team activity history using Kiva's application programming interface (API). This data is made publicly available by Kiva to aid application writers and researchers.

Although, on average, belonging to a team has a positive effect on lending, we observe considerable heterogeneity among the twenty thousand teams. While the top lending teams are characterized by their vibrant forums and lending, many teams become dormant shortly after their creation. Of the 20,000 lending teams, 43% have not made a single loan in the past twelve months.[2] Using downloaded communication data from the online forums, we explore factors that differentiate successful teams from dormant ones.

The results of our analysis will benefit Kiva and other online microfinance programs. More broadly, they can be applied to other online communities which provide public goods.

2.2 Literature Review

In this section, we will briefly review the social identity literature as well as the microfinance literature. The large body of theoretical and empirical work on social identity throughout the social sciences has established several robust findings regarding the development of a group identity and its effects. Most fundamentally, research shows that group identity affects individual behavior. Following Tajfel et al. (1971), which demonstrates ingroup favoritism and outgroup discrimination, many experiments in social psychology identify factors which enhance or mitigate ingroup favoritism. Furthermore, as a person derives self-esteem from the group she identifies

[2]We select a random sample of 200 teams from those below the top ten, and find that 77% have no forum messages.

with, salient group identity induces people to conform to stereotypes (Shih et al., 1999).

Since the seminal work of Akerlof and Kranton (2000), there has been increased interest in social identity research in economics, yielding new insights into phenomena which standard economic analysis on individual-level incentives proves unable to explain. Social identity models have been applied to the analyses of gender discrimination, the economics of poverty and social exclusion, the household division of labor (Akerlof and Kranton, 2000), contract theory (Akerlof and Kranton, 2005), economic development (Basu, 2006), and public goods provision. In the context of public goods provision, experimental research demonstrates that intergroup competition can increase public goods provision in the voluntary contribution game (Eckel and Grossman, 2005) and improve coordination in the minimum effort game (Gary Bornstein and Nagel, 2002; Croson et al., 2008; Chen and Chen, 2011) and the battle of sexes game (Charness et al., 2007). The main results are summarized in Akerlof and Kranton (2010).

A second stream of literature on the economics of microfinance largely focuses on the borrower side (Armendáriz and Morduch, 2010). With the recent popularity of large-scale peer-to-peer lending enabled by the Internet, a few recent studies on online microfinance examine lender motivations (Liu et al., 2012), their biases (Jenq et al., 2012) and sensitivity of giving to transaction costs (Meer and Rigbi, 2011). For example, Kiva lenders appear to favor more attractive, lighter-skinned and less obese borrowers. They also fund loan requests from borrowers who appear needy, honest and creditworthy more quickly (Jenq et al., 2012). In a field experiment at Kiva which varies the language of the loan requests, Meer and Rigbi (2011) find that untranslated requests take substantially longer to obtain full funding than translated ones, indicating that transaction costs arising from translations significantly deter funding speed. The authors also find that social distance plays a role in funding decisions. Closer to our study, Hartley (2010) reports observations of 120 lending teams on Kiva over a two-month period. The author compares team size and openness on group lending activities. Because of the small sample size, the study generates a number of useful hypotheses, without offering econometric analysis.

While Kiva loans are interest free, other microfinance websites allow users to make loans for profit. Lin et al. (2009) analyzes the effects of social networks from the borrower's perspective with the online peer-to-peer lending site Prosper.com. They find that the quality of relations between borrowers and lenders is more important than the borrowers positions in the network. The information asymmetry between borrowers and lenders will be alleviated if borrowers have verifiable ties to more lenders, i.e. more real lender friends who bid on their loans. It further results in a higher probability of funding, a lower probability of default and a lower interest rate. They also study the effects of group affiliation. The results show that members of alumni based groups and geography based groups have a greater chance of being funded because of high verifiability. Interestingly, members of religious groups benefit as well.

A possible interpretation of the results in the microfinance literature is the importance of social identity. Jenq et al. (2012) finds lender biases against borrowers with darker skin and Meer and Rigbi (2011) find lender biases against borrowers whose descriptions are not in English. Given the prevalence of lenders from the United States on Kiva, this could be a reflection of outgroup bias, or bias against those that are perceived as not in the lender's group. Darker skin tone and use of a foreign language highlight this perception and may contribute to the bias.

In the current study, we focus on the lender side and the effects of lending teams. This work contributes to various parts of the economic and social psychology literatures. As mentioned, social identity has been used to explain results in the microfinance literature. In this paper, we explicitly study the effects of identity on pro-social lending by estimating the degree to which social identity increases lending behavior. The research on social identity comprises theory, laboratory experiments, and field work, with contributions from both social psychologists and economists. The Kiva data that we use allows us to add to this literature with a large-scale empirical study on the effects of social identity. Also, we combine econometric and data mining techniques in our analysis to produce our estimates.

2.3 Kiva

Kiva, which was founded in October 2005, is the world's first peer-to-peer microfinance website. Kiva partners with microfinance institutions to allow its users to make loans of at least $25 to specific borrowers.[3] Most Kiva users, who we will call lenders, make at least one loan through Kiva. Table 2.1 displays some summary statistics relating to Kiva.

While a majority of lenders actually make a loan, one-third do not. This highlights the main issue that Kiva faces. The membership of Kiva, along with the number of loans made through Kiva, has increased greatly since Kiva's inception. Figure 2.1 displays the number of loans per month on Kiva from October 2005 to December 2010. However, only a few lenders give many loans, and many lenders give few to no loans, as shown in figure 2.2. This distribution of loans is very uneven, resulting in a group of "core" lenders and a large group of "peripheral" lenders. One issue, as discussed by Premal Shah, the president of Kiva, is that many Kiva lenders would lend once, then never come back to the site.[4] This is despite the fact that they were being repaid and could have made another loan for no extra money.

In August 2008, Kiva instituted their lending teams program, in part to combat this problem. Any lender is allowed to create or join as many teams as they wish.

[3]For details, see `http://www.kiva.org/about/how/more`

[4]Source: `http://www.youtube.com/watch?v=aEC3OwKWgfc`, retrieved on 6/5/2012.

Table 2.1: Summary Statistics

Number of Lenders	884,692
Lenders who have made:	
No loans	315,543
At least 1 loan	569,149
Avg. number of loans	13.41
At least 5 loans	246,673
Avg. number of loans	28.26
Lenders with:	
Location information	500,131
Motivation statements	129,731
Occupation information	436,986
Lenders who are:	
Male	175,219
Female	318,172
Companies	1,132
Families	3,558
Couples	3,171
Lenders who are members of:	
At least 1 team	159,833
Avg. number of teams	1.45
At least 1 top 10 team	37,394
Avg. number of teams	2.25
Number of Teams	22,322
Avg. Lenders per Team	11.92
Avg. Loans per Team	120.55

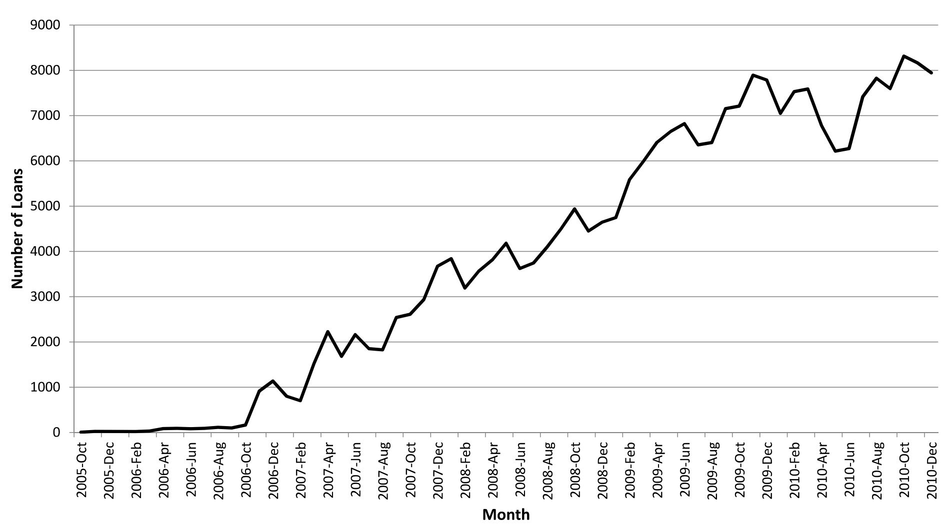

Figure 2.1: Loans per Month

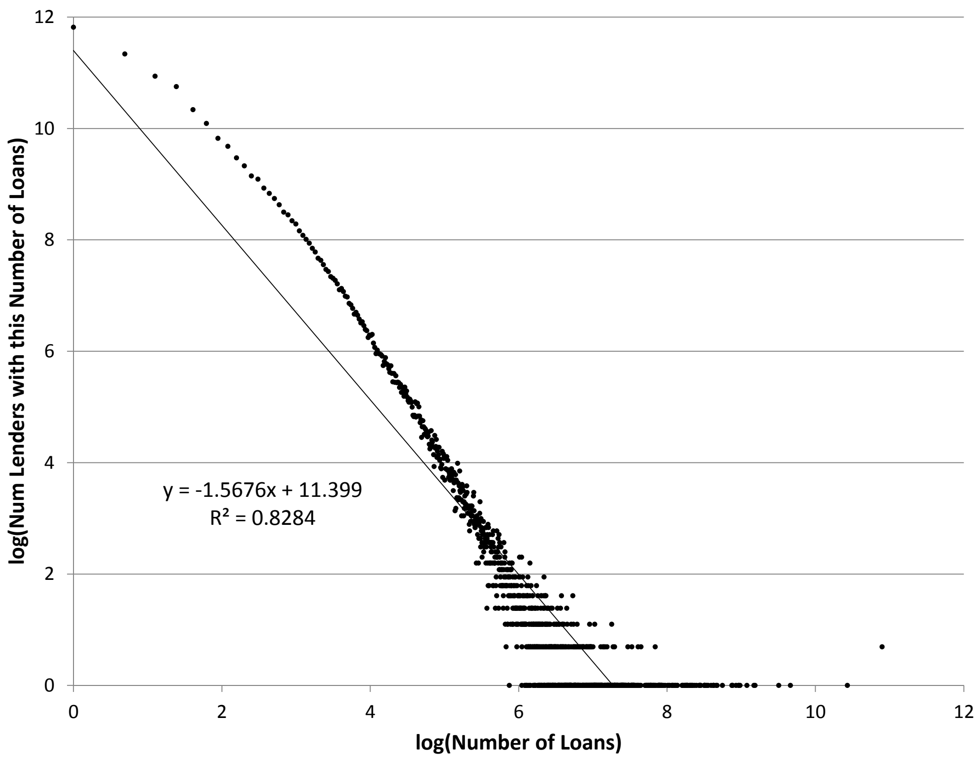

Figure 2.2: Loans per Lender

Once a team is created, it is displayed on Kiva's team leaderboard (http://www.kiva.org/community). Lenders who have joined at least one Kiva team makes loans as usual through Kiva, just like lenders who have not joined any teams. However, lenders who have joined teams are asked, when they make a loan, whether they want to assign that loan to any of their teams. For a given loan, a lender may choose to assign it to either no teams or one team. By default, the team leaderboard sorts the teams by the total loan amounts that their team members have assigned to them. As of June 2012, the top lending teams displayed on the leaderboard are "Atheists, Agnostics, Skeptics, Freethinkers, Secular Humanists and the Non-Religious," "Kiva Christians," "milepoint," "Team Europe," and "Friends of Bob Harris."[5]

Premal Shah suggests that the lending teams program was designed to increase lender engagement on Kiva and to make Kiva "as fun and compelling as possible". Another aspect of the lending teams program comes in the form of team competition. The captain (creator) of the "Atheist" team has stated that he believes "[t]he whole idea of teams in the Kiva context implies there should be competition."[6] With the lending teams system, Kiva seems to be attempting to increase the lending of the peripheral users while keeping the core users interested.

2.4 Data

In this section, we introduce the approaches we apply to collect and analyze behavioral data on Kiva users. The data regarding lending behavior on Kiva offers a unique opportunity for us to examine social identity in a real-world context.

[5]Source: http://www.kiva.org/community, retrieved on 6/5/2012.

[6]Source: http://www.kiva.org/team/atheists/messages?msgID=60182#msg_60182, retrieved on 9/20/2011.

2.4.1 Data Sources

We collect Kiva data from two main sources. First, we download data from Kiva's (API), located at `http://build.kiva.org/api`. This provides us with a snapshot of the information that Kiva collects about its lenders, teams and loans from February, 2012. Two important pieces of data that we collect from Kiva's API are the number of loans each lender has made and a list of teams that each lender has joined. For a full listing of the variables that we obtain from this source, see the appendix G. Some of the variables listed in that appendix, such as location information, are provided to us in free text form, which we must process before using them in our analysis.

Secondly, we recruit subjects to code other information which cannot be extracted from the API in a usable form. Three variables we obtain from this process are lender gender/group type, occupation, and motivation for lending. For the occupation and motivation variables, lenders provide free text describing their jobs and why they are making loans on Kiva.[7]

We have coders code a subset of lenders, then train classifiers based on these examples to automatically code the rest. For gender/group type, we ask the coders to look at each lender's username and profile picture to determine whether the lender is male, female, a couple, a company, a family, or another type of group. For occupation and motivation, we have our coders code the lenders' free-text descriptions into occupation and motivation categories. For a complete list of the categories we employ, see appendix H. Prior to this coding task, the researchers code a subset of these profiles and randomly place them among the lender profiles that the coders see. The coders are paid based on how many of these pre-coded profiles they match to ensure that they are paying attention to the coding task.

It is important to note that there are important variables that Kiva does not

[7] The prompts on Kiva are "Occupation" and "How would you decribe your work?" for lender occupation, and "I loan because:" for lender motivations. See Liu et al. (2012) for a more detailed analysis of Kiva lender motivations.

provide, likely due to privacy concerns. First, Kiva does not report the amounts that lenders lend. While we know the number of loans each lender gives, we do not know how much money is lent for each of these loans, except that the minimum loan amount is $25. Second, Kiva does not provide information on loan assignment to teams. We only know the aggregate number of loans that has been assigned to each team. This second variable would allow us to examine questions of multiple identities in the Kiva context, including which of a lenders numerous teams are most important to the lender, and how multiple identities affect lending behavior. In the future, Kiva may choose to provide this data, but at the moment we are unable to perform these analyses.

2.4.2 Similarity measures

A significant body of empirical network literature tries to explain why users join certain communities (Girvan and Newman, 2002; Newman, 2006; Backstrom et al., 2006; Leskovec et al., 2007; McPherson et al., 2001). It is often observed that people tend to associate with others whom they perceive as similar to themselves in some way, a phenomenon known as *homophily* or assortative mixing in the sociology literature. We wish to understand how homophily plays a role in the Kiva lending community. To that end, we derive two similarity measures which we use in our analysis.

The first is a location similarity measure. This provides a measure of how similar two lenders perceive themselves to be in location. We use a hierarchical network model to define the location similarity measure based on users' location information, as displayed in figure 2.3. This model has primarily been used to study social networks (Watts et al., 2002), though it has also been used to examine location similarity between individuals (Liben-Nowell et al., 2005).

The location similarity between two users i and j is denoted $l_{ij} \in \{0, 1, 2, 3\}$. This is defined as the level of their closest common parent node. For example, if

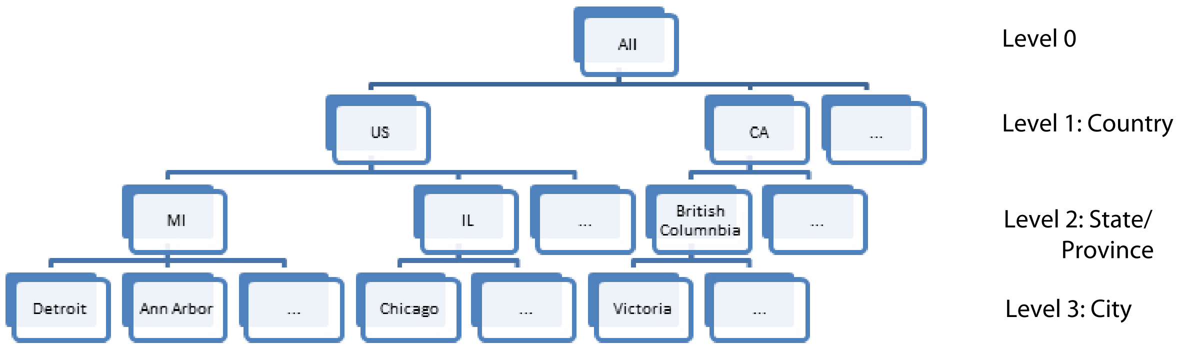

Figure 2.3: Location similarity hierarchy

two lenders are in two different countries, the similarity score will be 0. If Lender 1 is from Ann Arbor, MI and Lender 2 is from Chicago, IL, their closest common parent node is the US, giving these two lenders a location similarity of 1. If Lender 1 is from Ann Arbor, MI and Lender 2 is from Detroit, MI, their closest common parent node will be Michigan, giving them a location similarity of 2. Therefore, a higher location similarity between two users indicates closer jurisdictional proximity between two users. For this measure, we assign a location similarity of 0 to all pairwise comparisons where at least one lender does not provide location information to Kiva. When one lender does not provide this information, it makes sense that the lenders will perceive their location similarity to be 0.[8]

Also note that this measure does not refer to two lenders' geographic proximity. While geographic and jurisdictional proximity will be correlated, we suspect that people within the same jurisdiction but far apart geographically are more similar to each other than people who are close geographically but in different jurisdictions.

[8]This measure can also be seen as the worst case scenario, where we assume that each lender who does not provide location information lives in their own country. Other measures are possible here, but we feel this one more closely captures the similarity to others that the lenders themselves feel.

People who are in the same country or city are likely to perceive themselves as more similar due to social identity effects. These are less likely to occur due to mere geographic proximity.[9]

The second measure we employ is one of motivation similarity. This provides a measure of the perceived similarity between two lenders' motivations for lending. First, each motivation statement is converted into a $1 \times m$ vector, where m is the number of unique terms in the entire collection of statements. If a term appears in the statement, that term's component in the vector is 1, and 0 otherwise. Our motivation similarity measure then calculates the cosine between two lenders' motivation statements:

$$Similarity = \frac{A \cdot B}{\|A\| \, \|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}$$

where A and B are motivation statement vectors of the two lenders. This measure, known as *cosine similiarity* is commonly used in network analysis (Manning et al., 2008; Newman, 2010) and to compare texts in text mining (Tan et al., 2005). Alternatively, we could have derived a motivation similarity measure from our coded motivation categories, but cosine similarity is more objective than a measure derived from those codings, as it bypasses the determinations of human coders. Again, as with the location similarity measure, we assign a value of 0 to those who do not provide motivation statements.

Unfortunately, there will be many dimensions of homophily that we do not observe. While there may be relevant dimensions of homophily that might affect how a lender chooses her team membership, such as religious affiliation, the demographics that Kiva collects do not include these dimensions.

[9] For example, someone from Ann Arbor, MI will likely feel more similar to someone from the Upper Peninsula of Michigan than to someone from Columbus, OH, even though Ann Arbor is closer to Columbus than the Upper Peninsula.

2.5 Analysis

In this section, we analyze the data obtained from Kiva, examining which teams lenders choose to join and whether joining these teams increases lending. We also explore the factors that might differentiate active teams from dormant ones.

2.5.1 What teams do lenders join?

We first present two competing hypotheses regarding team selection. First, following the findings in the sociology literature on homophily, we posit that a lender is likely to join a team whose members are most similar to her along some dimension. For example, using the "search teams" function, a lender might find and join her city team, or she might join her alumni team. By browsing through various team homepages, she might join a team whose members' motivation statements resonate with her.

Hypothesis 8 (Homophily)**.** *A lender will join a team whose members are most similar to her.*

The possible types of similarity that we explore are location similarity and motivation similarity, as described in the data section. For each measure, we first calculate the lender-lender similarity between each pair of lenders. Then, for each lender, we find the average of these lender-lender similarities for each existing team, giving us a lender-team similarity between each lender and each team. Finally, we break the teams into those that the lender has and has not joined and take the average of these lender-team similarities in each subset, treating each team as a unit. This gives us an ingroup and an outgroup lender-team similarity measure for each lender. We perform a signed-rank test on these measures to examine which, if either, of these similarity measures can significantly predict whether a lender will join a given team.

For the motivation similarity measure, the average ingroup lender-team similarity is 0.014 and the average outgroup lender-team similarity is 0.009. A signed-rank test indicates that this difference is significant at the 1% level ($p < 0.0001$, two-sided test). For the location similarity measure, the average ingroup lender-team similarity is 0.417 and the average outgroup lender-team similarity is 0.099. A signed-rank test indicates that this difference is also significant at the 1% level ($p < 0.0001$, two-sided test). So, lenders are more locationally and motivationally similar to team members than to non-team members, suggesting that they might use these two measures when choosing their teams.

While homophily is a plausible reason for joining a team, because of the prominence of the leaderboard, we expect that lenders might be attracted to teams that rank highly. This could be related to status seeking in the economics literature, and preferential attachment in social networks (Barabási and Albert, 1999). Also, lenders may wish to join teams that are perceived to have a large impact on lending.

Hypothesis 9 (Team Rank)**.** *A lender is more likely to join a highly-ranked team.*

To explore this possibility, whenever a lender joins a team, we find the rank of that team based on the number of past loans that have been assigned to that team. First, figure 2.4 shows that the frequency with which lenders join teams of a given rank follows an inverse power law. A large number of lenders join the top-ranked teams, while a quickly shrinking number join lower-ranked teams. This suggests that rank may be a factor that lenders take into account when choosing teams.

To compare these two hypotheses, we compare the ranks of the teams that lenders join using these different measures. When a lender joins a team, that team has a rank based on lending history, as described above. We can also calculate ranks for that team based on the lender-team similarity measures. For each measure, we rank the teams based on how similar they are to the lender in question, and record the rank of the team that the lender joins. By comparing the ranks of the joined teams based on these similarity measures to those based on lending history, we can examine whether

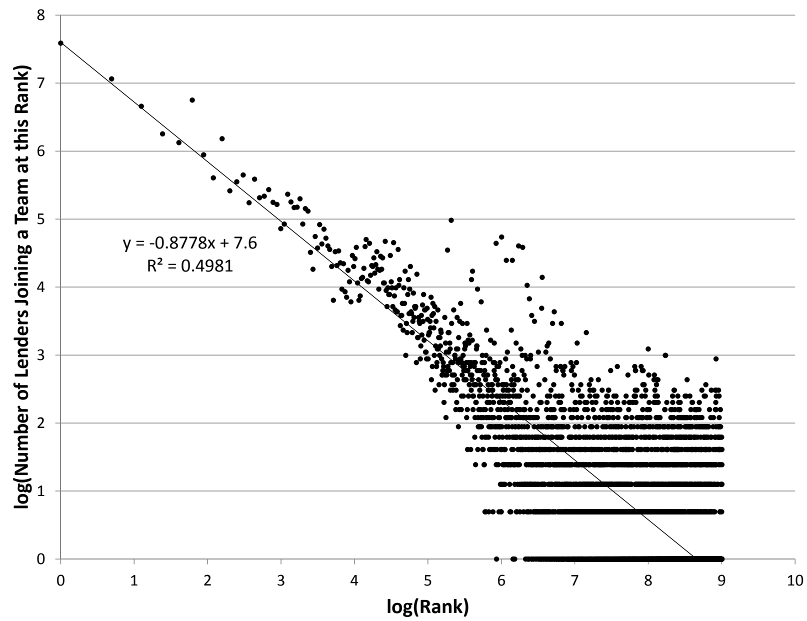

Figure 2.4: Log-log plot of the team ranks when lenders join each team.

homophily or team rank is more important to lenders when they are deciding on which teams to join.

The average team rank is 2261.61 using lending history, 3592.45 using lender-team location similarity, and 6543.80 using lender-team motivation similarity. Signed-rank tests show that all of these differences are significant at the 1% level ($p < 0.0001$, two-sided), indicating that lenders pay more attention to team rank using lending history than to either location or motivation similarity. This provides evidence that rank based on lending history is more important than homophily for the lenders who join teams on Kiva, lending support for hypothesis 9 over hypothesis 8. This result is likely due to the way that Kiva displays its lending teams. Kiva, by default, lists teams on the leaderboard ranked by their total lending amounts, which is highly correlated with the number of loans assigned to a team.[10] On the other hand, Kiva does not

[10]Kiva has not released data regarding lending amounts due to privacy concerns, so we perform the analysis using the number of loans instead. Using the team rankings from February 2012, the correlation between the rank of teams based on amount loaned and based on number of loans is 0.985.

prominently display information regarding the locations or motivations of current team members. Since lending history is made more salient than other information to the lenders, it is reasonable that lenders pay more attention to lending history when choosing their teams. Another possility, as mentioned in the previous section, is that there may be more relevant dimensions of homophily that we do not observe. If these dimensions could be taken into account, homophily might prove to be a more important consideration for subjects when choosing teams.

2.5.2 Does joining a team result in more lending?

We next discuss the effect of team affiliation on lending behavior. Applying the group-contingent social preference model of Chen and Li (2009) to the voluntary contribution public goods game, we obtain that a salient group identity leads to an increase of individual contributions to public goods. Thus, in the Kiva setting, we expect the following hypothesis to hold.

Hypothesis 10 (Effects of Teams on Lending)**.** *Joining a team increases a lender's lending activity level.*

Also, preliminary analysis shows that lenders who join teams lend substantially more on average than lenders who do not join teams. Figure 2.5 shows the average lending by lenders before and after teams were implemented by Kiva, broken down by those who joined at least one team and those who did not. Clearly, lenders who join teams loan much more than those who do not join teams. However, given the behavior of these lenders before the team system was implemented, it is possible that lenders who join teams are simply more active on Kiva in general.

There is evidence that the lenders who join teams are more engaged than those who do not. For instance, we can look at the proportion of the lenders who give Kiva demographic information about themselves. Of those who do not join teams, 11% give lending motivation statements, 53% give their locations, and 48% declare

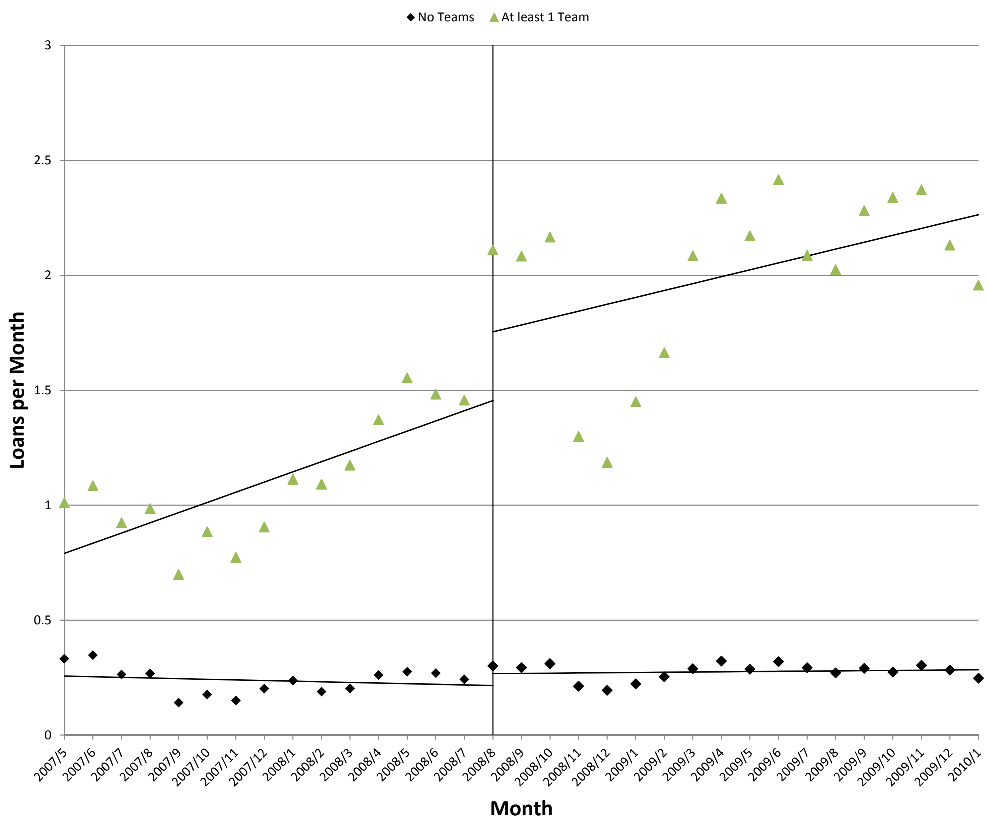

Figure 2.5: Average lending activity (loans per month), separated by whether or not lenders have joined at least one team on Kiva.

their occupations. In comparison, of those who join at least one team, 28% give motivation statements, 70% give their locations, and 56% state their occupations. Using chi-squared tests, the differences between those who join teams and those who do not are all highly significant (Pearson $\chi^2 > 3200$, $p < 0.001$ in all cases).

We can also examine the similarity measures for the different groups of lenders. For those who do not join teams, the average lender-team motivation similarity measure is 0.004 and the average lender-team location similarity measure is 0.087. For those who join at least one team, the average lender-team motivation similarity measure is 0.009 and the average lender-team location similarity measure is 0.099. Mann-Whitney U tests show that these differences are highly significant ($p < 0.0001$).

Without further analysis, it is unclear whether Kiva's team design has any effect on overall lending or whether it simply shifts the most active Kiva users into teams. To resolve this issue, we employ an instrumental variables analysis. Our econometric model is the following:

$$average\ loans_i = \beta_0 * constant + \beta_1 * joined\ team_i + \mathbf{B} \cdot Demographics_i + \epsilon_i$$

where $average\ loans_i$ is the average number of loans (loans/day) given by lender i, $joined\ team_i$ is a dummy variable for whether lender i has joined a team, and the $Demographics_i$ are demographic variables including location, gender/type, and occupation. We expect the variable $joined\ team_i$ to be correlated to the error term ϵ_i, and this is verified using a standard exogeneity test ($p < 0.0001$). Therefore, we instrument for $joined\ team_i$ using a version of the location lender-team similarity measure. This instrument is calculated by finding the average similarity measure between each lender and each team, then taking the maximum of these values (i.e. the similarity between the lender and her most similar team, regardless of her membership in that team). We expect this measure to affect lending activity only through its correlation with whether the lender joins a team, and we expect this correlation to exist due to the large number of location-based teams (e.g. In the top 10 teams by amount loaned, Team Europe, Australia and Team Canada are location-based teams).

The results of the two-stage least squares instrumental variables regression are displayed in table 2.2. This table shows the results of two specifications of the regression. Columns (1) and (2) show the results of the first-stage regressions while columns (3) and (4) show the results of the second-stage regressions. The F-statistic on the location similarity instrument is greater than 100 in both cases, eliminating potential weak-instruments problems. In both specifications, we only include lenders that have made at least one loan, ignoring those who have signed up for Kiva but not done anything on the site. Also, we include the number of days that the lenders have been members of Kiva, as well as the square of this term to control for lenders who have not been on Kiva for very long. Columns (1) and (3) show the results of a regression with no demographic variables included (i.e. the $Demographics_i$ term is dropped). Increasing the location similarity measure by 1, which is equivalent to changing the parent location node of a pair of lenders from a country to a state or from a state to a city, increases the probability of joining a team by 5.48%, and joining a team increases the number of loans given per day by 0.06. This is an increase of about 1.8 loans per month or at least $45 per month, since Kiva's minimum loan amount is $25 per loan.

The second specification includes the demographic variables gender, occupation and lending motivation. As described in section 2.4, all three of these variables are derived from human coders and trained classifiers. The occupation and motivation categories are treated as dummy variables. For this specification (columns (2) and (4)), we restrict the lenders included in the regression to individuals that provide occupation and motivation information. As is evident from table 2.2, columns (2) and (4) provide results that are quite different from those in columns (1) and (3). This reflects the fact that the more engaged lenders, which comprise about 10% of the entire population of lenders, are not very similar to the general population of lenders. For instance, in the restricted sample of lenders, an increase of 1 in the location similarity measure increases the probability of joining a team by 3.29%. Most interestingly, joining a team has no significant effect on the average number of loans per day. This perhaps reflects Kiva's purpose in implementing the lending teams system: the peripheral lenders increase their lending when they join teams, and the core lenders do not decrease their lending.

We also run these same two specifications, but restricting the analysis to lenders

Table 2.2: Instrumental Variables Regression

	First Stage: Joined Team		Second Stage: Average Loans	
	(1)	(2)	(3)	(4)
Joined Team			0.0598***	-0.0037
			(0.004)	(0.013)
Location	0.0548***	0.0329***		
Similarity	(0.001)	(0.002)		
Constant	0.1101***	0.4647***	0.0405***	0.0645***
	(0.002)	(0.012)	(0.001)	(0.007)
Female		-0.0365***		-0.0107***
		(0.004)		(0.001)
Occupation				
Entertainment		-0.0110*		-0.0027**
		(0.006)		(0.001)
Business		0.0006		0.0037***
		(0.006)		(0.001)
IT		0.0955***		0.0045**
		(0.007)		(0.002)
Education		0.0186***		-0.0063***
		(0.005)		(0.001)
Engineering		0.0539***		0.0062***
		(0.010)		(0.002)
Health Care		-0.0208***		0.0027*
		(0.008)		(0.001)
Home		-0.0573***		-0.0010
		(0.011)		(0.002)
Motivation				
Gnl. Altruism		-0.0023		-0.0036***
		(0.006)		(0.001)
Grp. Altruism		-0.0380***		-0.0039**
		(0.008)		(0.002)
Norms		0.0116**		-0.0002
		(0.005)		(0.001)
Tool		0.0325***		0.0052***
		(0.005)		(0.001)
Satisfaction		-0.0260***		0.0008
		(0.008)		(0.002)
Religious		0.0281***		0.0001
		(0.010)		(0.002)
External		-0.0972***		-0.0043
		(0.012)		(0.003)
Observations	569,149	69,890	569,149	69,890

Note: Significant at the * 10%, ** 5%, and *** 1% levels.

who join either 0 or 1 team. For the individuals who join only 1 team, it may be the case that their team identity is more concentrated and therefore stronger. The results of these two regressions are displayed in table 2.3. The main difference between these results and the ones displayed in table 2.2 is that for both those lenders who do and do not give demographic information, joining a team causes them to increase their lending, and by more than those who join more than 1 team. It is possible that having multiple identities dilutes the importance of each of those identities, and having only 1 will make the effect of joining a team as strong as possible. We leave this for future work.

The results indicate support for hypothesis 10, but only for lenders who do not give Kiva all of the demographic information that Kiva asks for. This could indicate that lending teams increase lender engagement with the Kiva website, having little effect on lenders who are already very active on the site. We next explore whether this or other possibilities explain the data.

2.5.3 Why do people who join teams lend more?

Akerlof and Kranton's (2000) social identity theory, lenders who join teams are influenced by the social norms of those teams. In light of the results in the previous section, these team norms might apply to engagement with the team rather than directly to increased lending. Joining certain teams may require subjects who are not otherwise engaged with Kiva to adopt those teams' norms of being engaged. To examine how this might occur, we examine the forum discussions between team members on Kiva.

Each Kiva lending team has a restricted forum that only members of that team can access. Given the many different teams, we take advantage of teams differences that can be observed from forum conversations to analyze which aspects of teams are associated with increased lending. We examine the forums of a random subset

Table 2.3: Instrumental Variables Regression: Lenders who join only 1 team

	First Stage: Joined Team		Second Stage: Average Loans	
	(1)	(2)	(3)	(4)
Joined Team			0.0712***	0.0316**
			(0.005)	(0.014)
Location	0.0418***	0.0247***		
Similarity	(0.000)	(0.002)		
Constant	0.1047***	0.3889***	0.0384***	0.0401***
	(0.002)	(0.012)	(0.001)	(0.006)
Female		-0.0308***		-0.0076***
		(0.004)		(0.001)
Occupation				
Entertainment		-0.0086		-0.0014
		(0.006)		(0.001)
Business		-0.0027		0.0029***
		(0.006)		(0.001)
IT		0.0778***		0.0021
		(0.007)		(0.001)
Education		0.0035		-0.0054***
		(0.005)		(0.001)
Engineering		0.0455***		0.0021
		(0.010)		(0.002)
Health Care		-0.0175**		0.0014
		(0.008)		(0.001)
Home		-0.0401***		0.0007
		(0.011)		(0.002)
Motivation				
Gnl. Altruism		0.0017		-0.0033***
		(0.006)		(0.001)
Grp. Altruism		-0.0283***		-0.0018
		(0.008)		(0.001)
Norms		0.0086*		-0.0002
		(0.005)		(0.001)
Tool		0.0244***		0.0021**
		(0.005)		(0.001)
Satisfaction		-0.0285***		-0.0008
		(0.008)		(0.001)
Religious		0.0315***		-0.0008
		(0.010)		(0.001)
External		-0.0719***		-0.0017
		(0.012)		(0.002)
Observations	542,019	62,117	542,019	62,117

Note: Significant at the * 10%, ** 5%, and *** 1% levels.

of 2000 teams. We note each team's number of loans, number of lenders, number of posts, proportion of lenders who make a post, number of links to specific loans, and number of plural versus singular pronouns.

The proportion of a team's members who post to the forums will indicate whether teams with actively lending members successfully encourage members to communicate with other team members. The number of links to specific loans measures the extent that the forums allow for coordination towards specific loans. This variable also provide a measure of coordination since posters who link to loans are generally attempting to bring attention to specific loans so that their team members might also loan to them. Finally, the number of plural versus singular pronouns that appear in the forums indicates how strongly lenders feel towards their teams. Teams that exhibit more group vocabulary may lend more for prestige reasons. This will be further analyzed later.

Table 2.4 displays the results of an OLS regression of average lending activity on forum characteristics. Only the number of links to specific loans per lender is associated with significantly different levels of lending activity. An increase of 1 link is associated with 0.006 more loans per lender per day, which is about 1 more loan per lender every 5 months. This is a very small effect. Note, however, that this analysis does not show a causal relationship. For future work, we plan to run a field experiment on Kiva where we will be able to manipulate forum characteristics to observe the causal effects of forum activity on lending. Also note that the data we are using here is sparse. Of the 2000 teams we sample, only 379 have any forum messages, and only 81 have more than 10 forum messages.

Another possibility is that of interteam competition. When a lender joins a team, particularly a highly-ranked team, the team leaderboard provides a level of prestige that the lender may wish to maintain or surpass. We examine whether teams that are close to overtaking another team or being overtaken by another team in terms

Table 2.4: Team Forum Analysis

	Dependent Variable: Number of Loans per Lender per Day
Proportion of Lenders who Post	-0.0001 (0.006)
Links per Lender	0.0064** (0.003)
Plural - Singular Pronouns per Word	0.0070 (0.028)
Constant	0.0659*** (0.004)
Observations	1,511
R^2	0.113

Note: Significant at the * 10%, ** 5%, and *** 1% levels.

of number of loans will attract more loans. First, on a monthly basis, we calculate each team's average rank, number of loans, and loan differences between the teams that are just above and below them in the rankings. For the rank and loan difference variables, we expect the month-to-month change to be more relevant to lending than the absolute values, so we calculate their first differences and lag them by one period. We also wish to examine an interaction term between a team's rank and their loan differences with other teams, so we center the rank and loan difference variables before calculating these interaction terms. This yields the following econometric model:

$$\begin{aligned} loans_{i,t} &= \beta_0 * constant + \beta_1 * loans_{i,t} + \beta_2 * rankchange_{i,t-1} + \\ &\quad \beta_3 * diffabovechange_{i,t-1} + \beta_4 * diffbelowchange_{i,t-1} + \\ &\quad \beta_5 * rankchange_{i,t-1} * diffabovechange_{i,t-1} + \\ &\quad \beta_6 * rankchange_{i,t-1} * diffbelowchange_{i,t-1} + \epsilon_{i,t} \end{aligned}$$

We run this as a fixed-effects regression with i denoting teams and t denoting months. Table 2.5 displays two specificaions of this regression. In column (1), we run this regression on all teams. Note that all of the included regressors are significant at the 1% level. There is a large amount of persistence in the number of loans a team gives in a month, as evidenced by the large coefficient on the lagged number of loans term. The other coefficients are smaller. When a team's rank worsens (i.e. moves farther away from rank 1) by 1, the average team's number of loans only increase slightly. When a team's loan difference to the team above it or below it decreases, the team lends more, indicating both a desire to maintain their current rank against other teams and improving their rank over other teams. Note that the coefficient on the "below" variable is about 5 times greater in absolute value than the coefficient on the "above" variable, showing that teams care more about maintaining their ranks than improving their ranks.

Table 2.5: Fixed-Effects Regression: Team Loans

Dependent Variable: Number of Loans per Month		
	(1)	(2)
Number of Loans	0.8139***	0.8016***
(Lagged)	(0.001)	(0.011)
Rank Change	0.0004***	0.1848**
	(0.000)	(0.087)
Loan Difference Change (Above)	-0.0027***	-0.0009
	(0.000)	(0.003)
Loan Difference Change (Below)	-0.0138***	-0.0130***
	(0.000)	(0.005)
Rank Change ×	0.0032***	-0.0005
Loan Difference Change (Above)	(0.000)	(0.001)
Rank Change ×	0.0004***	0.0184***
Loan Difference Change (Below)	(0.000)	(0.003)
Constant	2.7840***	79.3030***
	(0.018)	(5.005)
Observations	892,880	3,080
R^2	0.638	0.679

Note: Significant at the * 10%, ** 5%, and *** 1% levels.

In column (2), we restrict our attention to teams with an average rank over time of 100 or less. In other words, these are the teams that can be considered "top" teams.

There are some interesting differences between this specification and the one in column (1). First, the coefficient on the "Rank Change" variable is much higher. When the rank of a top team worsens by 1, the team members increase their lending the next month by 0.18 loans. Also, the change in the loan difference between teams and the teams above them in the rankings does not affect next month's loans significantly, while the similar difference to the team below them in the rankings has the same magnitude. This indicates that top teams care about maintaining their ranks much more than improving their ranks. Both when their ranks are threatened or actually worsened, top teams will increase their lending to compensate.

2.6 Conclusion

Kiva gives us an excellent opportunity to explore social identity in a real-world setting. We find that social identity plays an important role in lenders' choice of teams and in their lending behavior.

When lenders choose teams, we find that lenders pay more attention to team rank than to member characteristics. The ranks of the teams that lenders join are higher (i.e. closer to 1) if they are based on past lending history than if they are based on either locational or motivational similarity to the other people in the team. However, it also seems that they do take homophily into account to some extent, as the similarity measures are significantly higher between subjects and their team members than between subjects and non-team members. There are also likely dimensions of homophily that we cannot observe given the data the Kiva elicits from its lenders.

We suspect that lender preference for teams with high rank is largely due to the leaderboard that Kiva uses to display the lending teams. Kiva prominently displays a team's rank and makes teams with the highest rank easy to find and join. A possible way that lenders could choose their teams is to search for teams whose members share their interests or other characteristics. Another possibility is that the lenders limit their attention to the most successful teams before taking their shared characteristics into account. Kiva's team leaderboard could be pushing lenders towards this second option. This could explain our observation that most lending is attributed to a few top teams, with most teams left with little to no lending.

We also find that joining a team increases lending, but only for lenders who do not provide Kiva with demographic information. It is clear from simple analysis that lenders who join teams consistently lend more than those who do not join teams. By employing an instrumental variables analysis, we are able to show that teams are not simply gathering the most active lenders. Rather, lenders who join teams can be expected to make 1.8 more loans per month than their counterparts who do not join teams. If we restrict our attention to lenders who are engaged enough to enter location, occupation, and motivation information into Kiva's database, we find that these lenders who join teams do not make more loans than those who do not join teams. This could suggest that teams are causing lenders to be more engaged in Kiva, and that those who give Kiva demographic information, and are already engaged in Kiva, are not affected by joining a team.

By examining Kiva's team forums, we find that teams whose forum users post more links to specific loans are more active in lending. By obtaining the forum posts for 2000 random teams, we find that teams with more links in their forums lend more than other teams. This result indicates that the Kiva forums might be useful as a coordination device where the lenders of a team can share the information of borrowers they have found on Kiva. This could also be the mechanism by which lenders become more engaged through joining a team, though the observed effect is small. Further work is required on this front.

Finally, lenders in teams do seem to respond, to some extent, to interteam competition. Teams that are closer to being overtaken in the overall number-of-loan rankings lend more, and teams that are closer to overtaking another team also lend more. When we restrict our attention to top teams, we find that they do not react when they are closer to overtaking another team.

Kiva's implementation of lending teams seems to have increased loans. While this situation is good for Kiva, it is unclear whether Kiva could do better with its lending teams system. The current system takes advantage of social norms, but the possibility remains that lending would be higher if lenders would join teams whose motivations and other characteristics were more similar to their own. If Kiva were to collect more demographics from its users, it could perhaps suggest teams for users to join that would take greater advantage of homophily, and hopefully increase loans in the process.

We examine Kiva through the lens of social identity, in particular Akerlof and Kranton's formulation of social identity theory. Our analyses suggest that some Kiva lending teams have a social norm of increased lending, perhaps through increased engagement with the site and the teams. With our limited analyses on homophily, we find that team rank and past performance are more important to lenders than the similiarities between the lenders. However, our results on team competition suggest that team identities do matter, and that with either more demographic data or a closer analysis of the individual teams, we may find that homophily is important on Kiva.

CHAPTER III

Endogenous Groups and Cooperation

3.1 Introduction

Social groups are both pervasive and diverse. Such groups can include large stable demographic segments such as race, gender and religion, or they can include more specific malleable groups such as a school, employer or place of residence. In studying the effects of group membership on individual behavior, Tajfel and Turner (1979a) developed the concept of social identity, which is the part of a person's identity relating to membership in social groups. They first proposed social identity to explain ingroup bias, which is the observed tendency for people to treat members of their own groups better than members of other groups (Tajfel et al., 1971). Social identity theory attempts to explain how and why people join groups, and how they behave as members of those groups. Social identity has also been a consideration in economic analysis, as introduced by Akerlof and Kranton (2000).

In particular, social identity has been shown to affect people's behavior in economic situations. Generally, people in a group adjust their behavior to benefit that group. Tajfel and Turner (1979a) propose that this phenomenon is related to self-esteem. Once someone joins a group, that group's outcomes affect the member's self-esteem. When presented with comparisons to other groups, members either change their own behavior to improve their groups' outcomes or try to leave their groups. Alternatively, Akerlof and Kranton (2000) propose that people in a group internalize that group's behavioral norms, causing disutility to deviations from those norms. The internal process of group identification is a key part of these formulations of social identity theory. A person who identifies with a group ties the group's outcomes to her own utility, making her group membership salient. This process is therefore essential for social identity to change behavior.

Not surprisingly, many organizations use social identity as part of their organizational design. Public broadcasters such as National Publc Radio (NPR) and the Public Broadcasting Service (PBS) frequently run pledge drives to cover their costs. One of the many tactics they employ is to appeal to their listeners' and viewers' regional identities to elicit donations.[1] Another example is the annual blood drive competition between the University of Michigan and Ohio State University called the Blood Battle, started in 1982. Organizers of the Blood Battle have taken advantage of the long football rivalry between the two schools to spur blood donations. Over a three-week period in 2011, over 5000 pints of blood were donated by Blood Battle participants.[2]

Another example comes from an online peer-to-peer microfinance website, Kiva (kiva.org), founded in 2005. Users on Kiva browse through profiles of entrepreneurs in third-world countries and make zero-interest loans to specific entrepreneurs in order to help them fund their businesses. As of March 2012, the users of Kiva have lent over $294 million to entrepreneurs in over 60 countries. Since August 2008, Kiva has offered its users the option of joining lending teams. Each lending team is geared towards a specific group of lenders. In March 2012, the top five lending teams on Kiva, by total amounts loaned, were "Atheists, Agnostics, Skeptics, Freethinkers, Secular Humanists and the Non-Religious," "Kiva Christians," "Team Europe," "milepoint,"

[1] Source: http://www.slate.com/articles/arts/culturebox/2009/03/lets_get_those_phones_ringing.single.html, retrieved on 10/18/2011.

[2] Source: http://bloodbattle.org/, retrieved on 3/16/2012.

and "Australia".[3] Kiva users can freely choose which teams, if any, they join. A user can then assign each of her loans to one of her teams. The teams with the highest amount of total lending are prominently displayed on a leaderboard. In studying lending behavior on Kiva, Liu et al. (2012) find that belonging to a lending team is associated with increased pro-social lending.

The ability of Kiva users to choose their groups could explain this increased lending. People who can choose their groups are likely to strongly identify with those groups, increasing their desire to ensure those groups perform well compared to other groups. In laboratory settings, random assignment of subjects to groups is common. Without further group enhancements, these exogenously determined groups do not usually cause subjects to increase their pro-social behavior. Of course, groups in the real world generally use voluntary membership, so it is important to examine the effects of endogenously determined groups on pro-social behavior. This study is designed to identify and measure these effects.

In many cases, the possible effects of endogenous groups and of group competition are intertwined. On Kiva, as the Atheist and Christian teams are the top two teams in the overall lending rankings, there exists a sense of competition between the teams. A sentiment that exists in both teams' message boards is that they should "beat" the other team in lending. For example, the Atheist team captain has stated: "The whole idea of teams in the Kiva context implies there should be competition," so the view of at least some Kiva users is that teams exist for the purpose of team competition.[4] Using data from Kiva to explore the effects of endogenous groups will not be able to separate those from the effects of group competition.

This study addresses this issue by using a laboratory setting to separately test the effects of endogenous groups and of group competition on pro-social behavior.

[3] Source: http://www.kiva.org/community, retrieved on 3/16/2012.

[4] Source: http://www.kiva.org/team/atheists/messages?msgID=60182#msg_60182, retrieved on 9/20/2011.

The group-contigent social preference model first introduced in Chen and Li (2009) is used to model the situation and analyze the results.

In addition, by allowing for both endogenous groups and group competition, this study lends further insight into economic behavior by examining the influence of social identity. In particular, Tajfel and Turner (1979a) predict that a) a large part of social identity depends on the presence of social comparison, and b) people in unsatisfactory groups will either try to leave those groups or make them better. This study is one of the first, to my knowledge, that tests these hypotheses in an economic setting.

The rest of this paper is structured as follows: Section 3.2 reviews the literature on experiments in social identity, social comparison, and endogenous group formation. Section 3.3 presents the economic environment. Section 3.4 presents the experimental design. Section 3.5 presents the hypotheses. Section 3.6 presents the results of the experiment. Finally, section 3.7 concludes.

3.2 Literature Review

A common theme among experimental studies of social identity using exogenous groups is that categorization into groups is frequently insufficient to increase pro-social behavior toward group members. While studies in social psychology, led by Tajfel et al. (1971), show that subjects in exogenous groups will exhibit ingroup bias when their own payoffs are not affected, later economic studies show that further enhancements to group identification are required when their own payoffs are affected. For example, Eckel and Grossman (2005) find that "just being identified with a team is, alone, insufficient to overcome self-interest" in a public goods game using the Voluntary Contribution Mechanism (VCM). Instead, they find that increased group salience through activities such as group puzzle solving or intergroup tournaments is required to improve public goods provision. Weber (2006) shows that, while large groups in the minimum-effort game do not usually perform well, starting with a small group and slowly adding to the group increases performance. Charness et al. (2007) show that subjects act more in their own group's favor when being observed by that group. Chen and Li (2009) and Chen and Chen (2011) show that pre-game communication between subjects of the same group on an unrelated problem-solving task improves cooperation. In these cases, subjects exhibit cooperation only after their group identification is increased.

An exception occurs with the double-dictator game. When subjects are matched and given money to divide between themselves and others, several studies (e.g. Yamagishi and Kiyonari 2000; Fershtman and Gneezy 2001; Goette et al. 2006) find that subjects in exogenous groups exhibit ingroup bias. Yamagishi and Kiyonari (2000) also find that common knowledge of group membership is necessary for this effect to exist. These results suggest that ingroup bias works through generalized reciprocity among group members. Subjects expect that their own groups will give them more money, and therefore give more money to those group members.

There are several examples of individual-level social comparison being used to spur competition between people. Lin (2011) finds that on Amazon.com's Top Customer Reviewers "New Reviewers" leaderboard, reviewers are concerned with being overtaken by other reviewers and will write more reviews to prevent this. Chen et al. (2010) find that on movielens.org, a movie recommendation site, users will review more movies when given information about how many movies other users have reviewed. In August 2011, the International Association of Athletics Federations (IAAF) ruled that female marathon times would not be counted for records if they were set at mixed-gender events, due to the "benefits of pacesetting by faster male runners".[5] This suggests that the association believes social comparison in running provides enough of an advantage as to be unfair to women who run in events without social comparison.

[5] Source: http://www.nytimes.com/2011/09/22/sports/for-womens-road-records-only-women-only-races-will-count.html, retrieved on 9/25/2011.

This study explores whether and to what extent endogenous groups or group-level social comparison increase group identification. These manipulations are implemented as minimally as possible. While subjects choose which endogenous groups to join, they only interact with the other members of those groups through gameplay. In addition, while social comparison has been shown to be effective in both laboratory (Cason and Mui, 1998; Eckel and Wilson, 2007) and field settings (Frey and Meier, 2004; Shang and Croson, 2009; Chen et al., 2010), subjects are only able to compete through their group outcomes, not through their individual outcomes, making group identification a key factor in the effectiveness of the comparison.[6]

In most experimental studies with endogenous groups, experimenters attempt to find mechanisms that cause subjects to increase their pro-social behavior. One direction this research has explored is using subjects' past behavior or stated preferences to automatically sort them into groups. Page et al. (2005), for instance, sort subjects based on their stated group preferences. In one treatment, termed "Regrouping," subjects are allowed to see other subjects' past contributions to their public goods and to rank each other subject by preference for future partnership. The algorithm then places subjects into groups based on lowest mutual preference ranking. Gunnthorsdottir et al. (2007) run a similar experiment, but without the ability to rank others. Instead, their algorithm sorts subjects into groups based solely on the contributions to a public good in the past period. The highest contributors are placed in a group, then the next highest contributors are placed in another group, and so on. Both of these methods of automatic sorting result in an increase in contributions over random matching. Finally, Coricelli et al. (2004) find that allowing for unidirectional choice (the winner of an auction chooses her partner) rather than a bidirectional mechanism (subjects allocate money towards being partnered with other members, with the highest mutual allocations being paired) results in higher contributions. This study

[6]For a more comprehensive review of the literature on social comparison, see Chen et al. (2010).

suggests that choice is an important factor in increasing contributions.

Another related stream of research examines the effect of restricting group entry or exit, and of allowing subjects to expel group members. Cinyabuguma et al. (2005) and Croson et al. (2006) show that the threat of expulsion from a group can increase contributions to a public good. In each study, subjects are able to vote out group members. This threat is strong enough to reduce the number of free riders. A similar result is obtained by Brosig et al. (2005) when they allow subjects to communicate before voting someone out of the group. Charness and Yang (2010) and Ahn et al. (2008) examine the effect of restricting group entry through the use of majority vote for potential new members. Ahn et al. (2008) also examine the case where exiting a group requires the consent of a majority of the other members of a group. Both studies yields some success in increasing contributions, though the results are not significant in Ahn et al. (2008).

Within this stream of research, Ehrhart and Keser (1999) examine a voluntary contribution mechanism in a setting similar to that used in Tiebout (1956). Subjects are first randomly placed in one of three groups. They then play a round of a public goods game. After each play, subjects are told the payoffs achieved in each group and allowed to change their groups for a small fee. Interestingly, Ehrhart and Keser (1999) find that high-contributing subjects “flee” from free riders while free riders chase high contributors.

To my knowledge, only Charness et al. (2011) examine endogenous groups from a social identity perspective. Subjects can have both a team identity (teams solve a group problem) and an endowment identity (some people receive twice the endowment of others), and they play a public goods game with endogenous groups. When both identities are present, they find that the subjects choose to group themselves by their endowment rather than their team identity. Also, subjects with a team identity contribute more to the public good regardless of the identities of the subjects with whom they play the public goods game.

While all of these studies lend insight into group membership and individual behavior, only one explores group choice from a social identity perspective. The research on social identity has focused on exogenously-determined groups while the research on endogenously-determined groups has not focused on the effects of group choice. This study, by combining these two streams of research, allows for the direct evaluation of both social identity theory and of organizational designs.

3.3 Economic Environment

In this study, subjects first play the minimum-effort game, and then make a series of dictator game and other-other allocation decisions. During the minimum-effort game, groups will form and stabilize. The dictator game and other-other allocation decisions are used to measure the level of ingroup bias in the formed groups. While the dictator game measures how much subjects care about their own payoffs relative to those of others, the other-other allocation measures how much subjects care about their group members' payoffs relative to others' payoffs.

The situation is modeled using the group-contingent social preference model of Chen and Li (2009). This model uses an other-regarding parameter, α^g, which measures the degree to which a subject values the payoffs of others, as indicated by:

$$u_i(\pi_i, \Pi^g_{-i}) = \alpha^g \Pi^g_{-i} + (1 - \alpha^g)\pi_i, \tag{3.1}$$

where $g \in \{In, Out\}$ refers to those in/out of subject i's group, respectively. In the model, a subject holds a different value of α^g for each group. A higher α^g represents a greater level of other-regarding preferences towards members of group g.

Subjects play the minimum-effort game T times. This game is characterized by the following payoff function:

$$\pi_i(e_1, \ldots, e_n) = a \cdot \min\{e_1, \ldots, e_n\} - c \cdot e_i + b, \tag{3.2}$$

where n is the number of subjects playing the minimum-effort game in a group; $a, c,$ and b are real, nonnegative constants; and $e_i \in [\underline{e}, \overline{e}]$ is the effort provided by subject i.

The minimum-effort game is a game with multiple Pareto-ranked Nash equilibria. These occur when all players provie the same level of effort. When a logit equilibrium is derived for the game with social preferences taken into account, Chen and Chen (2011) show that increasing α^g increases equilibrium effort in the sense of first-order stochastic dominance. As long as $a > c$, increased equilibrium effort results in increased individual payoffs, utility, and group efficiency.

The minimum-effort game is also a potential game (Monderer and Shapley, 1996). This fact has been shown to be useful in predicting which equilibrium subjects will choose. In particular, Chen and Chen (2011) show that there is a cut-off cost, $c^* = a/n$, where n is the number of subjects in each group. If $c < c^*$, subjects are expected to converge to the most efficient Nash equilibrium, which occurs when all subjects choose the maximum possible effort. On the other hand, if $c > c^*$, the subjects are expected to converge to the least efficient Nash equilibrium, which occurs when all subjects choose the minimum possible effort.

To study the effect of endogenous groups, this experiment allows some subjects to change their groups after several plays of the minimum-effort game. To aid in comparing treatments, group sizes are kept constant. That is, there are N subjects total in a session, with G groups; each group has exactly $n = N/G$ subjects at all times. Thus, the assignment of subjects to groups is a house allocation matching problem. The algorithm used in this study is the Random Serial Dictatorship with Squatting Rights mechanism (RSDS), as described in Abdulkadiroğlu and Sönmez (2010). In the experiment, after the first period, subjects choose whether to "squat"(i.e., stay in their current groups). Subjects who do not squat then choose the group they want to join that is not yet full. While this algorithm is not necessarily ex-post individually rational or ex-post Pareto efficient, it is strategy-proof, conditional on not squatting. It is also a straightforward mechanism that is easy for the subjects to understand.

During the game, subjects receive one of two types of feedback. In the first, after every play of the minimum-effort game, a subject is shown her own payoff, the minimum effort within her own group, and the average payoff within her own group. In the other, a subject is shown her own payoff, the minimum effort in all groups, and the average payoff in all groups. This feedback may influence how subjects' other-regarding preferences change. When a subject is shown only her own group's results, only her other-regarding preferences towards her own group may change. Alternatively, when a subject is shown all groups' results, all of her other-regarding preferences might be affected. Furthermore, when a subject is shown all groups' results, she makes a social comparison between the groups, which could increase α^{In} and decrease α^{Out}.

3.4 Experimental Design and Procedure

Each session of the experiment consists of nine subjects, divided into three equal-sized groups. Subjects experience four stages. In the first three stages (Group Formation Stage, Group Stabilization Stage, and Fixed Groups Stage), subjects play the three-person minimum effort game. The Evaluation Stage consists of one period in which subjects make three dictator game decisions and three other-other allocations. The stages are displayed in Figure 3.1.

The experiment follows a 2×2 design. Along the first dimension, a subject's ability to choose and change her group is varied. In the "Exogenous Groups" treatment, subjects are randomly re-assigned to groups every two periods for the Group Formation and Group Stabilization Stages. In the "Endogenous Groups" treatment,

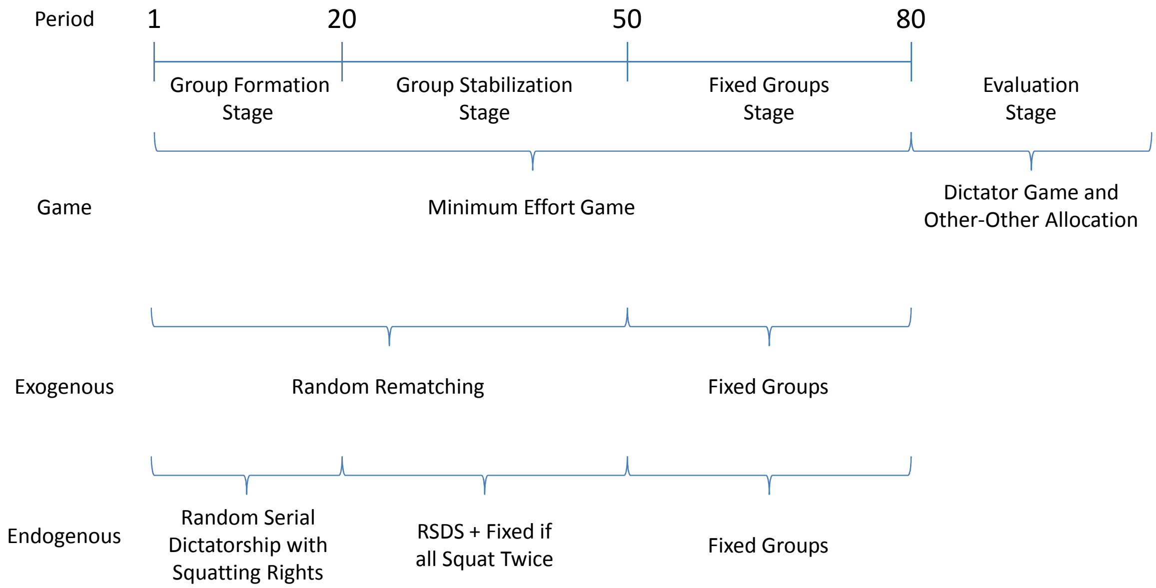

Figure 3.1: Experiment Stages

subjects have some control over their group membership.

In the **Group Formation Stage** (periods 1-20), beginning in the third period, each subject is asked whether she want to "squat" (i.e. stay in her current group) after every two periods of play. Each subject who does not squat is asked to state her preference ranking over the groups. Each subject is given a random priority and placed, in priority order, into her most preferred unfilled group. This mechanism is known as the random serial dictatorship with squatting mechanism (RSDS).

In the **Group Stabilization Stage** (periods 21-50), subjects follow the same procedure as in the Group Formation Stage, with one change. If every member of a group chooses to squat two times in a row, then that group becomes fixed and no one is allowed to enter or exit that group for the remainder of the experiment. Subjects in such fixed groups are no longer asked whether they want to squat or to rank the groups. All other groups, including those in the Exogenous treatment, also become fixed for the **Fixed Groups Stage** (periods 51-80). By the time the Evaluation Stage occurs, subjects have been in the same group for at least 30 periods. This scheme is summarized in Figure 3.1.

Along the second dimension, the feedback scheme is varied. In the "Own-Group Feedback" treatment, a subject is shown her own payoff as well as her group's minimum effort and average payoff after every period in the first three stages of the experiment. In the "All-Groups Feedback" treatment, a subject is shown her own payoff as well as the minimum effort and average payoff of all three groups, with the groups ranked by minimum effort, then by average payoff. This is to facilitate social comparison in an effort to further increase α^{In} while decreasing α^{Out}. A summary of the experimental design is displayed in Table 3.1.

Table 3.1: Experimental Design

Treatment		Number of Subjects
Exogenous	Own-Group Feedback	9×4
	All-Groups Feedback	9×4
Endogenous	Own-Group Feedback	9×4
	All-Groups Feedback	9×4

The minimum-effort game the subjects play follows that used in Goeree and Holt (2005) and Chen and Chen (2011). Subjects can choose any effort level between 110 and 170 in increments of 0.01. These numbers are chosen to prevent any obvious focal points, and to allow for a wider range of possible effort levels.[7] The cost parameter for the game is 0.50, which is well above the cut-off cost parameter for a three-person minimum effort game without other-regarding preferences taken into account ($c^* = \frac{1}{3}$). Therefore, according to the theory of potential games (Monderer and Shapley, 1996), subjects with no groups are expected to converge to the lowest effort equilibrium. All payoffs are reported to the subjects in points, which are converted to dollars at a rate of 650 points per dollar after the experiment is concluded. Given the parameters

[7]Subjects do, however, choose integer effort levels 97.9% of the time and effort levels divisible by 10 78.6% of the time. The standard minimum effort game used in experimental economics allows subjects to choose an integer effort level from 1 to 7 (Van Huyck et al., 1990).

used, a subject earns at least 25 points and at most 85 points per period, for a total of at least 2000 points and at most 6800 points across the first three stages of the experiment.

In the Evaluation stage, subjects make six simultaneous decisions. Three of these are dictator game allocations, and three are other-other allocations. For each dictator game allocation, a subject decides how to divide 100 points between herself and a respective random member of each group. That is, a subject makes a dictator allocation to a random member of Team A, a random member of Team B, and a random member of Team C. For the other-other allocations, a subject decides how to divide 100 points between a random member of one team and random member of another team (neither of whom can be the subject herself). The three other-other allocations cover the three possible pairs between the three teams. All decisions are implemented, with each subject receiving three dictator allocations (one from her own group), and six other-other allocations (two from her own group). Along with her own dictator decisions, a subject can earn up to 1200 points from this stage of the experiment.

To illustrate, the following outlines the procedure for a session of the Endogenous, All-Groups Feedback treatment. First, subjects are read the instructions out loud. Subjects then complete a quiz regarding the calculation of payoffs in the minimum-effort game to test their understanding of the game. Then, they begin the Group Formation Stage. For the first period, subjects state their preference rankings over the three groups, which are labeled "Team A," "Team B," and "Team C," respectively. Subjects are then assigned to a group, make effort decisions for the minimum-effort game, and are given feedback on all three groups. For period 2, subjects play the minimum-effort game (with the same group) and receive feedback. Starting with period 3, subjects repeat the procedure for the first two periods. However, at the beginning of every odd period, subjects first decide whether to squat. This continues in the Group Stabilization Stage (periods 21-50), except that if every member of a group decides to squat for two consecutive odd periods, the group becomes fixed. For every period of the Fixed Groups Stage (periods 51-80), subjects play the minimum-effort game and receive feedback without changing groups. Finally, for the Evaluation Stage, subjects make three dictator and and three other-other decisions. At the end of the experiment, subjects fill out a demographic survey and receive their earnings from the experiment.

As shown in Table 3.1, four independent sessions were run for each treatment, with nine subjects in each session. A total of 16 computerized sessions with 144 subjects were run in the Behavioral and Experimental Economics Lab at the University of Michigan in June of 2011. One pilot session with nine subjects was run before these 16 sessions. Nearly all subjects were drawn from the student body of the University of Michigan, with each subject participating in only one session.[8] Sessions lasted at most one hour each. All sessions were run using z-Tree (Fischbacher, 2007). Subjects were paid a $5 show-up fee in addition to their earnings for the experiment, which were rounded up to the next nearest dollar. The average payment was $15.41 (minimum $13; maximum $17). The experimental instructions are included in Appendix A and the survey with response statistics is included in Appendix B.

3.5 Hypotheses

This section presents the hypotheses regarding behavior in the various stages and treatments of the experiment. The general null hypothesis is that behavior does not differ between any pair of treatments. The alternative hypotheses are presented here. I first examine the minimum-effort game, which spans the Formation, Stabilization and Fixed Groups stages.

Each hypothesis is based on the other-regarding parameter, α^g. In the N (En-

[8] One subject was a staff member at the University of Michigan.

dogenous) treatment, subjects have some degree of control over the α^g that is relevant to them since they are able to change their groups and will tend towards higher levels of α^g. This will persist once the groups have become stable. As a result, subjects will give more effort in the N treatment than in the X (Exogenous) treatment.

Hypothesis 11 (Effort: Endogenous vs. Exogenous Groups)**.** *Subjects in the Endogenous (N) treatment will give more effort in the minimum-effort game than subjects in the Exogenous (X) treatment.* $(e_N > e_X)$

In the A (All-Groups Feedback) treatment, subjects are affected by social comparison. As described in the Economic Environment section, social comparison may result in both an increase in α^{In} and a decrease in α^{Out} for a given subject. Therefore, it is expected that subjects will give more effort in the A treatment than in the O (Own-Group Feedback) treatment.

Hypothesis 12 (Effort: All-Groups vs. Own-Group Feedback)**.** *Subjects in the All-Groups Feedback (A) treatment will give more effort in the minimum-effort game than subjects in the Own-Group Feedback (O) treatment.* $(e_A > e_O)$

Next, each subject in the N treatment periodically decides whether or not to squat, i.e. stay in their current groups. Subjects in the A treatment observe their groups' ranks while those in the O treatment do not. When a subject knows that her own group is a top-ranked group, she should be more likely to squat than a subject who does not know this information. Similarly, a subject who knows that her own group is a bottom-ranked group should be less likely to squat. Given these opposing effects, it is unclear how subjects in the A treatment will differ in their squatting decision from those in the O treatment.

Next, each subject in the N treatment who does not squat makes group choice decisions. The following hypothesis compares the ranks of the first-choice groups between subjects in the A and O treatments. This hypothesis relates to the bandwagon effect, whereby subjects want to join those groups that are the most successful. Since the subjects in the O treatment do not know the ranks of the different groups, their first-choice groups are not expected to have ranks as high as those of subjects in the A treatment.

Hypothesis 13 (Group Choice: All-Groups vs. Own-Group Feedback)**.** *Subjects in the All-Groups Feedback (A) treatment who choose not to squat will list higher-ranked groups as their first choice than subjects in the Own-Group Feedback (O) treatment who choose not to squat.*[9]

I now consider the dictator games and the other-other allocations. For these allocations, the group-contingent other-regarding preferences developed by the subjects over the first three stages of the experiment are expected to persist through the Evaluation Stage.

The dictator game used in this study does not pair subjects like those used in the experiments described in the Literature Review section, removing direct reciprocity. However, generalized reciprocity may still exist. Because subjects in the N treatment are able to choose their groups, $\alpha_N^{In} > \alpha_N^{Out}$. Since this is not true in the X treatment, $\alpha_N^{In} > \alpha_X > \alpha_N^{Out}$. Therefore, $\alpha_N^{In} > \alpha_X^{In}$ and $\alpha_X^{Out} > \alpha_N^{Out}$, resulting in the following hypothesis.

Hypothesis 14 (Dictator Game: Endogenous vs. Exogenous Groups)**.** *Subjects in the Endogenous treatment will allocate more points to their ingroup matches and fewer to their outgroup matches in the dictator game than subjects in the Exogenous treatment.* ($d_N^{In} > d_X^{In}$, $d_N^{Out} < d_X^{Out}$)

Similarly, for the All-Groups and Own-Group Feedback treatments, $\alpha_A^{In} > \alpha_A^{Out}$ due to the social comparison effect. With no social comparison in the O treatment, $\alpha_A^{In} > \alpha_O > \alpha_A^{Out}$. Therefore, $\alpha_A^{In} > \alpha_O^{In}$ and $\alpha_O^{Out} > \alpha_A^{Out}$, resulting in the following hypothesis.

[9] In this study, "highest rank" refers to rank 1 and the "lowest rank" refers to rank 3.

Hypothesis 15 (Dictator Game: All-Groups vs. Own-Group Feedback)**.** *Subjects in the All-Groups Feedback treatment will allocate more points to their ingroup matches and fewer to their outgroup matches in the dictator game than subjects in the Own-Group Feedback treatment.* ($d_A^{In} > d_O^{In}$, $d_A^{Out} < d_O^{Out}$)

The other-other allocations are based on social psychology research. Specifically, social psychologists consistently find ingroup bias in other-other allocations (e.g. Turner 1979). That is, subjects give more to members of their own group than to others. Along with the dictator game, the other-other allocation is used in this study to measure the magnitude of the group effect, if one exists.

Using the inequalities used to justify Hypotheses 14-15, the following equations can be established:

$$\alpha_N^{In} - \alpha_N^{Out} > \alpha_X^{In} - \alpha_X^{Out}$$

and

$$\alpha_A^{In} - \alpha_A^{Out} > \alpha_O^{In} - \alpha_O^{Out}.$$

These result in the next two hypotheses.

Hypothesis 16 (Other-Other Allocation: Endogenous vs. Exogenous Groups)**.** *In the other-other allocation with one ingroup member and one outgroup member, subjects in the Endogenous treatment will allocate relatively more points to ingroup members than to outgroup members when compared to subjects in the Exogenous treatment.* ($y_N > y_X$)

Hypothesis 17 (Other-Other Allocation: All-Groups vs. Own-Group Feedback)**.** *In the other-other allocation with one ingroup member and one outgroup member, subjects in the All-Groups Feedback treatment will allocate relatively more points to ingroup members than to outgroup members when compared to subjects in the Own-Group Feedback treatment.* ($y_A > y_O$)

Note that Hypotheses 14-17 do not compare intrasubject decisions. That is, all comparisons are between subjects rather than within subjects, reducing the possibility of an experimenter demand effect. By simultaneously viewing decisions related to ingroup and outgroup members, subjects may feel inclined to give more to ingroup members and less to outgroup members. However, since Hypotheses 14-17 examine the same decisions between different subjects, any observed differences will not be due to experimenter demand effects.

3.6 Results

In this section, I examine the effort and group-choice decisions in the minimum-effort game, as well as the dictator and other-other allocation decisions. To understand the underlying preference structure which gives rise to these decisions, I calibrate a learning model that incorporates social preferences.

3.6.1 Minimum-Effort Game

First, I examine the effect of the different treatments on the effort provided by subjects in the minimum-effort game. Figure 3.2 shows the maximum, median, and minimum effort in each period, with each line representing the average of the relevant statistic for a given treatment. For instance, in the minimum effort graph, the solid black line represents the minimum effort in each group for each period, averaged over the four sessions of the Endogenous, Own-Group Feedback treatment.

In Figure 3.2, note that a clear separation occurs between the four treatments, with the ordering of these treatments generally following the order predicted in Hypotheses 11 and 12. That is, $e_N > e_X$ and $e_A > e_O$.

Table 3.2 compares the minimum effort in the different treatments by stage. In each case, the relevant hypothesis is listed along with the result of a permutation test, with each session treated as one observation.

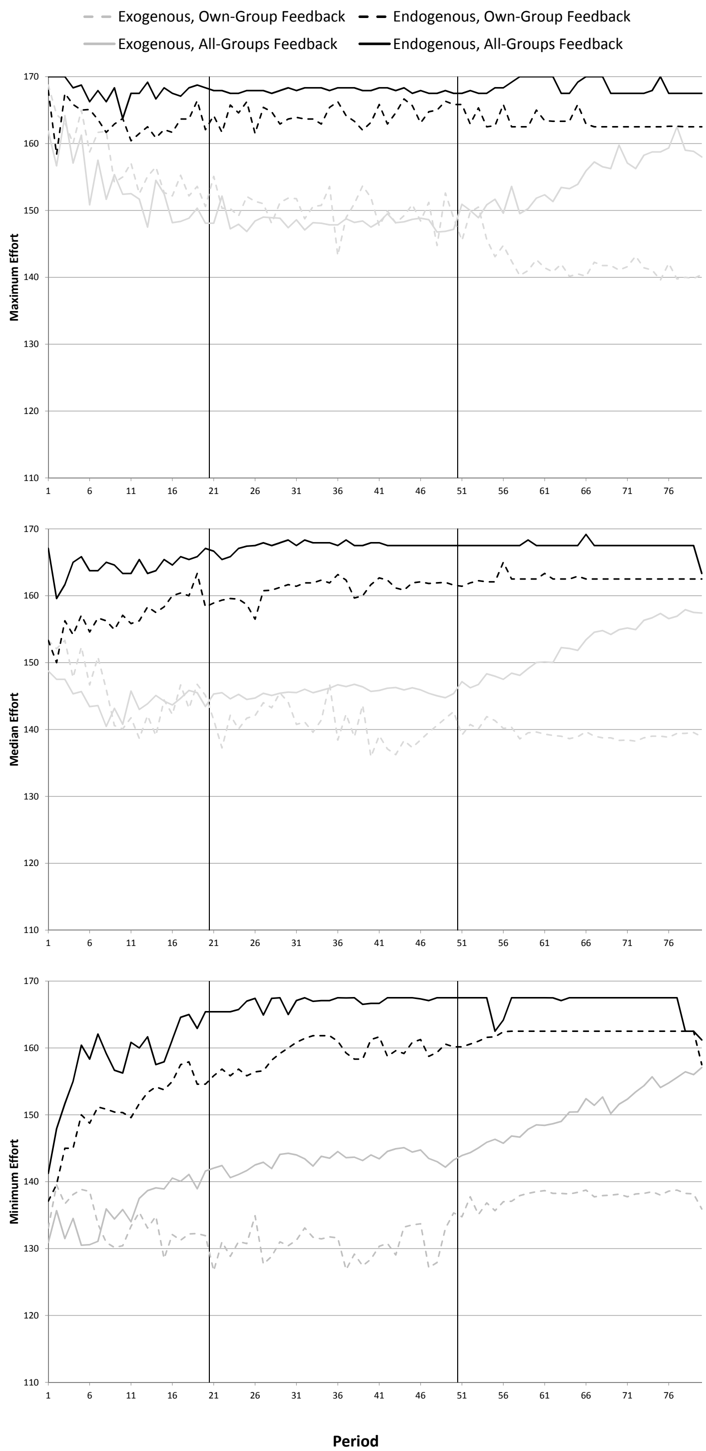

Figure 3.2: The respective figures represent the maximum (Top), median (Middle) and minimum (Bottom) effort in each treatment, with each period averaged by session. The two vertical lines denote the periods that separate the Formation and Stabilization stages, and the Stabilization and Fixed Group stages.

Table 3.2: Comparison of Average Effort between Treatments

Test/ Comparison	Stage	Alternative Hypothesis	Permutation Test p-values (1-sided)
Endogeneity NA v. XA	Formation	$H_{11}: e_N > e_X$	0.014**
	Stabilization	$H_{11}: e_N > e_X$	0.100*
	Fixed Groups	$H_{11}: e_N > e_X$	0.129
	Overall	$H_{11}: e_N > e_X$	0.071*
Endogeneity NO v. XO	Formation	$H_{11}: e_N > e_X$	0.043**
	Stabilization	$H_{11}: e_N > e_X$	0.014**
	Fixed Groups	$H_{11}: e_N > e_X$	0.043**
	Overall	$H_{11}: e_N > e_X$	0.029**
Feedback NA v. NO	Formation	$H_{12}: e_A > e_O$	0.057*
	Stabilization	$H_{12}: e_A > e_O$	0.171
	Fixed Groups	$H_{12}: e_A > e_O$	0.443
	Overall	$H_{12}: e_A > e_O$	0.129
Feedback XA v. XO	Formation	$H_{12}: e_A > e_O$	0.386
	Stabilization	$H_{12}: e_A > e_O$	0.200
	Fixed Groups	$H_{12}: e_A > e_O$	0.214
	Overall	$H_{12}: e_A > e_O$	0.257
Both NA v. XO	Formation	$H_{11}: e_N > e_X$ or $H_{12}: e_A > e_O$	0.014**
	Stabilization	$H_{11}: e_N > e_X$ or $H_{12}: e_A > e_O$	0.014**
	Fixed Groups	$H_{11}: e_N > e_X$ or $H_{12}: e_A > e_O$	0.014**
	Overall	$H_{11}: e_N > e_X$ or $H_{12}: e_A > e_O$	0.014**
Both NO v. XA	Formation	$H_{11}: e_N > e_X$ or $H_{12'}: e_A < e_O$	0.114
	Stabilization	$H_{11}: e_N > e_X$ or $H_{12'}: e_A < e_O$	0.143
	Fixed Groups	$H_{11}: e_N > e_X$ or $H_{12'}: e_A < e_O$	0.129
	Overall	$H_{11}: e_N > e_X$ or $H_{12'}: e_A < e_O$	0.143

Notes: (1) Significant at the: * 10%, ** 5%, and *** 1% levels.
(2) Each session is treated as one observation.

Using these permutation tests, the following results regarding the minimum-effort game are evident.

Result 6 (Effort: Endogenous vs. Exogenous Groups)**.** *Subjects in the Endogenous (N) treatment give more effort in the minimum-effort game than subjects in the Exogenous (X) treatment.*

Support. *Table 3.2 shows that both permutation tests comparing e_N to e_X, regardless of stage (Overall), find that e_N is significantly greater than e_X at the 10% level. The results are significant at the 5% level for the Own-Group Feedback treatment in all stages, and for the All-Groups Feedback treatment in the Formation stage. Also, the result is significant at the 5% level in all three stages when the $e_N > e_X$ is tested simultaneously with $e_A > e_O$ (i.e. $NAvs.XO$).*

The results show that subjects who choose their groups also choose higher efforts. This effect diminishes in the All-Groups Feedback treatment, but this is due to increasing effort in the Exogenous, All-Groups Feedback treatment rather than decreasing effort in the Endogenous, All-Groups Feedback treatment. This could be due to a boundary effect, since most groups' minimum effort in the Endogenous, All-Groups Feedback treatment is 170 (82.1% of groups overall; 87.4% after period 21). If they could have, these subjects may have given more effort, increasing the minimum effort in the NA treatment. These results provide support for Hypothesis 11 ($e_N > e_X$).

Result 7 (Effort: All-Groups vs. Own-Group Feedback)**.** *Subjects in the All-Groups Feedback (A) treatment do not give more effort in the minimum-effort game than subjects in the Own-Group Feedback (O) treatment.*

Support. *Table 3.2 shows that both permuatation tests comparing e_A to e_O, regardless of stage (Overall), find no significant difference. The result is significant only at the 10% level in the Formation stage for the Endogenous treatment. As before, the result is* *significant at the 5% level in all three stages when the* $e_N > e_X$ *is tested simultaneously with* $e_A > e_O$.

When comparing the All-Groups Feedback treatment to the Own-Group Feedback treatment, the only significant difference at the 5% level occurs when the permutation tests simultaneously test for a difference between the Endogenous and Exogenous treatments. While the results in Figure 3.2 indicate that providing subjects with feedback on all groups increases effort, this increase is not significant. Thus, the results do not support Hypothesis 12 ($e_A > e_O$).

Another interesting result is the increase in effort that occurs in the Exogenous, All-Groups Feedback treatment once the groups become fixed. The fact that this effect is not present in the Exogenous, Own-Groups Feedback treatment suggests that this is not a repeated-games effect. Rather, once subjects identify with their groups sufficiently, group-level social comparison is effective in increasing effort. The fact that this occurs as soon as the exogenous groups are fixed suggests that a minimal level of identification is necessary for group-level social comparison to have an effect on behavior.

To further test these hypotheses, I run a random-effects regression. Table 3.3 displays four specifications of this regression, one for each stage and one for all stages combined. Each specification controls for subject gender (which has been shown in the past to have a significant effect on behavior, e.g. Croson et al. 2008), whether the subject is employed full time, and the natural log of the period (to control for learning effects).

This regression supports the results found using the permutation tests. Specifically, subjects in the Endogenous treatment give 18.77 more units of effort than subjects in the Exogenous treatment while subjects in the All-Groups Feedback treatment give 6.33 more units of effort than subjects in the Own-Group Feedback treatment. The differences between the Endogenous and Exogenous treatments are significant at the

Table 3.3: Random Effects Regression: Treatment Effects on Effort

Dependent Variable: Effort				
	(1) Formation	(2) Stabilization	(3) Fixed Groups	(4) Overall
Endogenous	14.94***	21.31***	18.79***	18.77***
	(4.761)	(7.135)	(6.839)	(6.084)
All-Groups Feedback	3.15	5.74	9.04	6.33
	(4.729)	(7.104)	(6.897)	(6.083)
Female	-5.48***	-3.62	-2.53	-3.68
	(1.890)	(2.808)	(2.822)	(2.357)
Employed Full Time	-2.21	-1.46	1.15	-0.67
	(2.067)	(3.220)	(3.745)	(2.863)
Log(Period)	-0.07	2.18	11.31	2.55
	(3.264)	(4.223)	(12.244)	(2.571)
Constant	146.81***	138.97***	122.38***	139.85***
	(4.591)	(8.238)	(23.515)	(5.767)
Observations	2880	4320	4320	11520
R^2	0.1855	0.3017	0.2424	0.2452

Notes: (1) Standard errors are adjusted for clustering at the session level.
(2) Significant at the: * 10%, ** 5%, and *** 1% levels.

1% level. Also, females give 5.48 fewer units of effort than males in the Formation stage.[10]

In the minimum-effort game, allowing subjects to choose their groups significantly increases their effort. Social comparison alone is not shown to have an effect on effort. However, when social comparison is combined with endogenous group choice, effort is increased beyond that which group choice achieves (i.e. e_{NA} is weakly greater than e_{NO} in the Formation stage).

Now, in the Endogenous treatment, subjects state their preference rankings over the different groups but are not always assigned to their most preferred groups. Among these subjects, I examine how the realization of the group assignment affects the chosen effort. Table 3.4 displays the results of this regression, again by stage and overall. The three dummy variables "Squatted, "Joined Second Choice Group" and "Joined Third Choice Group" denote how much the subjects prefer the group they

[10] This result, in terms of direction, matches that found in Chen and Chen (2011).

joined. The excluded dummy variable is "Joined First Choice Group." While it would seem that the subjects who chose to squat are equivalent to those who joined their first choice group, there are strategic considerations in the squatting decision that make comparing the two sets of subjects problematic, so they are treated differently in this analysis.

Table 3.4: Random Effects Regression: Group Assignment on Effort

Dependent Variable: Effort			
	(1) Formation	(2) Stabilization	(3) Overall
All-Groups Feedback	6.67**	10.32*	8.17**
	(2.780)	(5.516)	(3.440)
Squatted	-4.18	4.78***	-0.82
	(3.316)	(1.510)	(2.409)
Joined Second Choice Group	-6.24	0.95	-3.21
	(4.451)	(0.899)	(3.409)
Joined Third Choice Group	-10.87**	2.74	-6.08
	(4.990)	(2.859)	(3.942)
Female	-3.07	-0.12	-2.34
	(3.133)	(2.959)	(3.111)
Employed Full Time	1.84	2.42	1.82
	(2.387)	(4.639)	(2.653)
Log(Period)	4.34*	-19.97**	2.44
	(2.449)	(9.085)	(3.997)
Constant	159.76***	177.72***	156.93***
	(4.065)	(13.960)	(3.053)
Observations	648	267	915
R^2	0.1482	0.2314	0.1164

Notes: (1) Standard errors are adjusted for clustering at the session level.
(2) Significant at the: * 10%, ** 5%, and *** 1% levels.

This regression shows some interesting group-assignment effects. In the Formation stage, subjects who are placed into their least favorite group give the least effort, with others give similar efforts. However, in the Stabilization stage, the squatters give more effort than the rest of the subjects. One possible explanation for this observation is that most subjects in the Stabilization stage (87.5%) stabilize (i.e. all members squat twice). Once a group stabilizes, their choices are no longer reflected in this analysis, so the Stabilization stage specification of this regression captures mainly the behavior of subjects whose groups have not yet become fixed. The difference between the specifications captures the behaviors of different sets of subjects.

Given this, I focus on the Formation stage specification. The most likely reason that subjects who are placed in their least favorite groups give less effort is that these subjects tried to but were unable to change groups. This would make their lowered effort a rational response to their group members' expected actions. Out of 140 instances where subjects were assigned to their third choice groups, their current groups were listed as their third choice groups 135 times. So, a large majority of the subjects who were assigned to their least favorite groups had recently been members of those groups and wished to leave. The realization of the group assignment negatively affects effort levels when subjects dislike their current groups and are not able to leave those groups.

3.6.2 Group Choice

Next, I examine the squatting and group-choice decisions made by subjects in the Endogenous treatment. Table 3.5 shows a probit regression, reporting marginal effects, with the dependent variable of whether or not the subjects decided to squat. The independent variables include the treatment, whether the subject's current group is the top-ranked group, and whether the subject's current group is the same as her original group.

The treatment variable, which shows whether the subject sees intergroup comparison information, has a weakly negative coefficient, with subjects in the All-Groups Feedback treatment 13% less likely to squat than those in the Own-Group Feedback treatment. This indicates that subjects who know their groups' current ranks are s-lightly more willing to explore the other groups than those that do not. While subjects whose groups are top-ranked squat more than 85% of the time in both treatments,

Table 3.5: Probit Regression: Treatment Effects on Squatting

Dependent Variable: Choose to Squat?	
All-Groups Feedback	-0.13*
	(0.070)
In Top-Ranked Group	0.26***
	(0.050)
Interaction of All-Groups Feedback	0.14***
and In Top-Ranked Group	(0.049)
In Original Group	0.26***
	(0.019)
Interaction of All-Groups Feedback	-0.20
and In Original Group	(0.161)
Observations	915
Pseudo-R^2	0.2510

Notes: (1) Standard errors are adjusted for clustering at the session level.
(2) Significant at the: * 10%, ** 5%, and *** 1% levels.

other subjects squat more than 70% of the time in the Own-Group Feedback treatment and about 40% of the time in the All-Groups Feedback treatment. Subjects in non-top-ranked groups would like to try joining the top-ranked group, but subjects who do not know their group ranks are not willing to risk leaving a top-ranked group.

The coefficients on the other terms in this probit regression provide evidence for how subjects with different amounts of information make their squatting decisions. Subjects in the O treatment 26% more likely to squat if they are in top-ranked groups than if they are not, even when they do not know this. Also, they are 26% more likely to squat if their current groups match their period-one groups. On the other hand, subjects in the A treatment are 14% more likely to squat than those in the O treatment if they are in top-ranked groups, but less likely to squat than those in the O treatment if they are in their original groups. A test that the "In Original Group" and "Interaction of All-Groups Feedback and In Original Group" coefficients does not reject the null that they sum to zero ($p = 0.742$). While subjects in both treatments are more likely to squat if they are in top-ranked groups, subjects in the O treatment are also more likely to squat in their original groups. When given less information about the ranks of the groups, subjects exhibit an unwillingness to change groups.

Next, if a subject chooses not to squat, I examine the rank of the group she lists as her first choice. Specifically, I look at whether or not she lists a top-ranked group as her first choice, comparing the probability that this occurs between the A treatment and the O treatment.

To perform this analysis, I define a variable that takes ties between the groups into account. Sometimes, the groups tie in both minimum effort and average payoff. In this case, the tied groups are all displayed to the subjects as having the highest rank. In the case that the tie occurs between the top two teams, the chance that a subject chooses a top-ranked team as her first choice is twice that in the other cases, assuming that the subject is choosing her first-choice team randomly.[11] This new variable gives this case half the weight of the other cases, and a zero if subject did not choose a top-ranked group as her first choice.

Result 8 (Group Choice: All-Groups vs. Own-Group Feedback)**.** *If they choose not to squat, subjects in the All-Groups Feedback treatment (A) list a top-ranked group as their first choices slightly more often than subjects in the Own-Group Feedback (O) treatment.*

Support. *The average of this new variable in treatments O and A are 1.464 and 1.048, respectively. The two distributions are significantly different at the 10% level but not the 5% level (Mann-Whitney* $U = 213.5$, $n_O = 28$, $n_A = 21$, *p-value = 0.0668 one-tailed).*

Note that the adjusted frequency of choosing a top-ranked team as the first choice for the two treatments is significantly different only at the 10% level. Even though the subjects in the O treatment are not told the group ranks, by using the minimum

[11] I drop all cases where all teams tie from this analysis.

effort of their own groups as guides, these subjects nearly match the subjects in the A treatment in listing the highest-ranked groups as their top choices. Thus, these results only weakly support Hypothesis 13.

These results support the social identity theory prediction of Tajfel and Turner (1979a). That is, subjects choose to remain with groups whose outcomes improve, regardless of whether they see the outcomes of the other groups. However, when presented with other groups' performances, subjects in first-place groups squat more often than subjects in second or third-place groups. Unfavorable social comparison causes subjects in lower-ranked groups to leave those groups.

3.6.3 Dictator and Other-Other Allocations

Next, I examine the treatment effects in the dictator game and in the other-other allocations. Figure 3.3 shows the average number of points allocated by a subject to her match(es) in each treatment. For the dictator game, the results of the two outgroup match decisions are pooled. So, if a subject is a member of group A, then her dictator allocations to members of groups B and C are averaged. Similarly, for the other-other allocations, the two decisions where the subject's own group is involved are pooled.

First, I examine whether subjects give more to ingroup members than to outgroup members. As discussed in the Hypotheses section, this is a within-subject comparison and may be affected by the experimenter demand effect. This is presented here as a verification of findings from social psychology regarding ingroup bias (e.g. Tajfel et al. 1971). In this study, a permutation test that subjects give their ingroup matches more than their outgroup matches in the dictator game yields a p-value of 0.019, and a permutation test that subjects give more to members of their own groups than to others in the other-other allocation yields a p-value less than 0.001. These results pool the different treatments but treat each session as one observation.

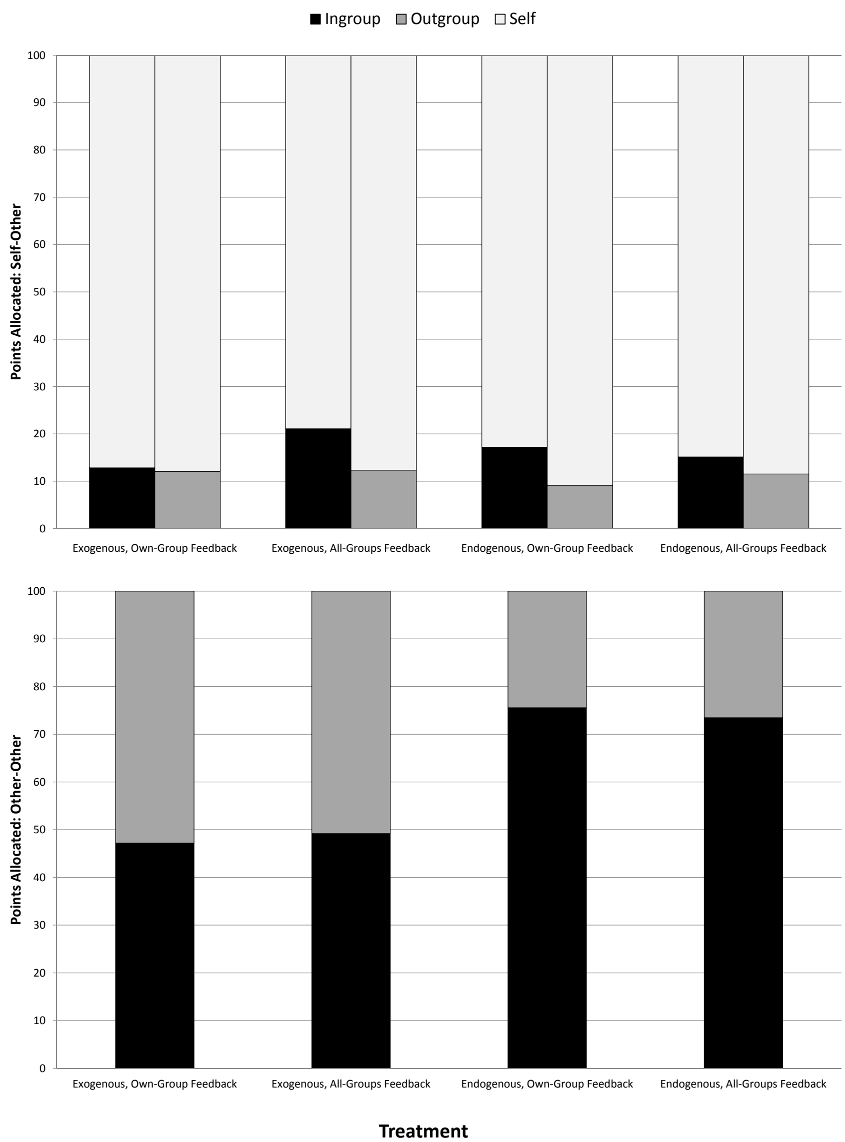

Figure 3.3: Allocations in the dictator game (Top) and other-other allocations (Bottom). For the dictator game, the number of points allocated to the two outgroups are pooled. For the other-other allocation, the ingroup-outgroup allocations are pooled over the two outgroups.

Next, I examine the treatment effects in the dictator game and the other-other allocations. Table 3.6 provides the comparison results for the different treatments. As in Table 3.2, the relevant hypothesis is listed along with the results of the permutation test. These lead to the following results.

Result 9 (Dictator Game). *In all treatments, subjects allocate the same number of points to an ingroup match in the dictator game. Likewise, in all treatments, subjects allocate the same number of points to an outgroup match in the dictator game.*

Support. *Of the twelve tests performed on Hypotheses 14 and 15, none yields results significant at the 5% level. Subjects generally keep the overwhelming majority of points for themselves.*

While ingroup bias exists in the dictator game, the degree of this bias does not change between treatments. One possible explanation for this observation is that behavior in the dictator game is resistant to weaker manipulations of social identity such as group choice and social comparison. Much stronger manipulations, such as the one used in Goette et al. (2006), have been shown to increase ingroup bias, but the manipulations used in this study do not.[12] Thus, I do not find support for Hypotheses 14 and 15.

Result 10 (Other-Other Allocation, Ingroup: Endogenous vs. Exogenous). *Subjects in the Endogenous (N) treatment allocate more points to ingroup members than do subjects in the Exogenous (X) treatment during the other-other allocation with an ingroup member and an outgroup member.*

Support. *Table 3.6 shows that all four tests performed on Hypothesis 16 ($y_N > y_X$) find that y_N is significantly greater than y_X at the 5% level.*

[12] Goette et al. (2006) find higher allocations when subjects are randomly assigned to Swiss military platoons in which members interact with each other for several weeks.

Table 3.6: Comparison of Allocations between Treatments

Test/ Comparison	Game	Recipient	Alternative Hypothesis	Permutation Test p-values (1-sided)
Endogeneity NA v. XA	Dictator	*In*	$H_{14a} : d_N^{In} > d_X^{In}$	0.9143
		Out	$H_{14b} : d_N^{Out} < d_X^{Out}$	0.4571
	Other-Other	*In*	$H_{16} : y_N > y_X$	0.0143**
Endogeneity NO v. XO	Dictator	*In*	$H_{14a} : d_N^{In} > d_X^{In}$	0.2429
		Out	$H_{14b} : d_N^{Out} < d_X^{Out}$	0.2571
	Other-Other	*In*	$H_{16} : y_N > y_X$	0.0143**
Feedback NA v. NO	Dictator	*In*	$H_{15a} : d_A^{In} > d_O^{In}$	0.6571
		Out	$H_{15b} : d_A^{Out} < d_O^{Out}$	0.6143
	Other-Other	*In*	$H_{17} : y_A > y_O$	0.5429
Feedback XA v. XO	Dictator	*In*	$H_{15a} : d_A^{In} > d_O^{In}$	0.0571*
		Out	$H_{15b} : d_A^{Out} < d_O^{Out}$	0.5429
	Other-Other	*In*	$H_{17} : y_A > y_O$	0.4143
Both NA v. XO	Dictator	*In*	$H_{14a} : d_N^{In} > d_X^{In}$ or $H_{15a} : d_A^{In} > d_O^{In}$	0.3857
		Out	$H_{14b} : d_N^{Out} < d_X^{Out}$ or $H_{15b} : d_A^{Out} < d_O^{Out}$	0.4714
	Other-Other	*In*	$H_{16} : y_N > y_X$ or $H_{17} : y_A > y_O$	0.0143**
Both NO v. XA	Dictator	*In*	$H_{14a} : d_N^{In} > d_X^{In}$ or $H_{15a'} : d_A^{In} < d_O^{In}$	0.8000
		Out	$H_{14b} : d_N^{Out} < d_X^{Out}$ or $H_{15b'} : d_A^{Out} > d_O^{Out}$	0.2429
	Other-Other	*In*	$H_{16} : y_N > y_X$ or $H_{17'} : y_A < y_O$	0.0143**

Notes: 1) Significant at the: * 10%, ** 5%, and *** 1% levels.
2) Each session is treated as one observation.

For the other-other allocation, the effect of endogenous groups is quite pronounced. While subjects give about 50 out of 100 points to both recipients in the X treatment, subjects in the N treatment give about 75 out of 100 points to members of their own group. These results provide support for Hypothesis 16.

Result 11 (Other-Other Allocation, Ingroup: All-Groups Feedback vs. Own-Group Feedback)**.** *Subjects in the All-Groups Feedback (A) treatment do not allocate more points to ingroup members than do subjects in the Own-Group Feedback (O) treatment during the other-other allocation with an ingroup member and an outgroup member.*

Support. *Of the four tests performed on Hypothesis 17 ($y_A > y_O$), only one finds a significant result at the 5% level, but this test also simultaneously tests whether $y_N > y_X$.*

As in the minimum-effort game, there is no significant effect of social comparison on other-other allocation behavior. The results in Figure 3.3 show that subjects in the X treatment give about the same to an ingroup match as to an outgroup match regardless of whether they are in the O or A treatment. Similarly, subjects in the N treatment do not differ their behavior towards the ingroup and the outgroup between treatments O or A. Thus, the results do not support Hypothesis 17.

To further explore these results, I run a series of OLS regressions on the subjects' allocation decisions. Table 3.7 shows five different regressions. The first two specifications relate to the subjects' dictator game allocations, and the last three relate to their other-other allocations. In the first specification, the dependent variable is the number of points allocated to an ingroup member. Here, it is clear that there are no differences between the treatments. However, there is a significant gender effect, with females giving 11.69 more points to their ingroup matches than males. Similarly, the second specification has the dependent variable of the number of points allocated to an outgroup member. There are also no treatment differences here either, and a weaker gender effect. These findings are in line with those in Result 9. Notice in particular that the payoffs that the subjects receive in the minimum-effort game have no effect on these dictator game decisions.

The dependent variable for the last three specifications is the number of points allocated to an ingroup member in an other-other allocation. Specification (3) shows the overall regression, which includes the decisions of all subjects. Specifications (4) and (5) split the groups into those that performed above and below the median in the minimum-effort game, in terms of total group payoff. In the overall regression, both the payoff in the minimum-effort game and the treatment have significant effects on the allocation to the ingroup member. In particular, subjects give above 12.60 more points to the ingroup member when they are in the Endogenous treatment. This is a substantial effect, since they given 100 points for this decision. When the regressions are performed separately on the high and low-performing groups, we can see that groups that performed more poorly rely on their previous payoffs to make their allocation decision, while those that performed well do not. Rather, the entire effect from being in the Endogenous treatment comes from those who are in groups that performed well. The subjects only cared about the groups they chose when those groups performed well. When they did not, the subjects adopt a reciprocity strategy, and did not care whether or not they were allowed to choose their groups.

Overall, the use of endogenous groups creates an effect that carries over to a one-shot game such as the other-other allocation. While it does not significantly affect a subject's attitude towards her ingroup in relation to herself, it does increase how much she cares about her ingroup in relation to the outgroup. On the other hand, group-level social comparison does not have this same effect. While there is a general increase in effort once the groups become fixed in the Exogenous, Own-Group Feedback treatment, this increase does not result in a greater degree of ingroup bias. Similarly, while subjects in the Endogenous, All-Groups Feedback treatment give

Table 3.7: OLS Regression: Treatment Effects on Allocations

Dependent	Dictator: *In*	Dictator: *Out*	Other-Other: *In*		
Variables:	(1)	(2)	(3) Overall	(4) Below Med	(5) Above Med
Endogenous	-0.20	-1.29	12.60**	2.26	17.21**
	(5.284)	(6.536)	(5.717)	(10.263)	(7.482)
All-Groups	2.60	1.18	-5.51	-1.50	-4.14
Feedback	(4.101)	(4.408)	(5.255)	(9.519)	(8.751)
Payoff in	-0.00	-0.00	0.01***	0.02***	0.04
Min-Effort	(0.004)	(0.004)	(0.004)	(0.006)	(0.024)
Female	11.79**	6.93*	-4.57	0.87	-6.43
	(4.086)	(3.807)	(4.859)	(6.318)	(8.380)
Employed	2.60	4.66	-6.85	-2.88	-13.33
Full Time	(9.869)	(9.713)	(8.775)	(9.924)	(11.332)
Constant	10.70	9.86	-17.02	-56.43*	-194.18
	(18.713)	(20.743)	(23.009)	(29.841)	(161.678)
Observations	144	144	144	72	72
R^2	0.056	0.030	0.245	0.167	0.242

Notes: (1) Standard errors are adjusted for clustering at the session level.
(2) Significant at the: * 10%, ** 5%, and *** 1% levels.

more effort than those in the Endogenous, Own-Group Feedback treatment, they do not allocate more points to their group members in the other-other allocation. While endogenous groups seem to increase subjects' group identification, group-level social comparison requires increased group identification to work, and does not itself contribute to that increase.

3.6.4 Calibration

To understand the underlying preferences that generate the behavior observed in the Fixed Groups Stage, I calibrate a stochastic fictitious-play learning model that incorporates social preferences. In this model, subjects hold beliefs regarding their opponents' play, and adjust these beliefs based on their opponents' actual play in previous periods. The player then, to some extent, best responds to these beliefs.

This procedure is similar to that used in Chen and Chen (2011). Specifically, this model includes three free parameters: a sensitivity parameter, λ, that denotes the extent to which a subject best responds to her beliefs; a discount factor, δ, that denotes how heavily her opponents' past behavior affects her beliefs; and an other-regarding parameter, α^g, from the group-contingent social preferences model.[13]

Player i's belief regarding the minimum of her matches' effort levels x_j in every period t depends on a weight function $w_i^t(x_j)$. This weight function assigns to each possible minimum effort the number of times she has observed that minimum effort in the past, discounted. Minimum efforts that have been observed in the past are more likely. The initial value of this weight function, $w_i^1(x_j)$, is left unspecified by

[13] λ ranges from 0 to ∞, with 0 indicating purely random play, and ∞ denoting best responding to the subject's belief regarding her matches' effort level. δ ranges from 0 to 1, and α^g ranges from -1 to 1.

the model. The updating rule is the following:

$$w_i^{t+1}(x_j) = \delta \cdot w_i^t(x_j) + \begin{cases} 1 & \text{if } x_j = x^t \\ 0 & \text{otherwise,} \end{cases} \tag{3.3}$$

where x^t is the minimum effort exerted by player i's matches in period t. δ is the discount factor on previous observations as described above. Player i's beliefs in period t regarding her matches' minimum efforts are:

$$\mu_i^t(x_j) = \frac{w_i^t(x_j)}{\sum_{x_k} w_i^t(x_k)}. \tag{3.4}$$

Next, the subject will best respond to these beliefs. Player i's expected utility for playing a strategy x_i is:

$$\overline{u}_i^t(x_i) = \frac{1}{(\overline{x} - \underline{x})} \sum_{x_j} [u_i(x_i, x_j) \cdot \mu_i^t(x_j)], \tag{3.5}$$

where $u_i(x_i, x_j)$, in this case, is derived from a combination of equations 3.1 and 3.2.

Using this expected utility, player i randomly chooses an effort level x_i with the distribution:

$$f_i^t(x_i) = \frac{\exp[\lambda \cdot \overline{u}_i^t(x_i)]}{\sum_{x_k} \exp[\lambda \cdot \overline{u}_i^t(x_k)]}, \tag{3.6}$$

To calibrate the model, a grid search is performed over the three parameters. Each set of parameters is evaluatied using the quadratic scoring rule described in Selten (1998). In any given round, let $f_{ij} = (f_{i1}, \ldots, f_{iK})$ be the predicted probability distribution over player i's strategies, where K is the number of strategies available to the players, and let $a_{ij} = (a_{i1}, \ldots, a_{iK})$ be the observed relative frequency distribution over player i's strategies, where $a_{ij} = 1$ if player i chooses action j, and zero otherwise. This score, $S_i(f)$, is calculated by $S_i(f) = 1 - \sum_{j=1}^{K} (a_{ij} - f_{ij})^2$. The estimates for the parameters are the values of λ, δ, and α^g that give the highest summed score for each subject.

I restrict my analysis to the Fixed Group Stage to obtain values of α^{In}. Unlike in Chen and Chen (2011), this analysis is performed on the subject level, giving each subject a different value of α^{In}. First, I perform a coarse calibration of the λ and δ parameters, which yields $\lambda = 4.5$ and $\delta = 0.7$ for all subjects. Next, I perform a finer calibration of the α^{In} parameter at the subject level, finding the value of α^{In} from -1 to 1 which best fits each subject's behavior. For each treatment, the results of this procedure are displayed in the form of cumulative distribution functions in Figure 3.4.

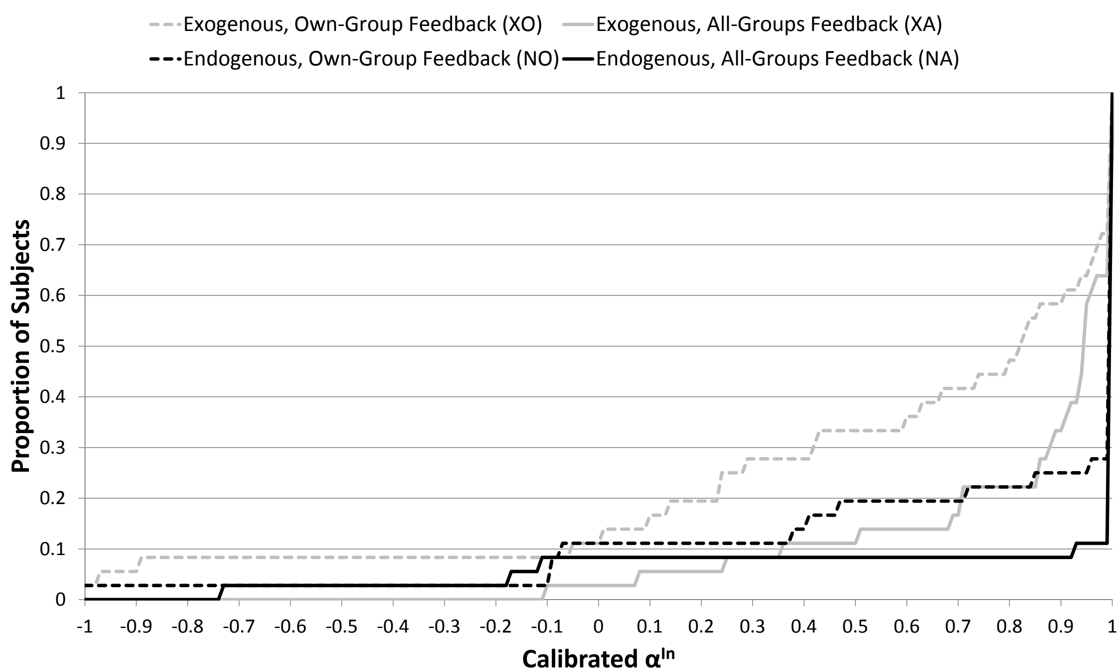

Figure 3.4: Cumulative Distribution Functions of the α^g parameter.

Of the 31 periods in the Fixed Group Stage, the first 21 (periods 50-70) are used to calibrate the learning model and the last 10 (periods 71-80) are used as a holdout sample for validation. Once the parameters are calibrated for an individual, a score is calculated for that subject's validation sample using those parameters. A test of the equality of the normalized scores for the calibration and validation sets does not reject the null of equality at the 5% level (Wilcoxon signed-rank test, $p = 0.412$), indicating that there is no overfitting.

Next, I test whether these distributions are equal across treatments, using a Kolmogorov-Smirnov test. The results of this test are shown in Table 3.8. These results show that only the NO and NA distributions are equal at the 5% level. Otherwise, the distributions follow the expected pattern, with the other-regarding parameter values of the NA and NO treatment being greater than the values of the XA and XO treatments. Endogenous groups seem to increase subjects' other-regarding preferences, while group-level social comparison has that effect once groups are fixed.

Table 3.8: Kolmogorov-Smirnov Tests of α^{In} Distributions

Kolmogorov-Smirnov Test		
Comparison	Statistic	p-value
NA v. XA	0.5278	0.000***
NO v. XO	0.4444	0.001***
NA v. NO	0.1667	0.615
XA v. XO	0.3333	0.022**
NA v. XO	0.6111	0.000***
NO v. XA	0.3611	0.010***

Note: Sig. at: ** 5% and *** 1% levels.

3.7 Conclusion

In this study, I evaluate the relative effectiveness of endogenous groups and group-level social comparison on pro-social behavior. I find that endogenous groups effectively increase subjects' group identification. On the other hand, group-level social comparison only works when subjects already identify with their groups, but does not itself increase group identification.

This study shows that subjects who choose their groups exhibit higher levels of other-regarding preferences than those who are randomly placed into groups. Social identity theory suggests that intergroup comparison is essential in creating social identity, but this result suggests that a subject's self-esteem is affected by her group's outcome even when there is no outside comparison. It makes sense that, by allowing subjects to choose their groups, they will care more about those groups. As mentioned, most previous experimental studies have generally found that simply placing subjects into groups is not enough to increase pro-social behavior. In these studies, some degree of socialization is required to induce a group effect, whether it be direct interaction with other subjects, communication with other subjects, or observation of other subjects' actions. Interestingly, in the present study, even when all of these techniques are put aside, subjects who can choose their groups act like subjects who have gone through one of these exercises. In that sense, allowing subjects to choose their groups acts like a socialization device. Of course, people in real groups, such as the ones on Kiva, both choose their groups and go through at least one of the other socialization techniques (communication in the case of Kiva), thus strengthening this result. Overall, this study provides evidence that groups with voluntary membership work naturally, without any outside intervention.

This social comparison result is in line with previous studies regarding individual-level social comparison. Unlike the case with groups and group-level social comparison, individuals will automatically care about comparisons to others and will likely adjust their behavior to improve those comparisons. Group-level social comparison simply requires the extra step of group identification to acheive the same effect.

This result provides justification for both the Blood Battle and Kiva's lending teams design. The Blood Battle takes advantage of an existing rivalry between the two schools, making group competition between the schools effective, even for a task unrelated to the original rivalry. By introducing lending teams, Kiva is able to take advantage of social identity to increase lending. Kiva users, by choosing their team membership, strongly tie their own self-esteem to the outcomes of these teams. This makes the team leaderboard that Kiva provides effective in encouraging users to increase their own lending. For Kiva and the entrepreneurs that receive these loans, this design is weakly better than not having lending teams.

One possible next step in this line of research is to examine the interaction of endogenous groups and further socialization. There are some situations, such as the dictator game in this study, that have been shown not to be affected by standard manipulations of social identity. By including endogenous groups, stronger group effects may be possible. Another avenue for future research is an examination of group labels. All groups in the real world have names, descriptions, and reasons for existence. This study, and nearly all experimental studies of social identity, suppress these labels. Very likely, groups with labels that have actual meaning will be able to create strong group identities. The strength of these identities should be compared to those created by endogenous groups, and by the socialization techniques used in other experiments.

APPENDICES

APPENDIX A

Theory

Proof of Proposition 1: Maximizing Equation (1.7) gives us a new threshold marginal cost value, which is a function of the group-contingent other-regarding parameter α_i^g,

$$c^*(n, \{\alpha_i^g\}_{i=1}^n) = \frac{1}{n - \sum_{i=1}^n \alpha_i^g}. \tag{A.1}$$

When $\alpha_i^I > \alpha_i^N > \alpha_i^O$, $\forall i$, the corresponding threshold marginal cost is as follows:

$$c^*(n, \{\alpha_i^I\}_{i=1}^n) > c^*(n, \{\alpha_i^N\}_{i=1}^n) > c^*(n, \{\alpha_i^O\}_{i=1}^n).$$

Furthermore, a more salient group identity increases α_i^I, which leads to an increase in the threshold marginal cost, $c^*(n, \{\alpha_i^I\}_{i=1}^n)$.

Proof of Proposition 3: Based on the standard assumption of the logit model that payoffs are subject to unobserved shocks from a double-exponential distribution, player i's probability density is an exponential function of the expected utility, $u_i^e(x)$,

$$f_i(x) = \frac{\exp(\lambda u_i^e(x))}{\int_{\underline{x}}^{\bar{x}} \exp(\lambda u_i^e(s))ds}, \qquad i = 1, \cdots, n, \tag{A.2}$$

where $\lambda > 0$ is the inverse noise parameter and higher values correspond to less noise.

Let $F_i(x)$ be player i's corresponding effort distribution. For player i, let $G_i(x) \equiv 1 - \prod_{k \neq i}(1 - F_k(x))$ be the distribution of the minimum of the $n-1$ other effort levels. Thus, player i's expected utility from choosing effort level x is:

$$u_i^e(x) = \int_{\underline{x}}^{x} y g_i(y) dy + x(1 - G_i(x)) - c[(1-\alpha_i)x + \alpha_i \int_{\underline{x}}^{\bar{x}} y dF_i(y)], \tag{A.3}$$

where the first term on the right side is the benefit when another player's effort is below player i's own effort, the second term is the benefit when player i determines the minimum effort, and the last term is the cost of effort weighted by player i's own effort and the average effort of others. The first and very last term of the right side of (A.3) can be integrated by parts to obtain:

$$u_i^e(x) = \int_{\underline{x}}^{x} \prod_{k \neq i}(1 - F_k(y)) dy - c(1-\alpha_i)x + c\alpha_i \int_{\underline{x}}^{\bar{x}} F(y)dy + \underline{x} - c\alpha_i \bar{x}. \tag{A.4}$$

Differentiating both sides of (A.2) with respect to x and using the derivative of the expected utility in (A.4), we obtain:

$$\begin{aligned} f_i'(x) &= \lambda f_i(x) \frac{du_i^e(x)}{dx} \\ &= \lambda f_i(x)[\prod_{k \neq i}(1 - F_k(x)) - c(1-\alpha_i)], \qquad i = 1, \cdots n. \end{aligned} \tag{A.5}$$

Using symmetry (i.e., dropping subscripts), further assuming $\alpha_i = \alpha$ for all i, and integrating both sides of (A.5), we obtain:

$$\int_{\underline{x}}^{x} f'(s)ds = \lambda \int_{\underline{x}}^{x} f'(s)[1 - F(s)]^{n-1} ds - c(1-\alpha)\lambda \int_{\underline{x}}^{x} f(s)ds.$$

Simplifying both sides, we obtain the first-order differential equation for the equilibrium effort distribution:

$$f(x) = f(\underline{x}) + \frac{\lambda}{n}[1 - (1 - F(x))^n] - c(1-\alpha)\lambda F(x).$$

The proofs of Propositions 4 and 5 use similar structure and techniques as those of the corresponding Propositions 4 and 5 in Anderson et al. (2001), with the marginal cost of effort, c, replaced by $c(1-\alpha)$. We present them here for completeness.

Proof of Proposition 4: Let the other regarding parameters be $\alpha_1 < \alpha_2$, and let $F_1(x)$ and $F_2(x)$ denote the corresponding equilibrium effort distributions. We want to show that $F_1(x) > F_2(x)$ for all interior x.

Suppose $F_1(x) = F_2(x)$ on some interval of x values. Then the first two derivatives of these functions must equal on the interval, which violates (A.5). Therefore, the distribution functions can only be equal, or cross, at isolated points. At any crossing, $F_1(x) = F_2(x) \equiv F$. From (A.2), the difference in slopes at the crossing is:

$$f_1(x) - f_2(x) = f_1(\underline{x}) - f_2(\underline{x}) - \lambda c(\alpha_2 - \alpha_1)F, \tag{A.6}$$

which is decreasing in F, and hence is also decreasing in x. It follows that there can be at most two crossings, with the sign of the right-hand side nonnegative at the first crossing and nonpositive at the second. Since the distribution functions cross at $\underline{x}$ and $\bar{x}$, these are the only crossings. The right-hand side of (A.6) is positive at $x = \underline{x}$ or negative at $x = \bar{x}$, so $F_1(x) > F_2(x)$ for all interior x. This implies that an increase in α results in a distribution of effort that first-degree stochastically dominates that associated with a smaller α.

Proof of Proposition 5: First, consider the case $c < c^*$, or $cn(1-\alpha) < 1$. We have to show that $F(x) = 0$ for all $x < \bar{x}$. Suppose not, and $F(x) > 0$ for $x \in (x_a, x_b)$.

From (1.9), we have:

$$\begin{aligned} f(x) &= f(\underline{x}) + \frac{\lambda}{n}[1-(1-F)^n] - c(1-\alpha)\lambda F \\ &= f(\underline{x}) + \frac{\lambda}{n}[1-(1-F)^n - cn(1-\alpha)F] \\ &> \frac{\lambda}{n}[1-(1-F)^n - F] \\ &= \frac{\lambda}{n}(1-F)[1-(1-F)^{n-1}]. \end{aligned}$$

Since density cannot diverge on an interval, $F(\cdot)$ must be zero on any open interval. Therefore, $F(x) = 0$ for $x < \bar{x}$.

Next, consider the case $c < c^*$, or $cn(1-\alpha) > 1$. In this case, we have to prove that $F(x) = 1$ for all $x > 0$. Suppose not, and $F(x) < 1$ for $x \in (x_a, x_b)$. From (1.9), we have:

$$\begin{aligned} f(\bar{x}) &= f(\underline{x}) + \frac{\lambda}{n}[1-(1-F(\bar{x}))^n] - c(1-\alpha)\lambda F(\bar{x}) \\ &= f(\underline{x}) + \frac{\lambda}{n} - c(1-\alpha)\lambda \\ &= f(\underline{x}) + \frac{\lambda}{n}[1-cn(1-\alpha)], \end{aligned}$$

which enables us to rewrite (1.9) as:

$$\begin{aligned} f(x) &= f(\bar{x}) - \frac{\lambda}{n}[1-cn(1-\alpha)] + \frac{\lambda}{n}[1-(1-F)^n] - c(1-\alpha)\lambda F \\ &= f(\bar{x}) + \frac{\lambda}{n}[cn(1-\alpha)(1-F) - (1-F)^n] \\ &> \frac{\lambda}{n}(1-F)[1-(1-F)^{n-1}]. \end{aligned}$$

Again, since density cannot diverge on an interval, $F(\cdot)$ must be one on any open interval. Therefore, $F(x) = 1$ for $x > 0$.

Finally, consider the case $c = c^*$, or $cn(1-\alpha) = 1$. In this case, (1.9) becomes:

$$\begin{aligned} f(x) &= f(\underline{x}) + \frac{\lambda}{n}[1-(1-F)^n] - c(1-\alpha)\lambda F \\ &= f(\underline{x}) + \frac{\lambda}{n}[1-(1-F)^n - F] \\ &= f(\underline{x}) + \frac{\lambda}{n}(1-F)[1-(1-F)^{n-1}]. \end{aligned}$$

This equation implies that the density diverges to infinity as $\lambda \to +\infty$, when $F(x) \neq 0$ or 1. Hence, $F(\cdot)$ jumps from 0 to 1 at the mode M. The above equation implies that $f(\underline{x}) = f(\bar{x})$, so the density is finite at the boundaries and the mode is an interior point. Using symmetry, we can rewrite (A.5) as $f^{'}(x) = \lambda f(x)[(1-F(x))^{n-1} - c(1-\alpha)] = \lambda f(x)[(1-F(x))^{n-1} - 1/n]$, or $\frac{f^{'}(x)}{\lambda f(x)} = (1-F(x))^{n-1} - 1/n$. Integrating both sides from $\underline{x}$ to $\bar{x}$ yields $\frac{1}{\lambda}\ln(f(\bar{x})/f(\underline{x})) = M - (\bar{x} - \underline{x})/n$, since $1-F$ equals one to the left of M and zero to the right of M. The left side is zero since $f(\bar{x}) = f(\underline{x})$, so $M = (\bar{x} - \underline{x})/n$.

Effort and Efficiency Benchmarks:

We use the equilibrium distribution described in Equation (1.9) to compute the expected effort and efficiency for different values of α. For each distribution, we assume that $\lambda = 0.125$, the value estimated by Goeree and Holt (2005). Summary statistics of this distribution for various values of α are included in Table A.1.

This table shows that the expected efficiency depends nonmonotonically on the exact level of α. As α increases from -1, the expected efficiency decreases until α reaches 0, then increases until α reaches 1. Given the above definition of efficiency, this behavior is expected. That is, at low values of α, subjects mostly give low effort. This results in a medium level of efficiency. At high values of α, subjects give high

Table A.1: Theoretical Distributions

	Effort		
α	μ	σ	Efficiency
-1.0	116.49	5.86	0.563
-0.8	117.40	6.59	0.558
-0.6	118.61	7.50	0.553
-0.4	120.30	8.69	0.546
-0.2	122.79	10.23	0.539
0.0	126.77	12.18	0.533
0.2	133.54	14.21	0.541
0.4	143.37	14.66	0.598
0.6	151.37	12.89	0.684
0.8	156.10	10.83	0.751
1.0	158.99	9.16	0.797

effort, resulting in a high level of efficiency. The lowest level of efficiency should be achieved when subjects giving low effort are paired with subjects giving high effort. This occurs more frequently when α is not extreme.

APPENDIX B

Experimental Instructions

(Potential of Social Identity)

We present the experimental instructions for the Enhanced Ingroup treatment. Instructions for other treatments are similar and can be found on the second author's website.

Economic Decision Making Experiment: Part 1 Instructions

This is an experiment in decision-making. The amount of money you earn will depend upon the decisions you make and on the decisions other people make. Your earnings are given in tokens. This experiment has 2 parts and 12 participants. Your total earnings will be the sum of your payoffs in each part. At the end of the experiment you will be paid IN CASH based on the exchange rate

$1 = 350 tokens.

In addition, you will be paid $5 for participation. Everyone will be paid in private and you are under no obligation to tell others how much you earn.

Please do not communicate with each other during the experiment unless asked to do so. If you have a question, feel free to raise your hand, and an experimenter will come to help you.

Before the experiment started everyone drew an envelope which contained either a Green or a Red slip. You have been assigned to the Green group if you received a Green slip, and the Red group if you received a Red slip. There are 6 people in each group. Your group assignment will remain the same throughout the experiment. That is, if you drew a Green slip, you will be in the Green group for the rest of the experiment, and if you drew a Red slip, you will be in the Red group for the rest of the experiment.

In Part 1 everyone will be shown 5 pairs of paintings by two artists. You will have 5 minutes to study these paintings. Then you will be asked to answer questions about two other paintings. Each correct answer will bring you 350 additional tokens. You may get help from or help other members in your own group while answering the questions.

After Part 1 has finished, we will give you instructions for the next part of the experiment.

Economic Decision Making Experiment: Part 2 Instructions

The next part of the experiment consists of 50 periods. In each period, you will be randomly matched with 1 other person in the room. If you are a member of the Green group, your match will always be a member of the Green group, and if you are a member of the Red group, your match will always be a member of the Red group.

You will be reminded every period of your own group and of your match's group. Your earnings for this part of the experiment depend on your choices as well as the choices of the people you are matched with.

Every period, each person will choose an effort level between 110.00 and 170.00. You will earn a number of tokens equal to the minimum effort level chosen by you and the person you are matched with, minus the cost of your own effort, which is 0.75 times your own effort choice. This is captured by the equation:

Payoff (Tokens) = Minimum Effort - 0.75*Your Effort

Note that the minimum effort here refers to the minimum of the effort levels chosen by you and your match. Refer to the handout for some examples. Note that there may be some case in which you earn a negative payoff. If your final payoff is negative, we will deduct that amount from your participation fee.

We will show you a running tally of the number of tokens you have earned from this part of the experiment, and after 50 rounds, we will add your earnings from Part 1 to this total and convert your total earnings into a dollar amount based on the exchange rate. We will also show you a list of your past effort choices and payoffs, as well as your matches' past effort choices and payoffs.

When you are ready to begin Part 2 of the experiment, please click OK.

APPENDIX C

Postexperimental Survey

(Potential of Social Identity)

Please answer the following survey questions. Your answers will be used for this study only. Individual data will not be exposed.

1. What is your age? (*Mean 21.37, Std Dev 3.27, Median 21, Min 18, Max 40*)

2. What is your gender? (*Male 48.53%, Female 51.47%*)

3. Which of the following best describes your racial or ethnic background? (*Asian 38.73%, Black 6.37%, Caucasian 42.16%, Hispanic 3.43%, Native American 0.49%, Multiracial 4.41%, Other 4.41%*)

4. In what country or region were you primarily raised as a child? (*US/Canada 74.51%, Africa 0.00%, Asia 23.53%, Australia 0.49%, Europe 0.98%, Latin America 0.00%, Middle East 0.49%*)

5. What is your marital status? (*Never Married 96.08%, Currently Married 3.43%, Previously Married 0.49%*)

6. How would you best describe your employment status? (*Employed Full Time 5.88%, Employed Part Time 38.24%, Not Employed 55.88%*)

7. How many siblings do you have? (*Mean 1.55, Std Dev 1.13, Median 1, Min 0, Max 6*)

8. Who in your household is primarily responsible for expenses and budget decisions? Please select all that apply (*Self 38.24%, Spouse 0.49%, Shared Responsibility with Spouse 3.43%, Parent(s) 64.22%, Other 1.47%*)

9. Have you ever voted in a state or federal government election (in any country)? (*Yes 53.92%, No 46.08%*)

10. Before today, how many times have you participated in any economics or psychology experimental studies? (*Mean 3.46, Std Dev 3.47, Median 2, Min 0, Max 20*)

11. In the past twelve months, have you donated money to or done volunteer work for charities or other nonprofit organizations? (*Yes 77.94%, No 22.06 %*)

12. On a scale from 1 to 10, please rate how much you think communicating with your group members helped solve the two extra painting questions, with 1 meaning "not much at all". (*Mean 6.04, Std Dev 2.90, Median 7, Min 1, Max 10*)

13. On a scale from 1 to 10, please rate how closely attached you felt to your own group throughout the experiment, with 1 meaning "not closely at all". (*Mean 3.97, Std Dev 2.67, Median 3, Min 1, Max 10*)

14. In Part 2 when you were asked to decide on an effort level, how would you describe the strategies you used? Please select all that apply (*I tried to earn as much money as possible for myself 46.08%, I tried to earn as much money as possible for me and my match 50.00%, I tried to earn more money than my*

match 17.65%, I gave high effort if my previous matches gave high efforts and low effort if my previous matches gave low efforts 27.45%, Other 14.22%)

15. Please tell us how your match's group membership affected your decision. If I had been matched with someone from the other group [my own group], (*I would have picked higher effort levels 16.67% [23.61%], I would have picked lower effort levels 8.33% [1.39%], I would not have changed my effort levels 69.44% [72.22%], Other 5.56% [2.78%]*)

16. On a scale from 1 to 10, please rate how familiar you were with the paintings made by Klee and Kandinsky before this experiment, with 1 meaning "not familiar at all". (*Mean 1.31, Std Dev 1.00, Median 1, Min 1, Max 6*)

APPENDIX D

Chat Coding Training Session Summary and Instructions

1. Summary of Coding Procedures:

For the line level, the coders are told to take each line of each communication log and sort it into one or more categories. These categories denote whether the line is (a) about the paintings, (b) about the experiment or experimenter, (c) about the subject's group, (d) an expression of excitement, or (e) irrelevant information or none of the other categories. The "paintings" category is further subdivided by whether the line shows (i) painting analysis, (ii) a question about the paintings, (iii) agreement with another participant regarding the paintings, or (iv) disagreement with another participant regarding the paintings. The coders are told that any line can be part of multiple categories.

Next, the coders examine the communication from the subject level. First, for each subject, the coders tell us (a) whether or not the subjects made initial guesses regarding the paintings' artists. In particular, the coders examine whether the subjects (i) made a guess about painting 6, (ii) guessed painting 6 correctly ("Klee"), (iii) made a guess about painting 7, or (iv) guessed painting 7 correctly ("Kandinsky"). We also ask the coders to rate, on a 1 to 5 scale, each subject's level of (b) engagement in the conversation, (c) assertiveness, (d) confidence, and (e) politeness. Finally, the coders examine each communication on a group level (i.e. each communication log as a whole). Again, the coders are asked to rate, on a 1 to 5 scale, each group's level of (a) agreement, (b) confidence, (c) excitement, and (d) politeness.

To ensure that the coders understood their task and the system they would use to complete the coding, we held a training session for the coders. This training session was held on March 1, 2010, and lasted 2 hours. During this session, we first read the instructions out loud, answering any clarifying questions along the way. A copy of these instructions is included in Appendix D. Next, we had the coders examine 2 example communication logs. For this purpose, we obtain communication logs from an experiment conducted by Chen and Li (2009) that used the same paintings and communication procedure as this experiment. After this was completed, the coders were given a week to code the rest of the communication logs. All coding, including the practice coding, was performed on Google Docs. The coders were paid $15 an hour, with a total average payment of $78.

2. Coding Training Instructions:

You are now taking part in a study that seeks to characterize the communication patterns in chat logs. Your participation will take the form of coding conversations on several factors.

After taking part in this training session, you will code other conversations at home using a web-based system (Google Docs) over the next week. We ask you to not communicate with others about the coding during the course of this week. Should you have any questions while coding on your own, please email me at [email redacted].

The purpose of this training session is to familiarize you with the coding methodology to be employed, and to ensure a common understanding of the factors used.

However, this does not mean that you should all give identical codings. We are interested in eliciting objective codings from impartial coders. We ask you to rely on your own judgment when coding.

In the chat logs that you will be coding, the participants were asked to complete a task. First, the participants were divided into 2 groups, named “Red” and “Green” (For the chat logs we will be examining for this training session, the participants were instead in a group called “Maize”). They were shown 5 paintings by Paul Klee and 5 paintings by Wassily Kandinsky (which were labeled 1a, 1b, 2a, 2b, etc.). Then, the participants were given paintings labeled 6 and 7, and were asked to identify which artist painted each painting. These chat logs are the discussions that the participants had with members of their group in order to try to solve this problem.

In this training session you will be asked to code two chat logs. For each chat log, you will be asked to code the conversation at three levels, as shown below:

1. For each line of conversation, code whether it is
    (a) about the paintings. For this category, code whether the line shows
        i. analysis (1 = yes, 0 = no)
        ii. a question (1 = yes, 0 = no)
        iii. agreement with another participant (1 = yes, 0 = no)
        iv. disagreement with another partcipant (1 = yes, 0 = no)
    (b) about the experiment or experimenter (1 = yes, 0 = no)
    (c) about the participant’s group (1 = yes, 0 = no)
    (d) an expresssion of excitement (1 = yes, 0 = no)
    (e) irrelevant information or none of the above categories (1 = yes, 0 = no)

2. For each chat participant, code whether that participant

    (a) made initial guesses about the paintings. For this category, code whether or not the participant

        i. made a guess about painting 6 (1 = yes, 0 = no)

        ii. guessed "Klee" for painting 6 (1 = yes, 0 = no)

        iii. made a guess about painting 7 (1 = yes, 0 = no)

        iv. guessed "Kandinsky" for painting 7 (1 = yes, 0 = no)

    (b) was engaged in the conversation (1 = not engaged at all 5 = very engaged)

    (c) was assertive (1 = not assertive at all 5 = very assertive)

    (d) was confident (1 = not confident at all 5 = very confident)

    (e) had a nice tone towards the other participants (1 = very cold 5 = very warm and friendly)

3. For each chat log, code whether the participants, as a group,

    (a) agreed with each other (1 = no agreement at all 5 = full agreement)

    (b) were confident (1 = not confident at all 5 = very confident)

    (c) were excited (1 = not excited at all 5 = very excited)

    (d) had a nice tone towards each other (1 = very cold 5 = very warm and friendly)

Note that each line of the chat log can be coded into multiple categories. For example, if someone says, "I think that 6 is Kandinsky. What do you think?", this would be coded as painting analysis (1.a.i.) and as painting question (1.a.ii.).

The procedure we will follow in the training session is as follows:

1. First, we ask that you fill out a background questionnaire, included with the papers handed to you in your training packet. Please hand those to [the experimenter] once you have completed them.

2. Next, click on the link to Training Chat 1.

3. You will code Training Chat 1, working individually. Please let us know when you have finished coding the chat log (don't forget the participant-level and group-level categories, located on different sheets of the spreadsheet). If you do not remember the definitions of any of the variables, you can refer to this instruction sheet or the "Coding Category Legend" file. You may ask us questions at any time.

4. When coding variables that are either 0 or 1, we have set the default value to 0. To code any of these variables to be 1, simply change the 0 to a 1.

5. When coding the participant-level variables, you will want to sort the chat lines by the participants so that you can easily see what each participant said during the chat. To do this, click on the cell that says "Participant". Then, click on "Tools", and then click on "Sort sheet by column B, A→Z". To revert to the default line order, click on the cell that says "Line Number", click on "Tools", and then click on "Sort sheet by column A, A→Z".

6. When everyone has completed coding the chat log, there will be a brief discussion, no longer than 30 minutes, regarding the coding activity. We will go over each line, asking each of you for your codings. We will also present our codings and why we coded them the way we did.

7. When all questions have been addressed, you will click on the link to Training Chat 2 and code that chat log.

Are there any questions? Before we start, we would like to ask you to please take the time to read each chat log carefully when coding. We have found that it takes between 15 and 30 minutes to code each chat log when evaluating them carefully. If there are no further questions, let's begin.

APPENDIX E

Additional Tables

Table E.1: Interclass Correlation Coefficients (ICC)

Category	ICC	St. Error
Line Level Coding		
Painting analysis	0.808	0.010
Question about paintings	0.706	0.013
Agreement regarding paintings	0.746	0.012
Disagreement regarding paintings	0.453	0.019
Experiment or experimenter	0.551	0.018
Group	0.550	0.018
Level of excitement	0.553	0.017
Irrelevant (none of the above)	0.622	0.016
Subject Level Coding		
Made a guess about painting 6	0.537	0.058
Guessed Klee for painting 6	0.522	0.059
Made a guess about painting 7	0.585	0.055
Guessed Kandinsky for painting 7	0.609	0.053
Level of engagement	0.744	0.040
Level of assertiveness	0.508	0.060
Level of confidence	0.346	0.064
Level of politeness	0.254	0.064
Group Level Coding		
Level of agreement	0.434	0.159
Level of confidence	0.449	0.157
Level of excitement	0.259	0.160
Level of politeness	0.018	0.126

Table E.2: Group Identity and Equilibrium: Probit Regression (Φ^{-1}(equilibrium) = $\beta_0 + \beta_1$*Ingrp+β_2*Outgrp+β_3*Ingrp*Enh+β_4*Outgrp*Enh+u_{it})

	Dependent Variable: Equilibrium
Ingroup	0.14
	(0.11)
Outgroup	0.02
	(0.11)
Ingroup*Enhanced	0.21**
	(0.10)
Outgroup*Enhanced	0.01
	(0.13)
Observations	5400
Pseudo-R^2	0.0584

Notes: Standard errors are adjusted for clustering at the session level. Significant at the: ** 5 percent level.

Table E.3: Average Efficiency by Session and Treatment

	Ingroup	Outgroup	Control
	0.65	0.49	0.56
Near-minimal	0.63	0.67	0.70
	0.70	0.68	0.63
Average	0.66	0.62	0.63
	0.85	0.80	0.58
Enhanced	0.90	0.50	0.80
	0.86	0.55	0.57
Average	0.87	0.62	0.65

Table E.4: Group Identity and Efficiency: Random-Effects
(Efficiency = $\beta_0 + \beta_1$*Ingrp+β_2*Outgrp+β_3*Ingrp*Enh+β_4*Outgrp*Enh+u_{it})

	Dependent Variable: Efficiency
Ingroup	0.02
	(0.04)
Outgroup	-0.03
	(0.06)
Ingroup*Enhanced	0.21***
	(0.03)
Outgroup*Enhanced	0.00
	(0.09)
Constant	0.64***
	(0.04)
Observations	5400
R^2	0.1251

Notes: Standard errors are adjusted for clustering at the session level. Significant at the: *** 1 percent level.

Table E.5: Effort Distributions

Treatment		Calibrated α^g	Predicted Mean	Predicted SD	Actual Mean	Actual SD
Near-Minimal	Control	0.41	145.65	14.90	133.28	13.07
	Ingroup	0.51	128.91	23.67	148.29	19.18
	Outgroup	0.65	164.37	18.61	157.46	20.96
Enhanced	Control	0.31	137.33	14.89	132.83	27.13
	Ingroup	0.83	139.94	19.09	166.15	6.78
	Outgroup	0.42	157.33	23.61	133.65	25.97

APPENDIX F

Reconciling Theory and Experiments

In this appendix, we apply our theoretical framework to previous experimental studies on coordination games, including the minimum-effort games, Battle of the Sexes, and the provision point mechanism. By incorporating group identity into the potential games framework, we can reconcile findings from previous studies and thus showcase the applications of our theory.

We first examine studies of the minimum-effort games that are successful in achieving higher effort levels contrary to the predictions of the theory of potential games. A summary of these studies and the other studies of the minimum-effort game mentioned in Section 1.2 is shown in Table F.1.[1] In addition to the parameter configurations of each experiment (strategy space, T, n, a, b and c), the last three columns present the cutoff marginal cost c^*, the theoretical predictions from standard potential maximization, and the empirical trend observed in the experiment, respectively. Recall that standard potential maximization theory predicts that choices converge to the low (high) effort equilibrium if $c > c^*$ ($c < c^*$). This prediction is consistent with the results from the three baseline studies by Van Huyck et al. (1990), Goeree and Holt

[1] Rather than exhaustively listing all experiments of the minimum-effort games, we instead present representative studies in each category.

(2005), and Knez and Camerer (1994), as well as many treatments in subsequent categories. Whenever the theoretical prediction is inconsistent with the observed trend, we put the treatment in bold face. In what follows, we discuss the three approaches used in the literature to achieve higher effort levels contrary to the theoretical predictions and how incorporating group identity into the potential function could reconcile theory and the empirical findings (Propositions 1 and 4).

Over two papers, Camerer and Knez (1994 and 2000) show that, if they use the same parameters as VHBB ($a = 0.2, b = 0.6, c = 0.1$) in the minimum-effort game, subjects will converge to the efficient equilibrium after 5 periods if $n = 2$, but not if $n = 3$. Using the phenomenon of "transfer of precedent," Camerer and Knez show that it is possible to make 3-player matches converge to the efficient equilibrium if the game is first played for 5 periods by two-player matches with the third player observing, and then for 5 more periods with all 3 players. The two-player matches establish a group norm of high effort that is consistent with potential theory, which is then transferred to the three-player matches. Allowing the third player to watch the other 2 players for 5 periods implicitly creates a group, establishes a group norm, and increases subjects' other-regarding preferences.

Weber (2006) shows that it is possible to apply Camerer and Knez's result successively to achieve higher effort levels in larger groups. Using parameters similar to VHBB ($a = 0.2, b = 0.2, c = 0.1$), Weber slowly grows the number of players in the minimum-effort game over 22 periods from $n = 2$ to $n = 12$. He shows that, if growth is too fast, or if no history is shown to the new players, then subjects converge to the least efficient equilibrium. If, on the other hand, the groups are grown slowly enough, and it is common knowledge that the new players observe the entire history of efforts provided, the entire 12-person group is able to achieve a minimum effort of 5 by the final period. Again, the observation of smaller groups facilitates the establishment of group norms.

Gary Bornstein and Nagel (2002) use a different method, intergroup competition, to promote higher effort levels. Taking essentially the same game as VHBB ($a = 20, b = 60, c = 10$), Bornstein et al. divide subjects into two competing groups of size $n = 7$. The group with the higher chosen minimum effort level is paid according to the normal payoff function, while the group with the lower chosen minimum effort level is paid nothing (in the case of a tie, everyone is paid according to half the normal payoff function). This revised payment method changes the game. In particular, the set of Nash equilibria is expanded. It is still a Nash equilibrium for every member of both groups to give the same level of effort, but it is also a Nash equilibrium for the members of one group to all give the same effort, and two members of the other group to give a lower effort (the rest of the members of this other group can give any level of effort and still preserve the Nash equilibrium). While the potential function is also changed in this scenario, the potential maximizing Nash equilibrium remains the equilibrium in which every member of both groups gives the minimum possible effort of 1. So, if social preferences are ignored, then the prediction of potential theory is that players will converge to the least efficient equilibrium. In another treatment, the subjects are all paid according to the normal payoff function, but are also given the extra information of what the minimum effort level is in the other group (this information is withheld in the control). This separates the effect of receiving this information from the actual competition. While Bornstein et al. find that the extra information has no effect (the control yields an average effort of 3.6 while the information treatment yields an average effort of 3.5), there is a significant increase in chosen effort with intergroup competition (average effort 5.3). In another session, instead of punishing the losing group, the winning group receives a bonus. This yields an average effort of 4.5, also significantly higher than in the control or information sessions. By explicitly tying the subjects' payoffs to the choices of the group, and by making the 2 groups compete with each other, Bornstein et al. create a very strong ingroup and outgroup effect that is able to raise the threshold c^* above the marginal cost of 10 used in the experiment.

Another approach to increase effort is to facilitate communication across group members. Specifically, Chaudhuri et al. (2009) suggest that giving subjects advice from previous subjects of the experiment can increase effort in large groups. Using the same parameters as VHBB and $n = 8$, the authors attempt to induce higher effort by providing subjects with full histories of previous sessions of the experiment, and by providing advice about the game given by previous subjects. Most of this advice suggests that players always give the highest effort. While this is not successful in most treatments, all of which have "private advice"(all subjects receive the advice but this is not common knowledge), the subjects do converge to the highest effort level when the advice is "public"(common knowledge). One plausible interpretation is that, communication between subjects creates an ingroup effect strong enough to induce high efforts, even if subjects in a session simply receive communication from a third party, as long as it is common knowledge that this communication is taking place.

Brandts and Cooper (2007) also examine the effect of communication in the minimum-effort game. Communication in this study is achieved through a manager, who is the only subject allowed to talk to the other 4 subjects in a "firm." These 4 other subjects are workers of the firm who play a minimum-effort game ($a = 6$ or $14, b = 200, c = 5$) with efforts restricted to 0, 10, 20, 30 or 40. The manager's payoff is also positively related to the minimum effort given by the 4 workers. Brandts and Cooper run three different treatments. In the first, the manager cannot communicate with the other subjects, but can control their financial incentives. In the second, managers can send messages to the other subjects (after the 10th period). This treatment is the most similar to the study run by Chaudhuri et al. The only difference here is that the third-party communicator has a stake in the game being played between the other players. In the third treatment, managers can send messages to other subjects and the subjects can send messages to the manager (also after the 10th period). The main result of this paper is that more avenues of communication lead to higher minimum effort levels. The two-way communication treatment yields higher minimum effort levels than the one-way communication treatment, and the same is true for the one-way communication treatment compared to the no-communication treatment. This result holds even when they consider only the sessions with minimum effort levels of 0 after the 10th period. The effect of communication in a coordination game may work through a different channel than other-regarding preferences, such as trust or learning (see Brandts and Cooper (2007) for a list, based on the content of the messages sent by the managers). However, discussions with the authors reveal that the most successful messages appeal to a group identity.

In addition to the minimum-effort game, experimental studies of the provision point mechanism (PPM) indicate that competition between groups increases the likelihood of successful coordination to an efficient equilibrium. The PPM is proposed by Bagnoli and Lipman (1989), with the property that it fully implements the core in undominated perfect equilibria in an environment with one private good and a single unit of public good.[2] In a complete information economy, agents voluntarily contribute any nonnegative amount of the private good they choose and the social decision is to provide the public good if and only if contributions are sufficient to pay for it. The contributions are refunded otherwise. This mechanism has a large class of Nash equilibria, some of which are efficient while others not. Among a large number of experimental studies of this mechanism, two studies highlight the effects of group competition in equilibrium selection, even though neither was explicitly designed to test group effects. First, Bagnoli and McKee (1991) study the mechanism with several independent groups simultaneously in the same room and publicly post-

[2]With multiple discrete units, the theoretical results also hold, but there have been very few experimental studies of the multiple-unit case.

ed contributions for all groups. They find public good is provided in 86.7 percent of the rounds. Second, Mysker et al. (1996) use the same parameters but with single, isolated groups. The latter is not nearly as successful as the former in coordinating to the efficient outcome. In the Bagnoli and McKee study, the efficient equilibrium contribution is a modal distribution, while in the Mysker, Olson and Williams study, contributions are evenly distributed along the strategy space. From the perspective of potential games, we can show that, in general, PPM is not a potential game.[3] However, with group competition, it can be transformed into a potential game where the potential maximizing equilibrium is the set of efficient equilibria.

Another well-studied coordination game is the Battle of the Sexes game (BoS hereafter). Charness et al. (2007) report a series of experiments on the effects of group membership on equilibrium selection in BoS games (as well as the prisoner's dilemma games). In treatments where groups are salient, the authors find that group membership significantly affects the rate of successful coordination. Taking a version of BoS such as the one on the left in the table below (Charness et al., 2007), it is straightforward to show that it is a potential game with the potential function given by $P = 4p_1p_2 - p_1 - 3p_2$, where p_i denotes the probability with which player i chooses A. Hence the potential is maximized by the mixed strategy equilibrium $(p_1 = 0.25, p_2 = 0.75)$. This prediction is consistent with the findings of Cooper et al. (1989), who show that subjects converge to a frequency of choices that is close to the mixed strategy equilibrium in BoS. If we transform the game to incorporate the effects of group identity, we obtain the game on the right, with the new potential function $P = 4(1+\alpha)p_1p_2-(1+3\alpha)p_1-(3+\alpha)p_2$, which is again maximized at its mixed strategy equilibrium. It is straightforward to show that the probability of coordination, $p_1p_2 + (1 - p_1)(1 - p_2)$, is increasing in α. This leads to a directional prediction that the

[3] A counter example can be constructed from an example used in Menezes et al. (2001). Let $x_i = \{0, c\}$. In the two-player case, $\pi_i(c, c) - \pi_i(c, 0) - [\pi_i(0, c) - \pi_i(0, 0)] = -v_i$, which violates the definition of potential games when $v_1 \neq v_2$.

probability of coordination is higher for ingroup matching compared to the control and outgroup matching, and increases with the salience of group identity.

Original BoS

	A	B
A	3, 1	0, 0
B	0, 0	1, 3

Transformed BoS

	A	B
A	$3+\alpha$, $1+3\alpha$	0, 0
B	0, 0	$1+3\alpha$, $3+\alpha$

In sum, we find that social identity, group competition, and group norms improve coordination in games with multiple Nash equilibria. Incorporating group identity into potential games provides a unifying framework which reconciles findings from a number of coordination game experiments.

Table F.1: Summary of Studies Regarding the Minimum-Effort Game

	Study	Treatment	Efforts	T (# of rounds)	n (# per match)	$\pi_i = a\min(x) - cx_i + b$ $P = a\min(x) - c\sum_i x_i$ a	b	c	Threshold $c^* = \frac{a}{n}$	Theoretical Prediction	Observed Trend
Baseline	Van Huyck, Battalio, Beil (1990)	Large Groups	{1,...,7}	10	14-16	0.20	0.60	0.10	0.01	Low	Low
		No Cost	{1,...,7}	5	14-16	0.20	0.60	0.00	0.01	High	High
		2-person	{1,...,7}	7	2	0.20	0.60	0.10	0.10	High	High
	Goeree, Holt (2005)	2-person, c=1/4	[110,170]	10	2	1.00	0.00	0.25	0.50	High	High
		2-person, c=3/4	[110,170]	10	2	1.00	0.00	0.75	0.50	Low	Low
		3-person, c=1/10	[110,170]	10	3	1.00	0.00	0.10	0.33	High	High
	Knez, Camerer (1994)	3-person	{1,...,7}	5	3	0.20	0.60	0.10	0.07	Low	Low
		6-person	{1,...,7}	5	6	0.20	0.60	0.10	0.03	Low	Low
Transfer of Precedent	Camerer, Knez (2000)	2-person	{1,...,7}	5	2	0.20	0.60	0.10	0.10	High	High
		2-person	{1,...,7}	5	2	0.20	0.60	0.10	0.10	High	High
		3-person	{1,...,7}	5	3	0.20	0.60	0.10	0.07	Low	Low
		2-, 3-person	{1,...,7}	5	2→3	0.20	0.60	0.10	0.07	**Low**	**High**
	Weber (2006)	No Growth	{1,...,7}	12	12	0.20	0.20	0.10	0.02	Low	Low
		No History	{1,...,7}	22	2→12	0.20	0.20	0.10	0.02	Low	Low
		Fast Growth	{1,...,7}	22	2→12	0.20	0.20	0.10	0.02	Low	Low
		Slow Growth	{1,...,7}	22	2→12	0.20	0.20	0.10	0.02	**Low**	**High**
Inter-Group Competition	Bornstein, Gneezy Nagel (2002)	No Comp.	{1,...,7}	10	7	20	60	10	2.86	Low	Low
		Info	{1,...,7}	10	7	20	60	10	2.86	Low	Low
		Group Comp.	{1,...,7}	10	7	20	60	10	2.86	**Low**	**High**
Communication	Chaudhuri, Schotter, Soper (2001)	Low Cost	{1,...,7}	10	8	0.20	0.60	0.10	0.03	Low	Low
		Progenitor	{1,...,7}	10	8	0.20	0.60	0.10	0.03	Low	Low
		History, Advice	{1,...,7}	10	8	0.20	0.60	0.10	0.03	Low	Low
		Advice	{1,...,7}	10	8	0.20	0.60	0.10	0.03	Low	Low
		Public Advice	{1,...,7}	10	8	0.20	0.60	0.10	0.03	**Low**	**High**
	Brandt, Cooper (2007)	Computer	{0,...,40}	20	4	10	200	5	2.50	**Low**	**High**
		No Comm.	{0,...,40}	20	4	9.3*	200	5	2.33	Low	Low
		One-way Comm.	{0,...,40}	20	4	9.3*	200	5	2.33	**Low**	**High**
		Two-way Comm.	{0,...,40}	20	4	9.9*	200	5	2.48	**Low**	**High**

*Chosen by subjects; average reported

APPENDIX G

Kiva Data

This appendix contains a list of the data that we are able to collect from Kiva. We break this data into those that concern lenders, teams, and loans.

1. Lenders
   (a) Kiva API
      i. name
      ii. profile picture
      iii. country code: lender's country
      iv. whereabouts: lender's detailed location information in free text form (i.e. state/province, and city information)
      v. member since: when the lender joined Kiva
      vi. occupation: lender's occupation information in free text form
      vii. I loan because: lender's motivation statement
      viii. number of loans: number of loans given by lender
      ix. teams: list of teams that the lender has joined

(b) Incentivized coding

  i. gender or group type: lender's gender or group type. These are either hand coded as described in section 2.4 or predicted using a name dictionary from the US Census. Possible values include: Male, Female, Couple, Family, Company, Other Group.

  ii. motivation category: 5250 randomly selected motivation statements hand coded, the rest classified using a trained machine learning classifier

  iii. occupation category: 5250 randomly selected occupation descriptions hand coded

2. Teams

   (a) name

   (b) category

   (c) whereabouts

   (d) We loan because

   (e) team since: when the team was founded

   (f) membership type: open or closed to all Kiva lenders

   (g) member count: number of members

   (h) loan count: number of loans belonging to the team

   (i) loaned amount: total amount of lending attributed to the team

3. Loans

   (a) id

   (b) posted date

(c) loan amount

4. Lenders, Loans

   (a) the list of loans which the lender borrowed to

5. Teams, Loans

   (a) the list of loans belonging to the team

6. Teams, Lenders

   (a) the list of public lenders belong to the team

   (b) team join date: when a team member joined the team

APPENDIX H

Motivation and Occupation Categories

This appendix contains a list of the categories for motivation statements and occupations that we use in the incentivized coding. Subjects take each motivation statement and occupation and place them into one or more of the following categories. Afterwards, we take these examples and train machine-learning classifiers to classify the rest.

Motivation categories:

1. **General altruism** (Gnl. Altruism): e.g., "I believe in a global community."
2. **Group-specific altruism** (Grp. Altruism): e.g., "I want to help women succeed in business and in life."
3. **Empathy**: e.g., "I am disabled and I know what it's like to feel helpless."
4. **Reciprocity**: e.g., "I am very fortunate to have several people in my life to lend me a hand when I needed help. I hope that I can do the same for someone."
5. **Equality and social safety net** (Equity): e.g., "I want to help others who are less fortunate. Everyone deserves a fair chance."

6. **Social responsibility and social norms** (Norms): e.g., "I have the ability and I'm lucky enough to be able to."

7. **Effective development tool** (Tool): e.g., "I believe in change through bottom-up initiatives and sustainable business models."

8. **Personal satisfaction** (Satisfaction): e.g., "It makes my heart smile."

9. **Religious duty** (Religious): e.g., "I believe that sometimes God works thru people to answer prayers. What a privilege!"

10. **External reasons** (External): e.g., "It's for a community service project at my university."

Occupation categories:

1. **Art, Media and Entertainment** (Entertainment): Artist, Musicians, Directors, Designers, Writers, Journalists, Producers, Editors

2. **Entrepreneurship** (Entrepreneur): Business Owners, Entrepreneurs, Venture Capitalists, Angel Investors, Business Incubator Managers

3. **Business and Finance** (Business): Financial Managers, CEOs, Management Consultants, Accountants, Bankers, CPAs, Bookkeepers, Loan Officers, Auditors, Business Analysts

4. **Information Technology** (IT): Computer Programmers, Software Developers, Systems Administrators, Web Developers, Technical Support Specialists

5. **Higher Education** (Higher Ed.): Professors, College/University Students, Researchers

6. **Primary, Secondary, and other Education** (Education): Teachers, Educators, Non-College Students, School Administrators

7. **Engineering**: Mechanical Engineers, Electrical Engineers, Chemical Engineers, Civil Engineers, Environmental Engineers, Systems Engineers, Computer Engineers

8. **Health Care**: Doctors, Dentists, Nurses, Medical Assistants, Medical Students, Emergency Medical Personnel

9. **Household Management** (Home): Homemakers, Stay-at-Home Parents, Child Care Providers, Housekeepers

10. **Retail**: Salespeople, Retail Buyers, Cashiers, Clerks, Store Managers, Sales Managers

11. **Law and Government** (Government): Lawyers, Judges, Court Personnel, Elected Officials, Civil Servants, Military Personnel, Firefighters, Police Officers, Social Workers

12. **Non-profit**: Any work with non-profit organizations

13. **Retired**: Any lenders who say they are retired

APPENDIX I

Experimental Instructions

(Endogenous Groups)

Experimental Instructions

Welcome!

This is an experiment in decision making. The amount of money you earn will depend on the decisions you make and on the decisions other people make. At the end of the session, you will be paid your earnings plus a $5 show-up fee in cash. Everyone will be paid in private and you are under no obligation to tell others how much you earn.

Please do not communicate with each other during the experiment. If you have a question, feel free to raise your hand, and an experimenter will assist you.

<u>Part 1: Effort</u>

The first part of the experiment lasts 80 rounds. For each round, you will be asked to make one or more decisions.

Each round, everyone will choose an effort level between 110.00 and 170.00 in increments of 0.01. You will earn a number of points equal to the minimum effort level chosen by the members of your team, minus the cost of your own effort, which is 0.50 times your own effort level. This is captured by the equation:

Your Earnings (Points)= Minimum Effort (Your Team)- 0.50 * Your Effort

Your team consists of 3 people in this room, including yourself. **Every 650 points you earn is worth $1 to you.**

In the folder on your desk, you should find an earnings matrix for the first part of the experiment. Along the left side of this matrix, you will see your effort level, and along the top of this matrix, you will see the minimum effort level of the **other** members of your team. The cells show how many points you earn in each case.

Note that the earnings matrix does not show all of the possible effort levels that you can choose. If you want to choose an effort level that is not listed on the earnings matrix, please use the equation discussed previously.

As an example, assume that you decide to contribute an effort level of 130.00, and your team members decide to contribute effort levels of 120.00 and 150.00. Then:

***Your Earnings (Points)*=120.00- 0.50 * 130.00 = 55.0**

since the minimum effort in your team was 120.00 and your effort is 130.00.

On the earnings matrix, you should look at the row labeled "130" and the column labeled "120". The result should still be 55.0 points.

After we have finished reading these instructions, you will be asked to calculate the earnings of the other members of your team in this example to ensure that you understand the experiment.

Teams

In the first round, you will tell us your first and second choice teams. In a random order, you will be placed into your most preferred team with fewer than 3 members in it. So, the 3 people who are randomly chosen to go first are placed into their first choice teams. Then, the next person is placed into his or her first choice team if that team does not have 3 people in it yet. Otherwise, he or she is placed into his or her second choice team. This continues until 2 teams have 3 people in them. Then, all remaining people are placed into the last team. Note that you may be placed into your last choice team.

Every 2 rounds after the first round, you will be able to change your team. First, you will tell us whether you would like to stay in your current team. If you say you want to stay, then you will remain in your current team. Otherwise, you will again tell us your first and second choice teams. Following the same procedure as in the first round, you will be placed into a team. Starting on round 21, if every member of your team decides to stay in the team 2 times in a row, then your team will be fixed. Once this occurs, no one will be able to leave or join your team for the rest of the experiment. After round 50, regardless of everyone's choices, all teams will be fixed.

At the end of every round, we will show you how all teams have performed. We will publicly display each team's minimum effort and average earnings, and privately show you your own earnings. This information will be ranked by each team's minimum effort, and then by each team's average earnings if multiple teams' minimum efforts are equal.

Part 2: Dividing Points

The second part of the experiment lasts only 1 round. For each of 6 situations, you will be given 100 points and asked to divide them between 2 people in the room.

In 3 situations, you will be asked to divide the points between yourself and one other person. In 3 other situations, you will be asked to divide the points between 2 other people.

The people that you give these points to will have those points added to their earnings for the experiment. The points that you decide to keep for yourself are added to your own earnings. **As before, every 650 points you earn is worth $1 to you.**

After this, you will be asked to fill out a demographic survey.

If you need to make any calculations, please use the blank sheet of paper provided in the folder. Are there any questions?

Earnings Matrix

Minimum Effort Given by Other Team Members

Amount of Effort You Give

	110	115	120	125	130	135	140	145	150	155	160	165	170
110	55.0	55.0	55.0	55.0	55.0	55.0	55.0	55.0	55.0	55.0	55.0	55.0	55.0
115	52.5	57.5	57.5	57.5	57.5	57.5	57.5	57.5	57.5	57.5	57.5	57.5	57.5
120	50.0	55.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0
125	47.5	52.5	57.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5	62.5
130	45.0	50.0	55.0	60.0	65.0	65.0	65.0	65.0	65.0	65.0	65.0	65.0	65.0
135	42.5	47.5	52.5	57.5	62.5	67.5	67.5	67.5	67.5	67.5	67.5	67.5	67.5
140	40.0	45.0	50.0	55.0	60.0	65.0	70.0	70.0	70.0	70.0	70.0	70.0	70.0
145	37.5	42.5	47.5	52.5	57.5	62.5	67.5	72.5	72.5	72.5	72.5	72.5	72.5
150	35.0	40.0	45.0	50.0	55.0	60.0	65.0	70.0	75.0	75.0	75.0	75.0	75.0
155	32.5	37.5	42.5	47.5	52.5	57.5	62.5	67.5	72.5	77.5	77.5	77.5	77.5
160	30.0	35.0	40.0	45.0	50.0	55.0	60.0	65.0	70.0	75.0	80.0	80.0	80.0
165	27.5	32.5	37.5	42.5	47.5	52.5	57.5	62.5	67.5	72.5	77.5	82.5	82.5
170	25.0	30.0	35.0	40.0	45.0	50.0	55.0	60.0	65.0	70.0	75.0	80.0	85.0

APPENDIX J

Postexperimental Survey (Endogenous Groups)

Please answer the following survey questions. Your answer will be used for this study only. Individual data will not be released.

1. What is your age? (*Mean 22.24, Std Dev 4.18, Median 21, Min 18, Max 54*)
2. What is your gender? (*Male 51.39%, Female 48.61%*)
3. Are you an undergraduate or a graduate student? (*Undergraduate 73.61%, Graduate 19.44%, Neither 6.94%*)
4. Which year of your current educational program did you complete in April/May of 2011? (*1st Year 9.72%, 2nd Year 20.83%, 3rd Year 25.69%, 4th Year 29.86%, Higher Year 9.03%, N/A 4.86%*)
5. Which of the following best describes your racial or ethnic background? (*Asian 34.03%, Black 15.97%, Caucasian 40.97%, Hispanic 0.69%, Native American 0.69%, Multiracial 6.25%, Other 1.39%*)

6. In what country or region were you primarily raised as a child? (*US/Canada 81.94%, Africa 0.69%, Asia 15.28%, Australia 0.00%, Europe 0.69%, Latin America 0.00%, Middle East 1.39%*)

7. What is your marital status? (*Never Married 96.53%, Currently Married 2.78%, Previously Married 0.69%*)

8. How would you best describe your employment status? (*Employed Full Time 11.81%, Employed Part Time 52.08%, Not Employed 36.11%*)

9. How many siblings do you have? (*Mean 1.64, Std Dev 1.38, Median 1, Min 0, Max 8*)

10. Who in your household is primarily responsible for expenses and budget decisions? Please select all that apply. (*Self 56.94%, Spouse 0.69%, Shared Responsibility with Spouse 1.39%, Parent(s) 54.17%, Other 0.00%*)

11. Have you ever voted in a state or federal government election (in any country)? (*Yes 63.89%, No 36.11%*)

12. Before today, how many times have you participated in any economics or psychology experimental studies? (*Mean 6.94, Std Dev 8.43, Median 5, Min 0, Max 50*)

13. In the past twelve months, have you donated money to or done volunteer work for charities or other nonprofit organizations? (*Yes 74.31%, No 25.69 %*)

14. On a scale from 1 to 10, please rate how closely attached you felt to your own group throughout the experiment, with 1 meaning "not closely at all". (*Mean 5.78, Std Dev 3.37, Median 6.5, Min 1, Max 10*)

15. How would you describe the strategies you used in the Effort part of the experiment? Please select all that apply. (*I tried to earn as much money as*

possible for myself 76.39%, I tried to earn as much money as possible for my team 54.17%, I tried to earn more money than the other members of my team 9.72%, I tried to ensure that my team earned more money than the other teams 20.14%, Other 9.72%)

16. How did you decide how to divide the points between different people? Please select all that apply. (*I tried to earn as much money as possible for myself 78.47%, I tried to earn as much money as possible for my team 34.72%, I tried to earn more money than the other members of my team 11.11%, I tried to ensure that my team earned more money than the other teams 25.69%, Other 14.58%*)

BIBLIOGRAPHY

BIBLIOGRAPHY

Abdulkadiroğlu, Atila and Tayfun Sönmez, "Matching Markets: Theory and Practice," 2010. Prepared for the 2010 Econometric Society World Congress.

Ahn, T.K., R. Mark Isaac, and Timothy C. Salmon, "Endogenous Group Formation," *Journal of Public Economic Theory*, April 2008, *10* (2), 171–194.

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *The Quarterly Journal of Economics*, August 2000, *115* (3), 715–753.

_ **and** _ , "Identity and the Economics of Organizations," *Journal of Economic Perspectives*, Winter 2005, *19* (1), 9–32.

_ **and** _ , *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, 1 ed., Princeton, NJ: Princeton University Press, 2010.

Anderson, Simon P., Jacob K. Goeree, and Charles A. Holt, "Minimum-Effort Coordination Games: Stochastic Potential and Logit Equilibrium," *Games and Economic Behavior*, February 2001, *34* (2), 177–199.

Arifovic, Jasmina and John O. Ledyard, "Individual Evolutionary Learning, Other-regarding Preferences, and the Voluntary Contributions Mechanism," 2009. Working Paper.

Armendáriz, Beatriz and Jonathan Morduch, *The Economics of Microfinance*, 2 ed., Cambridge, Massachusetts: The MIT Press, 2010.

Backstrom, L., D. Huttenlocher, J. Kleinberg, and X. Lan, "Group formation in large social networks: membership, growth, and evolution," in "Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining" 2006, pp. 44–54.

Bagnoli, Mark E. and Bart L. Lipman, "Provision of Public Goods: Fully Implementing the Core Through Private Contributions," *Review of Economic Studies*, October 1989, *56* (4), 583–601.

_ **and Michael McKee**, "Voluntary Contribution Games: Efficient Private Provision of Public Goods," *Economic Inquiry*, April 1991, *29* (2), 351–366.

Ball, J R, "'Space missions' Focus on Team Building," *The Greater Baton Rouge Business Report*, March 1999, *17* (15), 31.

Barabási, A.L. and R. Albert, "Emergence of scaling in random networks," *Science*, 1999, *286* (5439), 509.

Basu, Kaushik, "Participatory Equity, Identity, and Productivity: Policy Implications for Promoting Development," 2006. Center for Analytical Economics Working Paper 06-06.

Blume, Lawrence E., "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, July 1993, *5* (3), 387–424.

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

Bornstein, Uri Gneezy Gary and Rosemarie Nagel, "The Effect of Intergroup Competition on Group Coordination: An Experimental Study," *Games and Economic Behavior*, October 2002, *41* (1), 1–25.

Brandts, Jordi and David J. Cooper, "It's What You Say, Not What You Pay: An Experimental Study of Manager-Employee Relationships in Overcoming Coordination Failure," *Journal of the European Economic Association*, December 2007, *5* (6), 1223–1268.

Brewer, Marilynn B., "The Psychology of Prejudice: Ingroup Love and Outgroup Hate?," *Journal of Social Issues*, Fall 1999, *55* (3), 429–444.

Brosig, Jeannette, Magdalena Margreiter, and Joachim Weimann, "Endogenous Group Formation and the Provision of Public Goods: The Role of Promises and Lies," 2005. Mimeo.

Camerer, Colin F., *Behavioral Game Theory: Experiments on Strategic Interactions*, 1 ed., Princeton, NJ: Princeton University Press, 2003.

Cason, Timothy N. and Vai-Lam Mui, "Social Influence in the Sequential Dictator Game," *Journal of Mathematical Psychology*, June 1998, *42* (2–3), 248–265.

Charness, Gary and Chun-Lei Yang, "Endogenous Group Formation and Efficiency: An Experimental Study," in "BQGT '10 Proceedings of the Behavioral and Quantitative Game Theory: Conference on Future Directions" 2010.

_ **and Matthew Rabin**, "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, August 2002, *117* (3), 817–869.

_ **, Luca Rigotti, and Aldo Rustichini**, "Individual Behavior and Group Membership," *American Economic Review*, September 2007, *97* (4), 1340–1352.

_ **, Ramón Cobo-Reyes, and Natalia Jiménez**, "Efficiency, Team building, and Identity in a Public-goods Game," 2011. Working Paper.

Chaudhuri, Ananish, Andrew Schotter, and Barry Sopher, "Talking Ourselves to Efficiency: Coordination in Inter-Generational Minimum Effort Games with Private, Almost Common and Common Knowledge of Advice," *The Economic Journal*, January 2009, *119* (534), 91–122.

Chen, Roy and Yan Chen, "The Potential of Social Identity for Equilibrium Selection," *American Economic Review*, October 2011, *101* (6), 2562–2589.

Chen, Yan and Sherry Xin Li, "Group Identity and Social Preferences," *American Economic Review*, March 2009, *99* (1), 431–457.

_ , F. Maxwell Harper, Joseph Konstan, and Sherry Xin Li, "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens," *American Economic Review*, September 2010, *100* (4), 1358–1398.

Cheung, Yin-Wong and Daniel Friedman, "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior*, April 1997, *19* (1), 46–76.

Cinyabuguma, Matthias, Talbot Page, and Louis Putterman, "Cooperation Under the Threat of Expulsion in a Public Goods Experiment," *Journal of Public Economics*, August 2005, *89* (8), 1421–1435.

Cooper, David J. and Carol Kraker Stockman, "Fairness and learning: an experimental examination," *Games and Economic Behavior*, October 2002, *41* (1), 26–45.

Cooper, Russell W., Douglas DeJong, Robert Forsythe, and Thomas Ross, "Communication in the Battle of the Sexes: Some Experimental Results," *The RAND Journal of Economics*, Winter 1989, *20* (4), 568–587.

Coricelli, Giorgio, Dietmar Fehr, and Gerlinde Fellner, "Partner Selection in Public Goods Experiments," *Journal of Conflict Resolution*, June 2004, *48* (3), 356–378.

Cox, James C., Daniel Friedman, and Steven Gjerstad, "A Tractable Model of Reciprocity and Fairness," *Games and Economic Behavior*, April 2007, *59* (1), 17–45.

Crawford, Vincent P., "An 'evolutionary' Interpretation of Van Huyck, Battalio, and Beil's Experimental Results on Coordination," *Games and Economic Behavior*, February 1991, *3* (1), 25–49.

_ , "Adaptive Dynamics in Coordination Games," *Econometrica*, January 1995, *63* (1), 103–143.

_ **and Bruno Broseta**, "What Price Coordination? The Efficiency-Enhancing Effect of Auctioning the Right to Play," *American Economic Review*, March 1998, *88* (1), 198–225.

Croson, Rachel, Enrique Fatás, and Tibor Neugebauer, "Excludability and Contribution: A Laboratory Study in Team Production," 2006. Under Review.

_, **Melanie Marks, and Jessica Snyder**, "Groups Work for Women: Gender and Group Identity in Social Dilemmas," *Negotiation Journal*, October 2008, *24* (4), 411–427.

Deaux, Kay, "Social Identification," in E. Tory Higgins and Arie W. Kruglanski, eds., *Social Psychology: Handbook of Basic Principles*, New York, NY: The Guilford Press, 1996.

Eaton, B. Curtis, Mukesh Eswaran, and Robert J. Oxoby, "'Us' and 'Them': the origin of identity, and its economic implications," *Canadian Journal of Economics*, August 2011, *44* (3), 719–748.

Eckel, Catherine C. and Philip J. Grossman, "Managing Diversity by Creating Team Identity," *Journal of Economic Behavior & Organization*, November 2005, *58* (3), 371–392.

_ **and Rick K. Wilson**, "Social Learning in Coordination Games: Does Status Matter?," *Experimental Economics*, September 2007, *10* (3), 317–329.

Ehrhart, Karl-Martin and Claudia Keser, "Mobility and Cooperation: On the Run," 1999. CIRANO Scientific Series 99s-24.

Falk, Armin and Urs Fischbacher, "A Theory of Reciprocity," *Games and Economic Behavior*, February 2006, *54* (2), 293–315.

Fehr, Ernst and Klaus M. Schmidt, "The Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics*, August 1999, *114* (3), 817–868.

Fershtman, Chaim and Uri Gneezy, "Discrimination in a Segmented Society: An Experimental Approach," *The Quarterly Journal of Economics*, February 2001, *116* (1), 351–377.

Fischbacher, Urs, "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, June 2007, *10* (2), 171–178.

Frey, Bruno S. and Stephan Meier, "Social Comparisons and Pro-Social Behavior: Testing Conditional Cooperation in a Field Experiment.," *American Economic Review*, December 2004, *94* (5), 1717–1722.

Fudenberg, Drew and David K. Levine, *The Theory of Learning in Games*, 2 ed., Cambridge, MA: MIT Press, 1998.

Girvan, M. and Mark E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences of the United States of America*, 2002, *99* (12), 7821.

Goeree, Jacob K. and Charles A. Holt, "An Experimental Study of Costly Coordination," *Games and Economic Behavior*, May 2005, *51* (2), 349–364.

Goette, Lorenz, David Huffman, and Stephan Meier, "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups," *American Economic Review*, May 2006, *96* (2), 212–216.

Gunnthorsdottir, Anna, Daniel Houser, and Kevin McCabe, "Disposition, History and Contributions in Public Goods Experiments," *Journal of Economic Behavior & Organization*, February 2007, *62* (2), 304–315.

Hartley, Scott, "Kiva.org: Crowd-Sourced Microfinance & Cooperation in Group Lending," 2010. Harvard working paper.

Heap, Shaun P. Hargreaves and Daniel J. Zizzo, "The Value of Groups," *American Economic Review*, March 2009, *99* (1), 295–323.

Jenq, Christina, Jessica Pan, and Walter Theseir, "What Do Donors Discriminate On? Evidence from Kiva.org," 2012. Manuscript.

Knez, Marc and Colin Camerer, "Creating Expectational Assets in the Laboratory: Coordination in 'Weakest-Link' Games," *Strategic Management Journal*, Winter 1994, *15* (S1), 101–119.

_ and _ , "Increasing Cooperation in Prisoner's Dilemmas by Establishing a Precedent of Efficiency in Coordination," *Organizational Behavior and Human Decision Processes*, July 2000, *82* (2), 194–216.

Krippendorff, Klaus, *Content analysis: An introduction to its methodology*, 2 ed., Thousand Oaks, CA: Sage Publications, 2003.

Leskovec, J., L.A. Adamic, and B.A. Huberman, "The dynamics of viral marketing," *ACM Transactions on the Web (TWEB)*, 2007, *1* (1), 5.

Levine, David K., "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, July 1998, *1* (3), 593–622.

Liben-Nowell, David, Jasmine Novak, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins, "Geographic routing in social networks," *Proceedings of the National Academy of Sciences of the United States of America*, August 2005, *102* (33), 11623–11628.

Lin, John C., "Using Leaderboards to Motivate Contribution to Public Goods: A Field Study on Amazon," 2011. Working.

Lin, Mingfeng, Siva Viswanathan, and N. Prabhala, "Can Social Networks Mitigate Information Asymmetry in Online Markets?," in "Proceedings of the 30th International Conference on Information Systems" 2009.

Liu, Yang, Roy Chen, Yan Chen, Qiaozhu Mei, and Suzy Salib, ""I loan because...": understanding motivations for pro-social lending," in "Proceedings of the fifth ACM international conference on Web search and data mining" WSDM '12 ACM New York, NY, USA 2012, pp. 503–512.

Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schtze, *Introduction to Information Retrieval*, 1 ed., New York, NY: Cambridge University Press, 2008.

McKelvey, Richard D. and Thomas R. Palfrey, "Potential Games," *Quantal Response Equilibria for Normal Form Games*, July 1995, *10* (1), 6–38.

McLeish, Kendra N. and Robert J. Oxoby, "Identity, Cooperation, and Punishment," 2007. Institute for the Study of Labor Discussion Paper 2572.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, August 2001, *27*, 415–444.

Meer, Jonathan and Oren Rigbi, "Transactions Costs and Social Distance in Philanthropy: Evidence from a Field Experiment," 2011. Manuscript.

Menezes, Flavio M., Paulo K. Monteiro, and Akram Temimi, "Private provision of discrete public goods with incomplete information," *Journal of Mathematical Economics*, July 2001, *35* (4), 493–514.

Monderer, Dov and Lloyd S. Shapley, "Potential Games," *Games and Economic Behavior*, May 1996, *14* (1), 124–143.

Mysker, Michael B., Peter K. Olson, and Arlington Williams, "The Voluntary Provision of a Threshold Public Good: Further Experimental Results," in R. Mark Isaac, ed., *Research in Experimental Economics*, Vol. 6, Greenwich, CT: JAI Press, 1996.

Newman, Mark E. J., "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, 2006, *103* (23), 8577.

_ , *Networks: An Introduction*, 1 ed., New York, NY: Oxford University Press Inc., 2010.

Page, Talbot, Louis Putterman, and Bulent Unel, "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency," *The Economic Journal*, October 2005, *115* (506), 1032–1053.

Rabin, Matthew, "Incorporating Fairness into Game Theory and Economics," *American Economic Review*, December 1993, *83* (5), 1281–1302.

Selten, Reinhard, "Axiomatic Characterization of the Quadratic Scoring Rules," *Experimental Economics*, June 1998, *1* (1), 43–62.

Shang, Jen and Rachel Croson, "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods," *The Economic Journal*, October 2009, *119* (540), 1422–1439.

Shih, Margaret, Todd L. Pittinsky, and Nalini Ambady, "Stereotype Susceptibility: Identity Salience and Shifts in Quantitative Performance," *Psychological Science*, January 1999, *10* (1), 81–84.

Sobel, Joel, "Interdependent Preferences and Reciprocity," *Journal of Economic Literature*, June 2005, *43* (2), 392–436.

Tajfel, Henri and John C. Turner, "An Integrative Theory of Intergroup Conflict," in Stephen Worchel and William G. Austin, eds., *The Social Psychology of Intergroup Relations*, Monterey, CA: Brooks/Cole, 1979.

_ and _ , "The Social Identity Theory of Intergroup Behavior," in Stephen Worchel and William G. Austin, eds., *Psychology of Intergroup Relations*, Chicago, IL: Nelson-Hall, 1979.

_ , Michael Billig, R. Bundy, and Claude L. Flament, "Social Categorization and Inter-Group Behavior," *European Journal of Social Psychology*, April/June 1971, *1* (2), 149–178.

Tan, Pang-Ning, Michael Steinbach, and Vipin Kumar, *Introduction to Text Mining*, 1 ed., Boston, MA: Addison-Wesley Longman Publishing, 2005.

Tiebout, Charles M., "A Pure Theory of Local Expenditures," *Journal of Political Economy*, October 1956, *64* (5), 416–424.

Turner, John C., "Social Categorization and Social Discrimination in the Minimal Group Paradigm," in Henri Tajfel, ed., *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations*, London: Academic Press, 1979.

Ui, Takashi, "A Shapley Value Representation of Potential Games," *Games and Economic Behavior*, April 2000, *31* (1), 121–135.

Van Huyck, John B., Raymond C. Battalio, and Richard O. Beil, "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, March 1990, *80* (1), 234–248.

Watts, Duncan J., Peter Sheridan Dodds, and M. E. J. Newman, "Identity and Search in Social Networks," *Science*, May 2002, *296* (5571), 1302–1305.

Weber, Roberto A., "Managing Growth to Achieve Efficient Coordination in Large Groups," *American Economic Review*, March 2006, *96* (1), 114–126.

Yamagishi, Toshio and Toko Kiyonari, "The Group as the Container of Generalized Reciprocity," *Social Psychology Quarterly*, June 2000, *63* (2), 116–132.