

Houston Shaner
Economics Honors Thesis
Draft #4

A Laboratory Evaluation of Partner Selection in Group Lending Models

This study attempts to simulate the group loan model employed by many Microfinance Institutions (MFIs). Specifically, an experiment using Vanderbilt University undergraduates sought to determine if positive assortative matching (PAM) occurred in the context of a pool of risky and safe borrowers and a type of joint liability group loan. The simulation manipulated two independent variables in a 2x2 factorial design with four conditions. The first independent variable tested a fixed distribution of the group loan between partners versus a participant-selected distribution of the loan. The second independent variable tested learning across iterations of the lending simulation by manipulating the display of simulation results. Direct evidence from the simulation provides only limited support for the PAM outcome, though indirect evidence provides both further support and further qualifications. A 2x2 ANOVA reveals a main effect of the results display factor and an interaction between both factors. Overall, heavily qualified empirical support is found for PAM, and learning effects and other potential confounds seem to be present.

Introduction

Microfinance institutions (MFIs) have attracted a great deal of attention in part thanks to their mission of alleviating poverty but also because they often employ unconventional lending techniques in doing so. Though MFIs throughout the world employ a wide variety of lending mechanisms, group lending, in which two or more borrowers jointly take on liability for a loan, is perhaps the most well known of such mechanisms. Given that many poor people in developing countries lack access to capital, MFIs are often thought to be remedying a particular market failure by providing credit or financial services to those credit-worthy people who could gain from such opportunities but cannot acquire them.

Many potential benefits to group lending or joint liability contracts have been hypothesized; there is widespread belief that microfinance programs, and group lending in particular, exhibit exceptional performance in terms of welfare improvement and loan repayment (see Morduch (1999) for a relevant discussion). Repayment rates for microfinance group lending programs are often as high as 98%. To explain such real-world results, the theoretical microfinance literature often points to the ability of group lending to overcome one or more of

the following problems growing out of an information asymmetry: adverse selection, moral hazard, and strategic default by successful borrowers (Guttman 2006). This theoretical literature, though, is highly divided in its discussion of just *how* group lending resolves such problems. To avoid a wide-ranging discussion of such posited explanations of group lending's success, consider the categorization of group lending theories by Cassar et al. (2007) into three conceptual mechanisms for addressing information asymmetry: relational social capital, informational social capital, and the "pure" incentive effects of joint liability (p.F86). This paper is concerned primarily with this latter category and its ability to address the issue of adverse selection. More specifically, the experimental method discussed later is designed to test and extend the incentive structure discussed in a theoretical paper by Ghatak (1999). While the technical model constructed by Ghatak will be discussed and extended in more detail in a later section, here I will provide a brief overview of the paper's assumptions and conclusions.

The scenario developed by Ghatak (1999) provides a rationale for the efficiency of group lending mechanisms. He assumes at the outset that monitoring costs by a financial intermediary like an MFI are prohibitively high given a lack of information about local borrowers and limited ability to impose non-monetary default penalties. Moreover, the lending bank faces an information asymmetry. There are two types of borrowers, risky and safe, with lower and higher chances of successful projects, but despite knowing the overall distribution of borrowers, the bank lacks information regarding any given borrower's type while the locals all know each other's chance of success. Ghatak shows that this information asymmetry produces adverse selection akin to Akerlof's lemon market, driving up the interest rate to the point where many safe borrowers exit the market and risky borrowers dominate the borrowing pool.

To compensate, Ghatak's lender introduces a joint liability contract for two person groups. If her project succeeds, a borrower pays a fixed fee, plus a joint liability payment if her partner fails. Since the expected joint liability penalty of any borrower depends upon the probability of success of her partner, everyone prefers a safe partner. That partner, of course, would also prefer a safe borrower, and Ghatak shows that the joint liability contract creates a separating equilibrium in which borrowers only pair with another borrower of their same type, a phenomenon known as "positive assortative matching." Thus, though the bank imposes a uniform interest rate, risky borrowing groups effectively face a higher borrowing cost due to a higher likelihood of paying the joint liability penalty. This outcome approximates the socially efficient situation in which the bank could discriminate between borrowers. Safe borrowers are attracted back into the market; the overall borrowing pool improves; and the bank charges a lower interest rate.

In particular, I seek to investigate empirically several questions regarding the famed group lending mechanism in the context of Ghatak's model. First, do joint liability provisions empirically induce positive assortative matching? Second, given group lending, what conditions regarding the contract or group composition are necessary for the program to be successful in some sense? This second question will be considered primarily in the context of different distributions of the group loan.

The theoretical literature on group lending beyond the examples previously discussed is well developed and ranges over a wide number of topics, but "empirical work testing the effect[s] of the specific instrument of joint liability...has lagged behind" (Ghatak 1999). This may be due to a reliance on econometric studies using large-scale data sets derived from direct field work (see, for example, Townsend & Ahlin (2007)). Given the rural and remote nature of

many microfinance programs, one might expect the flow of such data to be infrequent and lumpy, with the result that the theoretical literature moves forward far faster. Additionally, some variables that might play crucially important roles in theoretical models (for example, the riskiness of a certain borrower's investment) are virtually unobservable by researchers hoping to collect data in the field (Abbink et al. 2006).

In contrast to the type of studies exemplified by Townsend and Ahlin, other researchers have taken to using experiments or microfinance-related games to test theoretical hypotheses. For example, Irlenbush et al. (2006) used a laboratory experiment to test the effects of the level of the interest rate on repayment and group dynamics under a joint liability contract. Similarly, Karlan (2005) conducted an experiment (albeit with Peruvian microcredit borrowers) to test the relationship between interpersonal trust and group lending success. Abbink et al. (2006) even use an experiment with undergraduate students to examine dynamic incentives and group size in microfinance.

This thesis will adopt the empirical methodology of these latter researchers and conduct a microfinance-related experiment using lending games and undergraduate participants. The experimental approach, of course, cannot simulate cultural or other conditions that may differ drastically in field work, but such non-economic differences may violate the assumptions of models based upon rational choice theory anyway. A laboratory experiment does, however, provide a much more practical means of testing the sorts of hypotheses found in Ghatak's model, allows the experimenter to control and observe certain key variables, and provides some novel data where many field data sets may already be thoroughly hashed and rehashed. Laboratory experimentation is particularly appropriate for several of the otherwise hard-to-observe variables featured in the Ghatak model.

Motivated by the questions concerning positive assortative matching above, the experiment reported here employed a 2x2 factorial design and hence four experimental conditions. The first independent variable, distribution of the total group loan, seeks to test an extension of the Ghatak model to certain types of side payments discussed in more detail in the next section; the manipulation of the second independent variable should provide evidence for or against learning across the course of the experiment as a way to eliminate the potentially confounding influence of participant misunderstandings or mistakes in a one-shot experiment.

This paper will proceed through four further sections. The second section will elaborate on the Ghatak model of joint liability and will adapt the model to the particular design of the experiment. This section thus provides the hypotheses tested in the experiment. The third section will discuss the experimental design in detail. The fourth section will develop a game theoretic model of the experimental design. The fifth section presents and analyzes the results of the experiment. The sixth section provides a discussion of the results and concludes the paper. Ultimately, the results of the experiment provide some support for the Ghatak model, though such support is rather weak and points to the importance of considerations external to the model. The qualifications of the evidence in favor of positive assortative matching may provide fruitful guides for further empirical work.

Adaptation of the Ghatak Model

The theory of joint liability presented in the Ghatak paper requires several key assumptions concerning the economic environment. There are two types of borrowing entrepreneurs in the local area distinguished by the probability of success on their investment projects. The first type of borrower, “safe” borrowers, have a probability p_s of succeeding in their individual projects where $0 < p_s < 1$. The second type of borrower, “risky” borrowers, have a

probability p_r of project success where $0 < p_r < p_s < 1$. All borrowers are risk neutral and require minimum inputs of capital and inputs of labor for each investment project. Each borrower supplies the amount of labor needed for the project but has zero wealth. Consequently, borrowers require a loan to secure the needed capital input to their projects but cannot supply collateral.

The outcome of a borrower's investment project is either success or failure. The outcome of any given borrower's project is not influenced by the likelihood or outcome of any other borrower's project. In the context of this experiment, a successful investment project will produce a return $R(p)$ where $R(p)$ is a positive function of a borrower's probability of success. If a project is unsuccessful, the borrower enjoys no return on the investment, and $R(p) = 0$. Moreover, the returns to the two types of projects are such that the expected returns for safe and risky borrowers all investing a given amount are equivalent, or $p_s R(p_s) = p_r R(p_r)$.

A significant information asymmetry exists between borrowers and the lender. The local pool of prospective entrepreneurs is evenly divided between safe and risky types. All borrowers in the area know both the distribution of borrowers and the particular type of each individual borrower. While the lending institution entering the local area is aware of the distribution of local borrowers, it cannot identify the probability of success of any given borrower. Additionally, the lender can only observe the outcome of a borrower's investment project (and do so with full certainty at no cost) but cannot observe the returns to the investment. Hence, the contract used by the lender takes the form of debt rather than equity (Diamond 1984). Enforcement of a contract is costless and guaranteed, eliminating the prospect of strategic default.

The special sort of debt contract modeled by Ghatak, a joint liability group loan, is an agreement by the lender to provide a fixed and uniform amount of capital L to a group comprised of two borrowers. If the project of one of these borrowers is unsuccessful, she defaults on the

loan and owes the bank nothing due to a limited-liability constraint. If her project is successful, however, she owes the bank a standard interest rate r . The contract also contains a joint liability payment c that a successful borrower will pay if her partner's project is unsuccessful and her partner is forced to default.

Thus the expected utility of a borrower with probability of success p pairing with a partner with probability of success p' is modeled by the equation

$$EU_{p, p'}(r, c) = pR(p) - rp - cp(1 - p')$$

From this equation, Ghatak derives his first lemma (Lemma #1):

A borrower of any type prefers a safer partner, but the safer the borrower herself, the more she values a safer partner.

The proof of Lemma #1 is fairly straightforward. Ceteris paribus, the net gain to a borrower with probability p from having a safe partner with probability of success p_s compared to having a risky partner with probability of success p_r reduces to

$$EU_{p, p_s} - EU_{p, p_r} = cp(p_s - p_r)$$

As a result, the borrower with probability p should be willing to pay up to $cp(p_s - p_r)$ for the opportunity to partner with a borrower of probability p_s over a partner of probability p_r .

Clearly, then, a borrower whose probability $p = p_s$ will be willing to pay more for a safer partner than a borrower whose probability $p_r < p_s$.

Ghatak claims that as a result of Lemma 1 safe borrowers will outbid risky borrowers for the opportunity to partner with other safe borrowers. What medium of exchange the borrowers (who have zero wealth at the outset) use to make "bids" is somewhat unclear in the Ghatak paper. Ghatak has a footnote discussing "side payments" that borrowers might make to one another. Here is the footnote (footnote 13):

Since we assume borrowers have no wealth that can be used as collateral, when we talk about side payments among borrowers, we mean that these transfers take forms that are not feasible with the bank. For example, borrowers within a social network can make transfers to each other in ways that are not possible with an outsider (namely, the bank), such as providing free labor services, or writing contracts based on the output (as opposed to outcome) of their projects.

Incorporating the first type of side payment (non-monetary incentives) into an experiment raises theoretical and experimental design issues that could be avoided with other approaches to willingness to pay. The second type of side payment mentioned in the footnote essentially amounts to offering equity in one's investment project. Since all borrowers have identical expected returns on their investments, a given share of equity in one type of project should have the same expected return as a similar share of equity in the other type of project. However, equity contracts between local borrowers raise the question of contract enforcement after the realization of the stochastic investment process. Enforcement of such borrower-borrower contracts is not explicitly modeled in the Ghatak paper. Moreover, there might be simpler, more logistically efficient methods of incorporating side payments into an experimental design.

This experiment takes a slightly different approach to side payments and willingness to pay by allowing borrowers to divide ex ante the total amount L of the group loan provided by the lender before the realization of investment outcomes. That is, a borrower can outbid another borrower for a certain loan partner by offering the target partner a greater share of the group loan to be invested in the target's project.

The experimental design breaks down the return variable $R(p)$ in the Ghatak model into two components. The first component $B(p)$ multiplies the amount invested by the borrower by a positive factor if the borrower is successful. If the borrower is unsuccessful, $B(p) = 0$. Again, to

equalize expected returns for borrower types, $p_s B(p_s) = p_r B(p_r)$. The second component M is the amount of L that the borrower has agreed to receive from the division of the group loan.

Since the posted offer system (discussed below) used in the experimental design involves proposer and responder aspects of a participant's role in addition to the safe/risky borrower distinction, proposers in the experimental design can "bid" for the opportunity to pair with a given responder by offering a larger share M to the responder. Thus, willingness to pay should be expressed in the experimental design by the variable M .

Ghatak's first lemma should still hold in this particular experiment. It has already been established above that, given fixed and equal divisions of the group loan (an implicit assumption in Ghatak's paper), safer proposers should value pairing with safe responders more than risky proposers. Since all borrowers have the same expected return on investment, the expected decrease in utility from offering a certain M to a responder is identical for both risky and safe proposers. Thus, risky proposers have no incentive to offer a larger M to safe responders than that offered by safe proposers. Consequently, safe proposers should outbid risky proposers for the opportunity to pair with safe responders, and the positive assortative matching result derived in Ghatak should still hold. Ultimately, then, the experiment will test whether such matching does occur.

Rewriting the expected utility functions for this specific experiment requires two equations, one for proposers and one for responders in the posted offer system. If M denotes the amount of the loan L that a proposer is willing to provide to a responder, then the expected utility for a proposer with probability p pairing with a responder with probability p' can be written as

$$EU_{p,p'} = pB(p)[L - M] - rp - cp(1 - p')$$

while the expected utility for the responder can be written as

$$EU_{p', p} = p' B(p') [M] - rp' - cp'(1 - p) \quad .$$

Experimental Method

Factorial Design

The experiment to test positive assortative matching utilizes a mixed 2x2 factorial design. The first, within-subjects factor is the division of the group loan (analogous to the M variable discussed in the model) which takes on two levels. In the first level, the division of the group loan is fixed exogenously at a 50-50 split between partners. That is, one partner receives half of the loan and the other partner receives the other half of the loan. In the second level, the participants may determine the division of the loan using the posted offer system. Here, participants in the proposer role can set a value to M in their offer to the pool of responders. Both levels of this factor are experienced by all participants. Boundary constraints are imposed on the possible values of M proposed so that all parties at least break even if successful and forced to repay the bank.

This offer factor serves two purposes in the experimental design. First, taking the broad theoretical categories of Cassar et al., the offer factor can help illuminate whether “pure” incentive effects, in this case willingness to pay as expressed by M , or more “social” considerations not as well considered in rational choice theory motivate participant decisions. My particular concern derives from a potential confound related to categorizing participants into distinct borrower types. The mere act of dividing participants into distinct groups creates the potential for in-group bias in partner selection,¹ but it may be possible to discern the importance of willingness to pay if participants demonstrate positive assortative matching when they can

¹ Work by Tajfel (1982) and Tajfel & Billig (1974) on minimal groups suggests that even the act of dividing participants into random or arbitrary groups will induce participants to strongly favor the members of their own group over those in the out-group. Other work (cf. Hogg & Abrams 2001) has found that participants will even forego monetary gain to deprive out-group members of higher monetary gains.

negotiate the division of the loan but not when they are selecting partners only by viewing a borrower's type. Secondly, manipulation of the offer variable provides a way to test the particular sort of side payment mechanism discussed in the model section. If the Ghatak paper is correct that safe borrowers are willing to pay more for safe partners, this hypothesis should be supported in the experimental outcome and can be evaluated against a control condition in which participants do not explicitly produce bids or other expressions of willingness to pay.

The second, between-subjects factor is the display of investment project results for a participant and her chosen partner and takes on the self-explanatory levels of "shown" and "hidden." This factor is manipulated to help determine whether learning effects are present across the periods of the simulation. Participants may be influenced as much by the ex post outcomes of their respective projects in the last period as by the ex ante probabilities of the next period. For example, a participant who initially adopts the strategy of homogeneous matching predicted by Ghatak may become discouraged and switch strategies following a series of unlucky outcomes for his project or that of his partner. For this between-subjects factor, half of the participants experience the "shown" level while the other half of the participant pool experiences the "hidden" level.

This factor is not designed to detect all instances of learning effects, and some such effects are still possible. Due to logistical concerns, only partial counterbalancing (with respect to participant roles) was employed. There are roughly 30 iterations of the lending simulation in each condition, and, though the experiment cannot explicitly identify sources of learning outside of the display of project results, if 30 iterations are not enough to produce sustained convergence towards positive assortative matching, the results would raise questions about the real-world applicability of the Ghatak model.

Logistics & Materials

Twenty-four participants were recruited for the experiment, which took place in a Vanderbilt University computer lab and lasted approximately an hour and a half. Participants experienced the experimental conditions through the use of the computer program *Z-tree* (Fischbacher 2007). Each participant was identified by the identification number of the computer at which he or she sat. Initially, participants were randomly assigned to a borrower type, role within the posted offer system, and between-subjects (results display variable level) based on this identification number.

Participants carried out a group lending simulation using the *Z-tree* software. For each experimental condition, participants completed two *Z-tree* treatments (a “treatment” in *Z-tree* is the completion of a set number of iterations or periods of a program). In the first treatment, each participant was assigned by identification number to one of four roles: safe proposer, risky proposer, safe responder, or risky responder. “Safe” roles correspond to Ghatak’s safer borrowers and have a higher probability of success on their investment projects; “risky” roles have a lower probability of success but equal expected return on their projects. Proposer roles created offers for division of the loan in the posted offer system of the program, while responders chose among the listed proposals displayed by borrower type and the M variable (amount of loan offered to the responder). A responder accepting another participant’s proposal would take on that proposer as a partner in a two-person group loan. For participants in the “results shown” manipulation, the outcomes and profits of the relevant projects were displayed after the completion of the partner-matching stage. Participants in the “results hidden” manipulation were directed to a brief waiting screen after all proposals and acceptance of proposals had been made through the posted offer (though of course the program still calculated the stochastic outcomes,

repayments, and profits). This first lending simulation was repeated for 15 periods in which new proposals and groups were formed.

In each condition, participants completed a second treatment in which each participant switched to the opposite borrower type and role within the posted offer system. For instance, safe proposers became risky responders, while risky proposers became safe responders. In this treatment, participants completed another 15 periods of the lending simulation. Thus, participants experienced some degree of counterbalancing with respect to borrower type and role.

The lending simulation proceeded sequentially through several stages. When the Z-tree treatment began, participants initially saw an introduction screen with written instructions for the simulation. These instructions included information on borrower types, probabilities of success, the loan contract, the matching system, the investment projects, and identification of borrower types by computer ID number. Once a participant had read and understood the instructions, he would click a button to proceed to the next stage. This following stage displayed the participant's own identification number as well as his own borrower type (denoted by either a "1" for safe or a "2" for risky). After taking in this stage, the participant would click a button to continue to matching process. If a participant was assigned to the safe proposer role, he would enter the safe proposer stage where he would be asked to enter the amount of the group loan (out of a total of 200 "points") he would be willing to offer to potential partners. If a participant was playing as a risky proposer for the treatment, he would proceed instead to the risky proposer stage where he would face the same command. Proposers could type a numeric amount into an input box and confirm their proposal by pushing a button. Participants were given 45 seconds to complete this proposal stage. For the fixed share offer conditions, all proposers were asked to enter "100" into the input box and could not confirm a proposal of a value other than 100 (otherwise an error

message would be displayed). For the variable share offer conditions, minimum and maximum bid constraints were imposed so that no player would be unable to repay in full interest and the liability penalty to the bank (i.e. receive such a low share of the loan that even a successful project would generate losses).

During this proposal stage, participants playing as either safe or risky responders would view a screen with a short message asking them to please wait for the next stage. After all proposers had confirmed their proposals, the responder stage of the posted offer system would begin. Participants assigned to safe responder roles would enter this stage first under the intuition in the Ghatak paper that all proposers are really bidding for the preferred safe partners. These safe responders were given a 30 second period to select a proposal before participants playing as risky responders were allowed to enter the stage and select a proposal to accept. A responder in this stage would see a list of potential contracts to accept characterized by two pieces of information, each proposer's identification number and the amount of the 200-point loan offered to the responders by each proposer. Responders could infer each proposer's borrower type by whether the identification number of the proposer was even or odd. Responders in the stage would select a proposer as a partner by clicking on a proposed contract in the list and confirming acceptance via a button. Accepting would pair the two players in a joint liability contract for the remaining stages. Because there was a one-to-one correspondence between all four roles in the experiment, every participant would have a partner at the end of the matching process and could not choose to continue without a partner.

The next stages displayed the results of the investment process. For participants in the "results shown" conditions, the first of these stages would display the outcome of their respective investment projects. Project outcomes were determined by sampling a normal random variable

with success cut-offs for safe and risky borrower types corresponding to their respective probabilities of success. That is, a borrower would need a realization of the random variable above a certain number to achieve “success” in an investment project. The screen would either display a message of success in the investment project or a message of failure in the investment project. The screen would also display the return on investment and interest owed to the bank. Both of these values were zero for failed projects. The next stage showed the results of the investment project of a participant’s partner. The random variable realizations for this stage were not actually those of the partner selected in the posted offer system but did simulate an independent realization of the random variable using the probability of success of the borrower type of the partner selected. In other words, if participant #1 had selected participant #2 as a partner, this stage for participant #1 would not necessarily show the same outcome for the “partner’s investment project” as that displayed in the previous stage on participant #2’s computer, but the outcome would be based on the probability of success associated with participant #2’s borrower type. This stage also displayed the joint liability penalty amount (if any) owed to the bank, the participant’s total repayment, and the participant’s total profit for the period. A period, again, consisted of one iteration of the lending simulation. Participants in the “results hidden” condition would only see a waiting screen displayed while those in the “results shown” condition saw their results. After all participants had completed the results stage, the simulation would begin again in the next period with the instructions screen until 15 periods had been completed for each treatment. Short breaks were provided between each treatment.

After being consented by the Principal Investigator, each participant was guaranteed a fee of \$10 in compensation for his or her time. As an incentive to maximize one’s points in the simulation, a lottery system was also added where three participants were randomly drawn as

winners at the end of the experiment. The lottery winners split a fixed cash prize based on the following process: a winner received the same proportion of the cash prize as the ratio of his total accumulated profit during the experiment to the total accumulated profit during the experiment of all three winners.

Game Theoretic Model of the Experimental Design

In the Ghatak paper, the timing and sequencing of actions by different agents is not explicitly considered in the description of an equilibrium. However, the experiment design and technology clearly add issues of sequential moves and strategic choices to the process of choosing partners for group loans. It is necessary therefore to further supplement the theory of the matching process developed earlier to account for the considerations imposed by the details of the experimental procedure.

A game-theoretic pay-off table can be a useful framework for examining how participants would be expected to make strategic choices in the lending simulation. It was suggested earlier that safe responders will essentially dictate whether positive assortative matching occurs since they should be the implicit targets of all proposals. Or, in a more flexible setting outside the experimentally-constrained posted offer system, proposers might be expected to make proposals *only* to safe responders and not more general posted offers capable of being accepted by anyone. Thus, safe responders would have the first pick of proposer partners, leaving the risky responders with whichever proposers were not chosen, and would presumably pick the proposal offering the highest expected utility. All proposers would strictly prefer having a safe responder accept their proposal of a given amount, so the key question then concerns what strategies would safe and risky proposers adopt in the formulation of their proposals in the posted offer system.

A continuum of proposals is technically possible, but strategies at the extremes become the major options for a player. Consider a risky proposer. He knows that a safe proposer is willing to pay more to a safe responder than he himself is willing to pay. Thus, the risky proposer knows that the safe proposer can and will bid \$1 more for the safe responder than the maximum he himself is willing to bid, denoted here as b^r . In fact, at any bid below b^r the risky proposer can be similarly outbid, so he will consider the minimum possible bid, denoted b_r . This minimum bid might give up on trying to secure a safe partner but allows the proposer to keep more of the group loan to invest in his own project. If the risky proposer keeps moving his bid lower towards the minimum b_r , eventually the safe proposer will be able to achieve a winning bid by bidding only just higher than b_r . However, bidding this low risks the possibility that the risky proposer will raise the bid back towards his maximum willingness to pay, at which point the safe proposer will bid just more than b^r . The safe proposer should therefore expect to bid near the minimum b_r or near the maximum b^r .

We can thus construct the pay-off matrix for both proposer types as shown in Table 1. All pay-offs listed in Table 1 are the net increase or decrease in expected utility relative to making a proposal of the b_r and ending up with a risky partner from that proposal. The top line in each cell represents the pay-off for the safe proposer and the bottom line the pay-off for the risky proposer.

Table 1: Proposer Pay-off Matrix
(Relative to a minimum bid of b_r)

		Risky Proposer	
		Max b^r	Min b_r
Safe Proposer	Max b^r	$+cp_s(p_s - p_r) - (b^r - b_r)p_sB(p_s),$ $-(b^r - b_r)p_rB(p_r)$	$+cp_s(p_s - p_r) - (b^r - b_r)p_sB(p_s),$ 0
	Min b_r	$0,$ $+cp_r(p_s - p_r) - (b^r - b_r)p_rB(p_r)$	$+cp_s(p_s - p_r),$ 0

Consider the decision of a safe proposer facing a proposal of b^r by a risky proposer. If the safe proposer makes a proposal of b_r (as in the bottom left cell), he could potentially improve his outcome by raising his own proposal to b^r . If the safe proposer does so, though, the risky proposer could be better off by lowering his own bid to b_r ; consequently, the risky proposer should expect that his maximum possible bid of b^r will be unsuccessful and achieve the outcome of the top left cell. When the safe proposer is facing a proposal of b_r by the risky proposer, the safe proposer is assured of having his proposal accepted by the safe responder and hence will enjoy some positive utility gain. However, he can further supplement this gain by lowering his own bid to the minimum b_r and then invest more of the group loan in his own project. Thus, the ultimate outcome predicted by Table 1 is that of the bottom right cell, where the risky proposer makes the minimum proposal b_r and the safe proposer makes a proposal just above b_r .

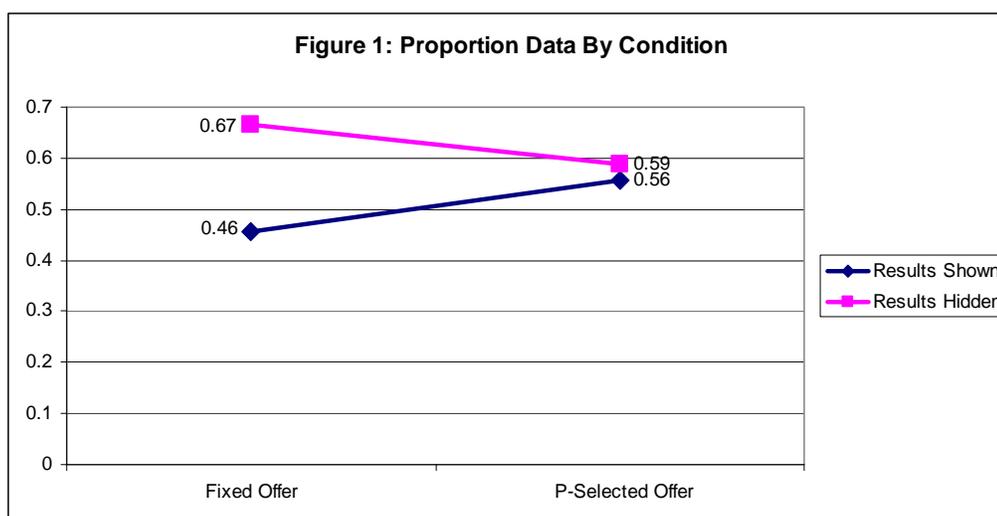
Data & Results

Data

The Z-tree program produces a large array of data from each treatment, most of which relates to the operation of the simulation. For example, Z-tree will include the final profit per participant, each participant's profit per round, time taken by a participant to create or accept a

proposal, and the parties to and terms of a contract dividing the group loan. To test the major hypothesis of the Ghatak model, that joint liability contracts will induce positive assortative matching, I primarily examine only the information produced in Z-tree's Contracts table. The raw data on each contract includes the Subject ID number for both the "buyer" and "seller" side of the posted offer system. I encoded each contract as a binomial outcome for positive assortative matching. "Success" represented a contract with two borrowers of the same type; "failure" represented a contract with heterogeneous borrower types.

Since there are 12 participants in each treatment and each contract is between a pair of participants, there are 6 contracts per period. There are 15 periods per Z-tree treatment, and 2 treatments per experimental condition. With four experimental conditions, there are 720 (6x15x2x4) contracts in the entire experiment and 180 contracts in each condition. In Figure 1, I provide a graph of the proportions data for each experimental condition. The analysis of the data reported below was carried out using the statistical package Stata[®].



Confidence Intervals

The fundamental prediction of the Ghatak model is that borrowers will self-sort into homogeneous borrowing groups. The immediate intuition for a statistical test, then, is to use

binomial proportions in standard hypothesis testing related to the probability of homogenous pairing. However, the Ghatak model does not explicitly incorporate unobservable sources of variation across individuals, and, as an example of rational choice theory, should predict that the prevailing strategy in this game (positive assortative matching) should occur in all instances. Thus, if one were to draw from the Ghatak paper an implicit null hypothesis for the parameter p expressing the probability of homogeneous matching in the population, one would be forced to use a null hypothesis $H_0: p = 1$. Obviously, a value at the boundary for the null hypothesis implies zero variance for the binomial random variable and makes hypothesis testing problematic. If there is any variation across individuals at all (or even any real source of stochastic variation), an experimenter would only need to keep sampling individuals until he found one instance of heterogeneous matching in order to reject the null.

Due to the infeasibility of hypothesis testing in this situation, I attempt to investigate positive assortative matching in the experiment using the alternative approach of confidence intervals. While confidence intervals do not provide the sharp distinction of either rejecting or failing to reject the null found in hypothesis testing, we can nevertheless glean some information regarding the proportion of positive assortative matching and how it compares to the Ghatak model. I thus examine positive assortative matching in each experimental condition using confidence intervals.

Even with confidence intervals, though, the Ghatak model presents a significant obstacle to the standard approach. Given that the model implicitly predicts that all contracts will exhibit positive assortative matching, i.e. the proportion of positive assortative matching will be equal to one, it is plausible to believe *a priori* that the observed proportion of positive assortative matching in the experiment will approach the boundary of one. Approaching the boundary,

though, creates serious problems for the use of the standard Wald confidence interval. The standard rule of thumb is that the normal approximation for the binomial holds only if $\min\{np, n(1-p)\} \geq 5$, a clear problem for $p = 1$ (Olivier & May 2007). The Wald interval, then, is valid only for p over the interval $(0,1)$. Moreover, Brown et al. (2001) demonstrate that the coverage properties of the Wald interval can vary wildly and erratically, especially for p near but not on the boundary. Brown et al. also show that the interval's coverage varies unpredictably more generally for "unlucky" (i.e. unpredictable) pairs of parameters. Considering the litany of other research pointing to potential problems with the Wald interval in binomial proportions near the boundary (cf. Agresti & Coull 1998 and Hanley & Lippman-Hand 1983), the use of the Wald interval for the data in this experiment does not seem appropriate *a priori*.

The argument could be made, after seeing the raw proportions data above in Figure 1, that concerns over the suitability of the Wald interval can be dismissed since the observed proportions are not particularly close to the boundary. However, proper experimental and analytical procedure dictates that the tests of the data be decided *a priori*; in the words of Olivier and May, "the data should not dictate the method of interval construction" (p.179). Consequently, the decision was made beforehand to use an alternative to the standard Wald interval.

Such an alternative discussed by Brown et al., Agresti and Coull, and Olivier and May is the interval developed by Wilson (Wilson 1927). Though two-sided, near the boundary the Wilson interval functions as a one-sided interval with a limit at the boundary, allowing the Wilson interval to handle extreme values of the parameter p without the problems of the Wald interval. Farther away from the boundary, the Wilson interval converges to the Wald interval as sample size increases, so the Wilson interval is not inappropriate *a priori* if we do not expect the

observed proportions to fall away far from 0 or 1 (Olivier and May 2007). I therefore report the Wilson 95% confidence intervals for each experimental condition in Table 2.

I have chosen here to examine the proportion of positive assortative matching for each experimental condition. Obviously, these proportions both pool the data in each condition across participants and periods and also disaggregate the data for the entire experiment. Pooling the data across periods raises the issue of learning effects discussed in the section on experimental design; however, the design of the experiment is constructed to allow for some insight into learning effects from the analysis of variance conducted below. The combination of the results of the confidence intervals and the ANOVA should thus help direct our intuitions towards what is happening with participant decisions across periods. Moreover, even building a case that positive assortative matching does or does not occur in each experimental condition as a whole should shed some light on the empirical validity of the Ghatak thesis. On the other hand, one might claim that the existence of positive assortative matching should be examined for the experiment as a whole across conditions. Since there is reason to believe that the manipulation of the independent variables in the experimental design, though, might produce significantly different outcomes in the response variable, aggregating the data across the entire experiment would quite clearly lose information related to the specific conditions under which the positive assortative matching hypothesis holds.

Table 2: Wilson 95% Confidence Intervals
(Proportion of Positive Assortative Matching)

<i>Condition</i>	<i>Obs</i>	<i>Mean</i>	<i>Std. Err.</i>	<i>Wilson Interval</i>	
With Results, Fixed Offer	180	.4556	.0371	.3845	.5285
With Results, Var. Offer	180	.5556	.0370	.4826	.6262
No Results, Fixed Offer	180	.6667	.0351	.5950	.7314
No Results, Var. Offer	180	.5889	.0367	.5159	.6582

As shown in Table 2, the first two conditions (in which the results of a participant's and her partner's investment project are displayed) produce confidence intervals that contain $p = .5$, which is the proportion that would be expected if matching occurred randomly rather than systematically. Not only are the results of these two conditions indistinct from the random case, but neither condition comes close to the upper boundary of 1 for p , implying that random matching seems to prevail over systematic positive assortative matching in both conditions. In contrast, the two conditions in which project results are not displayed produce intervals that fall between (and do not include) $p = .5$ and $p = 1$. In both such conditions, positive assortative matching is more likely than simple chance but does not occur at nearly the boundary frequency predicted in Ghatak. Thus, these last two conditions seem to provide qualified support for the Ghatak hypothesis. In no case does heterogeneous matching seem to be the most likely strategy adopted by participants.

Matching – Analysis of Variance

I next attempt to determine how the experimental manipulation of the independent variables affects the likelihood of positive assortative matching. Again, the experimental design manipulates both the display of project results and the ability of participants to vary their proposed offers; the former variable is included to test the presence of learning effects while the

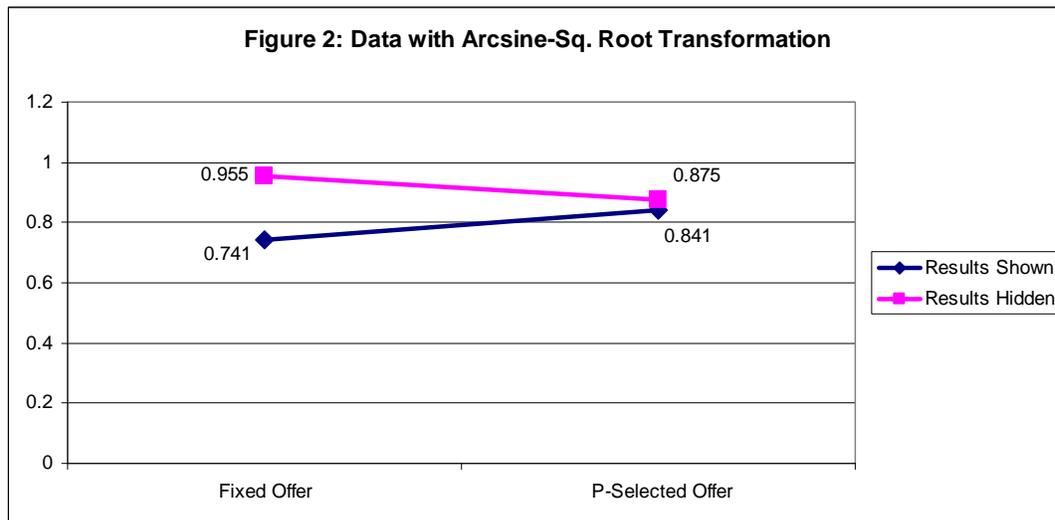
latter is meant to shed light on the relative importance of “pure” incentive effects as embodied in differences in monetary proposals. Thus, determining whether significant differences, in terms of borrower matching, exist across the manipulations of each of these independent variables helps impart the importance of potentially confounding variables like learning effects and social factors.

To determine if such differences exist, I use the analysis of variance (ANOVA) technique for a mixed factorial design to evaluate the presence of main effects of the independent variables as well as possible interaction effects. During the discussion of the ANOVA, I refer to the between-subjects variable (whether or not project results were shown) as *Showresults* and the within-subjects variable (whether the division of the group loan was fixed or participant-chosen) as *Chooseoffer*. The interaction of the two main variables is labeled *Showresults*Chooseoffer*. Both variables, again, have two levels.

The reader may have noticed that a binomial dependent variable might violate one or more of the assumptions of ANOVA. The analysis of variance assumes homogeneity of variance across conditions; since the variance of a binomial random variable is not independent of the mean, it's is all too possible that the assumption of homogeneity of variance does not hold. This problem is especially acute for situations where the proportion is near the boundary (Mimmack et al., p.334). In addition, ANOVA assumes normally distributed data. In response to such concerns about violation of assumptions, multiple sources recommend using an arcsine-square root transformation on proportion data to homogenize the variance and better approximate the normal distribution (cf. Cardinal et al. 2006 and Mimmack et al. 2001). If X is a binomial random variable expressed as a proportion, the arcsine-square root transformation, as the name suggests, produces the new variable X' pursuant to the equation:

$$X' = \arcsine(\sqrt{X})$$

Henceforth, the dependent variable used for ANOVA will be referred to as *Transform*. Figure 2 provides a visual representation of the transformed data. As with Figure 1, a quick inspection of the slopes in the graph leads one to expect an interaction effect between the main variables. It also appears likely from Figure 2 that there are also main effects present, though it is difficult to determine by inspection if the main effects are significant given the scale of the graph.



The output from Stata includes epsilon estimates, providing some information as to whether the arcsine-square root transformation is successful in avoiding violations of the assumptions of ANOVA. The two types of epsilon estimates, Greenhouse-Geisser and Huynh-Feldt, essentially denote whether the assumptions of normality of the data and homogeneity of variance across groups (assumptions that might together be packaged as “sphericity²”) have been met. Stata returned estimates of 1.0 for both the Greenhouse-Geisser and Huynh-Feldt estimates of epsilon, so no corrections to the p-values reported below were needed. Epsilon estimates below 1.0 would imply that the p-values reported would need to be adjusted upwards.

² Cf. “An Introduction to Sphericity,” located here: <http://homepages.gold.ac.uk/aphome/spheric.html>, for a more thorough explanation of epsilon estimates and sphericity in the context of ANOVA. This is one of a very few thorough, intelligible explanations of sphericity that I can find.

Table 3 reports the results of the mixed design factorial ANOVA using 48 observations. Each observation consists of the value of *Transform* for a given subject experiencing one of the levels of *Chooseoffer*, with the 24 subjects evenly divided between levels of *Showresults*. Division of the group loan as expressed in *Chooseoffer* has no significant effect on the transformed proportion of positive assortative matching at conventional levels of significance; display of the investment project results, though, has an effect on the transformed proportion of positive assortative matching at the 0.01 level. There is also a significant interaction between the two variables (at the 0.05 level of significance) such that allowing the participants to choose the division of the loan decreased positive assortative matching in the condition with hidden project results and increase positive assortative matching in the condition with project results displayed.

Table 3: ANOVA Results

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Between-subjects	.2612	23			
<i>Showresults</i>	.2612	1	.2612	12.02	0.0022***
Subjects WG	.4781	22	.0217		
Within-subjects	.1136	24			
<i>Chooseoffer</i>	.0025	1	.0025	.013	0.7200
<i>Chooseoffer*Showresults</i>	.1111	1	.1111	5.80	.0249**
Residual	.4217	22	0.0192		
Total	1.275	47			
Adjusted R-squared =	.2933				

Matching – Effect Size & Post-Hoc Comparisons

The omnibus F values reported in Table 3, i.e. the tests for whether the manipulations of the *Showresults* and *Chooseoffer* variables produce significant differences in matching, only speak to whether or not a main or interaction effect exists at all and provide no information on

the size of any of effects found. So, if the ANOVA results can show whether an independent variable like *Showresults* does significantly influence how borrowers match, another method is needed to estimate how much of an influence the variable has on the results. To that end, and based the results of the omnibus F tests in the ANOVA, I estimate the size of the effect of the *Showresults* and *Chooseoffer*Showresults* variables using the ω^2 (omega-square) statistic. Because the *Chooseoffer* variable by itself is not significant in the ANOVA results, I do not calculate its effect size.

The omega-square statistic estimates effect size in terms of the proportion of variability of the dependent variable accounted for by a given main or interaction effect (Sheshkin 2007, p.514). By Cohen's classification, *Showresults* has a large effect size ($\omega^2 > .14$) on the dependent variable, explaining roughly 18.7% of the variability of *transform* in the population. This result is relatively clear evidence that participants' learning across iterations of the lending simulation had a practically significant impact on borrower matching. Under the same classification, the interaction between the variables *chooseoffer* and *showresults* has a medium effect size ($.06 < \omega^2 < .14$), explaining about 7% of the variability of *transform*. The implications of this interaction are not immediately clear and will be discussed at greater length later.

The analysis of variance technique also leaves another question unanswered in addition to the size of the effects of significant independent variables. This second question concerns whether significant differences in matching outcomes exist among the combinations of independent variable manipulations as represented by the four experimental conditions. After main or interaction effects have been demonstrated by ANOVA, experimental psychology and other laboratory experiment researchers commonly employ a technique known as post-hoc, or *a*

posteriori, comparison to test whether significant differences in results exist between each possible pair of experimental conditions. These post-hoc comparisons provide a finer detail to the statistical picture of the experiment than the very blunt and general results provided by the ANOVA F tests. Where the ANOVA technique again only demonstrates which variables are significant, the post-hoc comparisons show which combinations of the levels of the independent variables are significantly different.

For post-hoc comparisons, I utilize Tukey's Honestly Significantly Different (HSD) test to preserve the experiment-wise Type I error rate. While there is a fairly wide menu of post-hoc tests to deploy for comparisons between individual experimental groups after the use of ANOVA³, Sheshkin (1997) recommends the Tukey HSD test as an acceptable compromise between overly conservative post-hoc comparisons that preserve experiment-wise Type I error rate and perhaps overly generous post-hoc tests that preserve the comparison-wise Type I error rate but not that of the whole experiment.

The Tukey HSD test may be contrasted to the more widely known *a priori* comparisons used in hypothesis testing (e.g. a two-sample Student's *t* test). Conducting a Student *t*-test to compare each possible pair of conditions would increase the chances of rejecting the null and therefore increase the Type I error rate because providing more observed *t* statistics simply gives the experimenter more chances to observe a value beyond the critical value. The Tukey HSD test employs a Studentized range statistic *Q* similar to a two-sample Student's *t* statistics but using a "generalized standard error" for all pairwise comparisons rather than the a standard error derived from the pooled variance of the two samples in the pair (Hinton 1995, p.132).

³ Examples of other post-hoc tests include Newman-Kuels, Scheffe, Fisher's LSD, Bonferroni-Dunn, and Dunnett tests. Cf. Sheshkin (1997) or the website http://web.uccs.edu/lbecker/SPSS/glm_1way.htm#6.%20Post%20Hoc%20Tests for more information on other post-hoc comparisons.

Stata does not come immediately equipped to handle post-hoc tests for two-way ANOVA, so to compensate I introduced a new labeling variable *condition*. The results-shown, fixed offer condition is group 1 in *condition*; the results-shown, variable offer condition is group 2 in *condition*. The results-hidden, fixed offer and results-hidden, variable offer conditions are labeled groups 3 and 4, respectively. These group numbers correspond to the group numbers in Table 4, where pairwise comparisons of conditions are displayed.

Tukey HSD *a posteriori* comparisons revealed significant differences between conditions 1 and 3 ($p < .05$) and conditions 2 and 3 ($p < .05$). In both pairs, condition 3 displayed a greater degree of positive assortative matching than the condition with results displayed. Differences between all other pairs of conditions were insignificant in post-hoc comparisons.

Table 4: Post-hoc Tukey HSD Results ($\alpha = .05$, $Q_{crit} = 3.927$)

<i>Group vs Group</i>	<i>Group Means</i>	<i>Mean Difference</i>	<i>Q</i>
1 vs 2	0.7403, 0.8220	0.0817	2.0447
1 vs 3	0.7403, 0.9840	0.2438	6.0993*
1 vs 4	0.7403, 0.8733	0.1330	3.3285
2 vs 3	0.8220, 0.9840	0.1620	4.0547*
2 vs 4	0.8220, 0.8733	0.0513	1.2839
3 vs 4	0.9840, 0.8733	0.1107	2.7708

Proposals – Hypothesis Testing

The previous analyses have examined the outcome of positive assortative matching from the top-down point of view of the experimental design. While the confidence intervals can tell us just how closely the experimental outcomes conform to full positive assortative matching, the ANOVA results and complementary tests identify differences between the experimental conditions. In principle, the preceding tests as a whole can help us see to what extent positive

assortative matching occurs as predicted by Ghatak and help isolate potentially confounding variables like learning effects.

However, there is another way to test the theory of the Ghatak paper based on the proof of the positive assortative matching result. The adaptation of Lemma #1 demonstrates that a safe proposer should be willing to bid more than a risky proposer for the opportunity to partner with a safe responder. Thus, the theory predicts that the proposals made by safe proposers should be higher than the proposal made by risky proposers, providing a hypothesis to test the validity of the model in addition to the final outcome of positive assortative matching. In fact, it is possible that an experimenter could observe full positive assortative matching caused by reasons other than those hypothesized by Ghatak. That is, heterogeneous matching would imply that the Ghatak model is flawed, but homogeneous matching would not necessarily imply that the model is correct in its explanation of what produces the result. So, the analysis of the experimental outcome needs to take the possibility of spurious results into account, and examining the proposals made by safe and risky proposers provides one way of doing so.

However, it must be noted that the responders evaluate proposals based not only on the amount of the group loan offered to them but on the type of the proposer as well. Consequently, I will here define the “implicit” proposal made by a safe proposer as the amount of the group loan explicitly offered to a responder plus the expected extra utility derived from having a safer partner and less chance of incurring the joint liability penalty. Since the expectation of this extra utility is equal to $cp(p_s - p_r)$ (recall Lemma #1) where p is the probability of success for the responder, the benefit of a safer partner differs between safe and risky responders. Based on the adaptation of the Ghatak model and strategic game developed earlier, I assume that safe proposers are effectively making proposals directed towards safe responders and hence use the

probability of success for a safe proposer, p_s , in calculating the “implicit” proposal of a safe proposer.

I therefore report for two experimental conditions the results of two-sample Student’s t tests of the null hypothesis $H_o: \mu_s = \mu_r$ against the alternative $H_1: \mu_s < \mu_r$, where μ_s represents the mean “implicit” proposal made by safe proposers in the condition and μ_r represents the mean explicit proposal made by risky proposers in the condition. Since the sample size is limited and safe proposers may be functionally out-bidding their risky competitors only by very small amounts, i.e. \$1, I choose this particular alternative hypothesis as an indictment of the theory. In other words, rejection of the null and acceptance of the alternative hypothesis would imply that the theory in Lemma #1 is incorrect. Failure to reject the null would leave the Ghatak theory intact though leave open the question of just how much safe proposers out-bid risky proposers, if they did so at all.

I conduct this hypothesis test only on the data from the second and fourth experimental conditions (where the participants are allowed to choose share offers) since by construction the proposals are identical in the first and third (fixed M) conditions. Table 5 displays the results of these hypothesis tests. Note that 90 observations for each sample are taken from a condition in Table 5. The term s_p^2 denotes the pooled variance of the two samples. The results show that we can reject the null for the Results Shown condition at a level of significance of less than 1% but cannot reject the null for the Results Hidden condition at conventional levels of significance. It seems that the risky proposers are out-bidding the safe proposers when results are shown but not when results are hidden.

Table 5: Two-sample t-test Results ($H_0: \mu_s = \mu_r$ vs $H_1: \mu_s < \mu_r$; $n = m = 90$)

Condition	μ_s	μ_r	s_p^2	t_{obs}
Results Shown	83.23	129.31	797.58	-10.9449***
Results Hidden	124.77	119.37	1127.25	1.0789

Proposals – Regression Analysis

The classical hypothesis tests involving mean implicit proposals just described examines the empirical validity of the positive assortative matching theory from the point of view of the proposers; by considering proposals in a different context, it is also possible to use the actions of the simulation’s responders to test Ghatak’s matching hypothesis. As already noted, one can, and indeed needs to, look deeper into the experimental data than the mere matching outcomes and see if the borrower pairings in the context of the proposals made are congruent with Ghatak’s theoretical derivation of positive assortative matching. The simulation’s safe responders effectively drive the matching results, whether homogenous or heterogeneous, both theoretically and operationally, so considering how proposals by the two different borrower types influence the decisions of safe responders should yield further insight into how one can explain the matching outcomes generated by the experiment.

I attempt to leverage this analysis of safe responder decisions using a linear regression framework. The regression model is specified as follows:

$$M_{ijt} = \beta_0 + \beta_1 B_{jt} T_{jt} + \beta_2 B_{jt} (1 - T_{jt}) + \zeta_t .$$

The dependent variable M_{ijt} represents the decision of a safe borrower i to match with a proposer j in period t of the simulation. This variable is encoded as a “1” if the accepted proposer is also a

safe borrower (homogenous matching) and “0” if the accepted proposer is of risky borrower type. The first explanatory variable $B_{jt}T_{jt}$ denotes the product of the monetary share offer proposal (B) and the borrower type (T , encoded “1” for safe type and “0” for risky type) of proposer j in period t . Note that the product term is zero if the proposer is a risky borrower. The coefficient of this variable (β_1), then, should estimate the ability of proposal by safe proposers to explain the variation in positive assortative matching by safe responders. In contrast, the coefficient β_2 should provide an estimate of how much of the matching result can be explained by the offers of risky proposers. The product term attached to this coefficient, $B_{jt}(1 - T_{jt})$, uses the information on the share offer proposal and borrower type of a particular proposer j in period t but now equals zero if the proposer is of safe borrower type and is strictly positive if the proposer is of risky borrower type. Thus, for any given proposer in a given period, only one of the product terms $B_{jt}T_{jt}$ and $B_{jt}(1 - T_{jt})$ will be greater than zero while the other will be exactly zero. The variables β_0 and ξ represent the constant term and shock term, respectively.

From a theoretical standpoint, one should be able to form several hypotheses about the coefficients in the model. First, the constant term β_0 effectively represents the probability of positive assortative matching if only proposer borrower type is considered and relative magnitudes of proposals are ignored. Since, *ceteris paribus*, all responders should strictly prefer a safe proposer as a partner, the estimate for β_0 should exceed one-half ($\beta_0 > 0.5$). Second, any proposer’s bid grows more attractive to a responder as the share of the group loan offered (B_{jt}) increases, so a larger offer by a safe proposer should increase the probability of homogenous matching while a larger offer by a risky proposer should have the opposite effect of decreasing the likelihood of homogenous matching. This intuition produces the twin hypotheses $\beta_1 > 0$ and $\beta_2 < 0$.

Table 6 presents the results of the regression analysis for three different sets of data. Column 1 displays the results for the regression run using all data from both experimental conditions (results-shown and results-hidden) in which participants selected the share offers made⁴. Column 2 displays the results from the variable data for the share offer, results-shown condition only. Column 3 presents the results of the regression using only data from the variable share offer, results-hidden condition. Each cell in the results table displays, from top to bottom, the correlation coefficient estimate (in bold), the robust standard error for that estimate, and the 95% confidence interval for the correlation coefficient. The reader should immediately note that all coefficient estimates are significant at the 1% level or better.

<i>Table 6: Regression Results</i>			
<i>Variable</i>	1	2	3
<i>B_{jt}T_{jt}</i>	.00213*** (.00036) [.00142, .00284]	.00387*** (.00086) [.00216, .00557]	.00089*** (.00034) [.00022, .00156]
<i>B_{jt}(1 - T_{jt})</i>	-.00610*** (.00038) [-.00685, -.00534]	-.00481*** (.00055) [-.00590, -.00372]	-.00790*** (.00065) [-.00919, -.00662]
<i>Constant</i>	.75074*** (.04308) [.66573, .83576]	.63340*** (.07590) [.482534, .78426]	.88404*** (.04523) [.79415, .97393]
<i>R²</i>	0.9102	0.8912	0.9610
<i>N</i>	180	90	90

The coefficient estimates are generally the same across all three regressions in terms of signs and significance. In each case, the signs for β_1 (positive) and β_2 (negative) turn out as expected. Thus, the results support the common-sense intuition that higher bids by safer

⁴ For the fixed share offer conditions, this regression would be uninteresting, since no information on proposal magnitude could be gleaned from the results.

borrowers increase the likelihood of positive assortative matching while higher bids by risky borrowers decrease the same phenomenon. Though the coefficient estimates are numerically low, one must remember that they represent the change in the probability of homogenous matching for each incremental increase in a bid and that the bids can range over numerically large intervals. Estimates for the constant term, representative of the probability of homogenous matching if bids are non-existent, are in all cases greater than 0.5 as expected. However, in the results-shown condition, the 95% confidence interval for β_0 is inclusive of 0.5, so a rejection of the null hypothesis $\beta_0 \leq .5$ cannot be ruled out for this condition. There is also a surprisingly large discrepancy between the β_0 confidence intervals for the results-shown and results-hidden conditions, so much so that the two do not overlap at all. Finally, in all three regression outputs, the absolute value of the β_2 estimate is greater than the absolute value of the β_1 estimate, though the confidence intervals are not distinctly different for the results-shown condition. One might infer from this difference that safe responders generally tended to be more sensitive to changes in bids by risky proposers than to changes in bids by safe proposers.

Discussion & Conclusion

This study has attempted to help explore the extent to which the positive assortative matching result derived by Ghatak can explain and predict human behavior in the context of microfinance group lending. In so doing, this study also contributes to the wider study of the microfinance by helping to support or weaken the case for “pure” incentive effects as an explanation of MFI success above or alongside Cassar et al.’s other “social” theoretical categories. Empirically testing the “pure” incentive theory as exemplified in Ghatak’s model can, if support for positive assortative matching is found, provide motivation for further theoretical development of “pure” incentives, while, if the Ghatak model is shown to be lacking, this study

can help build a case for more research into other, particularly social, considerations. The microfinance literature needs to generate more evidence based on the observed behavior of real, warm-blooded people if it is ever to determine which of these two theoretical routes it should primarily pursue, and the experiment explained here provides just such a means of leveraging human behavior to test theory. This section will conclude the paper by examining the results of the various statistical analyses as a whole and commenting on what lessons can be learned.

To What Extent Did Positive Assortative Matching Occur?

The extent to which the outcome of positive assortative matching occurred in the experiment provides the obvious starting point for an interpretation of the results. The Wilson confidence intervals come the closest to qualifying as true *prima facie* evidence with regards to the predictive power of the Ghatak theory, and here the support for the positive assortative matching result is, at best, rather weak. Only two of the four experimental conditions show a greater tendency towards positive assortative matching than would be expected by mere chance, but because the results are noticeably different across levels of the results display variable, there is a clear need to qualify the overall matching results with other explanatory variables, ostensibly learning effects. Thus, positive assortative matching seems more likely in some cases than others. However, the Wilson intervals do not go so far as to refute conclusively and outright the Ghatak hypothesis by showing that the opposite of positive assortative matching, namely, heterogeneous matching, prevails.

What Indirect Evidence on the Ghatak Theory is Available?

Though the direct evidence for the Ghatak models in terms of the Wilson intervals is not very strong, the experiment also provides the opportunity to consider indirect evidence testing intermediate results and key propositions of the model using the analysis of simulation proposals.

Again, as in the case with the Wilson intervals, the evidence is decidedly mixed, yet here the results seem to be mediated more by the simulation role of a participant than whether or not results are displayed. To indirectly support the Ghatak model, one can point to the regression analysis. Generally, the coefficient estimates imply that participants playing as safe responders reacted to variations in bids by different proposer types largely as the Ghatak model would predict. The constant term estimate also suggests that safe responders were, to a degree, taking into account proposer borrower type in line with the positive assortative matching theory, even if their considerations did not always show up prominently in the Wilson intervals. On the other hand, the implicit bid hypothesis testing shows proposer behavior running counter to the implications of positive assortative matching, at least when participants are aware of their project outcomes. Again, learning effects due to display of results seem to mediate the support of the Ghatak model.

In sum, the experiment produces support for the Ghatak model of positive assortative matching that is very limited and needs to be thoroughly qualified. It appears that the model only accurately predicts behavior under specific conditions, perhaps one-time rather than repeated loans as implied by the learning effects. However, there is nonetheless at least some support for the model, and the evidence against the positive assortative matching needs to be considered in light of potentially confounding variables and individual variation. If participants were clearly affected by considerations other than or in addition to the “pure” incentive effects of their joint liability loans, what were those external influences?

How Can the Limited Support for the Ghatak Model Be Explained?

In any experiment, the researcher must be cognizant of issues related to both internal and external validity. By “internal validity,” I mean the elimination of potentially confounding

variables other than those manipulated or measured by the experimenter. Upon reviewing the entirety of the evidence produced by the experiment and statistical analyses, I will argue here that certain psychological factors acted as key confounds to the matching outcomes. There are clear and conclusive indications of significantly different outcomes between the results-shown and results-hidden groups, most notably the Wilson intervals, the ANOVA *Showresults* variable, the bid hypothesis testing, and the differences between the second and third regression outputs. Moreover, the omega-square calculation suggests that the manipulation of the results display accounts for nearly a fifth of the variation in the matching outcome, certainly a sizable proportion. Though one cannot be entirely certain why such differences between the two experimental groups appeared, it is very tempting to attribute the disparity to learning effects. I included the display of results as an independent variable in the experimental design precisely to help identify participant learning across the simulations periods, and I am disposed to believe that participant learning does in fact explain the significance of the results manipulation. It is entirely plausible that the display of project results for one group influenced those participants to alter their strategies as the simulation progressed, in this case away from positive assortative matching. Perhaps the revealed ex post outcomes of the stochastic investment projects influenced the participants more than the ex ante probabilities one would expect to form the basis of a strategy. Though anecdotal evidence, the post-experiment questions of one participant provides some support for this possibility. The participant asked whether or not his strategy of preferring a single partner until that partner defaulted served as the hypothesis tested by the experiment.

One must also recognize that these learning effects, if they do exist, at least narrow when participants are allowed to select their own proposal amount, as implied by the results of the Tukey HSD comparisons. Visually, this narrowing appears in Figure 2. This possibility would

explain the significant interaction between *Chooseoffer* and *Showresults* found in the analysis of variance and estimated in the omega-square calculations. However, the learning effects do not entirely disappear due to the interaction since the regression analysis finds some substantial differences between results conditions.

The interaction between the offer and results variables does help resolve some of the social psychological considerations incorporated into the experimental design. The insignificance of the main effect of the *Chooseoffer* variable in the ANOVA results may at first appear to be evidence for a confounding influence of minimal group formation. That is, the fact that allowing participants to select the amount of their own proposals did not appear to influence the likelihood of matching might be taken as evidence that participants were psychologically tethered more to artificial “groups” of their own borrower types than to the incentive structure. However, the collective findings of the significant ANOVA interaction and the Tukey HSD test illustrate that in fact allowing participants to select their own offers did influence behavior, though only in the presence of the manipulation of the results variable and only to a limited extent as implied by the effect size estimate.

In addition to potential learning effects, one could argue that the particular simulation role (proposer or responder) also acted as a variable confounding the matching results. The regression analysis suggests that safe responders acted with some of the intuitions of the Ghatak model in mind, but the hypothesis testing of safe versus risky proposals finds that some proposers (those in the results-shown condition) acted incongruently to the Ghatak theory. Importantly, these roles were often played by the *same people* at different times. Each condition had two treatments where participants played as the opposite borrower type and opposite role, so a safe responder in one treatment would become a risky proposer in the other treatment and vice

versa. I can only speculate here that this within-individual variation in strategy across roles was a function of the posted offer system used in the experimental design. The technological necessity of the posted offer system for making and accepting proposals forced the proposers to make a single offer to responders generally without directly knowing the offers of other proposers.

Formulating an effective proposal in this environment may have been too cognitively burdensome for the participants; each proposer had to derive the strategic considerations of the simultaneous proposal while also presumable computing the amount of the proposal within the time constraint of a couple of minutes (albeit with the help of an onscreen calculator).

Participants may have simply relied on imperfect but quick heuristic reasoning or “gut” instinct rather than formal mathematical reasoning. A possible solution to this apparent issue would be to screen participants for mathematical or statistical knowledge during recruitment, but to do so for a microfinance study would raise even greater issues of external validity.

“External validity” refers to the generalizability of the results in the experimental laboratory to the larger “real world.” In the case of this particular experiment, that “real world” means the areas in which microfinance lenders actually operate. It must, of course, be admitted that an experiment testing positive assortative matching in the sort of field environments in which microfinance lenders operate would be preferable to testing undergraduates using computer software. However, economic experiments may avoid some of the generalizability concerns often directed towards psychology experiments due to the use of hypotheses based in rational choice theory. One should remember that rational choice theory predicts that the economic incentives in a person’s environment will dominate the influences of internal factors like personality or psychology. When testing subsets of rational choice theory, then, the questions of who is being tested and where should not be as important as the incentives with

which the participants are presented. In any case, until someone tests positive assortative matching in the field, laboratory experiments are the best available source of empirical data, and simply noting that the experiment recruited the famed “college sophomore” does not in and of itself refute the findings.

Directions for Future Research

Taking the empirical investigation to the field is an obvious potential direction for future research as a means of putting to rest concerns over the external validity of the laboratory experiment. Another means of boosting external validity may be to increase the size of the lending group beyond the two-person design used here. Ghatak generalizes his theoretical result of positive assortative matching beyond two borrowers in the appendix to his paper, but given the constrained computer resources and participant pool to which I had access, a two-person group simulation was by far the most feasible option logistically and technologically for this experiment. Likewise, a system or framework for participant matching other than the posted offer system used in Z-tree software here could help eliminate concerns over the strategic and computational complexity of the simulation. Ideally, such a system, on a computer or otherwise, would allow participants to make multiple proposals targeted to specific responders rather than a one-shot, made-to-everyone proposal. This system might be more intuitive to and easily understood by the participants. I do not know exactly how such a system could be constructed or programmed since the posted offer method is the only viable option I could personally configure in Z-tree, but I do not doubt that someone with more technological expertise could devise a better alternative. Until that happens, though, and issues of external validity are more fully resolved, the Ghatak model of positive assortative matching will need to wait for more substantial and less limited support for its ability to explain and predict human behavior.

References

Abbink, Klaus, Bernd Irlenbusch, and Elke Renner. "Group Size and Social Ties in Microfinance Institutions." *Economic Inquiry*. 44 4 (2006) 614-28.

Agresti, Alan, and Brent A. Coull. "Approximate is better than 'exact' for interval estimation of binomial proportions." *The American Statistician*. 52 2 (1998) 119-27.

- Brown, Lawrence D., T. Tony Cai, and Anirban DasGupta. "Interval Estimation for a Binomial Proportion." *Statistical Science*. 16 2 (2001) 101-33.
- Cardinal, Rudolf N., and Michael R. F. Aitken. *ANOVA for the Behavioural Sciences Researcher*. Routledge, 2006.
- Cassar, Alexandra, Luke Crowley, and Bruce Wydick. "The Effect of Social Capital on Group Loan Repayment: Evidence from Field Experiments." *The Economic Journal*, 117 (2007) F85-F106.
- Cohen, Jacob. *Statistical Power Analysis for the Behavioral Sciences*. 2nd Ed. Lawrence Erlbaum Associates, 1988.
- Diamond, Douglas W. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies*. 51 3 (1984) 393-414.
- Fischbacher, Urs. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*. 10 2 (2007) 171-78.
- Ghatak, Maitreesh. "Group lending, local information and peer selection." *Journal of Development Economics*. 60 (1999) 27-50.
- Guttman, Joel M. "Assortative Matching, Adverse Selection, and Group Lending." Networks Financial Institute Working Paper. September 2006.
- Hanley, J.A., and A. Lippman-Hand. "If nothing goes wrong, is everything alright?" *Journal of the American Statistical Association*. 249 (1983) 1743-45.
- Hinton, Perry Roy. *Statistics Explained: A Guide for Social Science Students*. Routledge, 1995.
- Hogg, M.A. and Dominic Abrams. *Intergroup Relations: Essential Readings*. Psychology Press, 2001.
- Irlenbusch, Bernd, Klaus Abbink, and Elke Renner. "Interest Rates in Group Lending: A Behavioural Investigation." *Pacific Economic Review*. 11 2 (2006) 185-99.
- Karlan, Dean S. "Using Experimental Economics to Measure Social Capital and Predict Financial Decisions." *American Economic Review*. 95 5 (2005) 1688-99.
- Mimmack, G. M., Gillian Mimmack, Gary Manas, Denny Meyer, D. H. Meyer, and G. J. Manas. *Introductory Statistics for Business: The Analysis of Business Data*. South Africa: Pearson, 2001.
- Morduch, Jonathan. "The Microfinance Promise." *Journal of Economic Literature*. 37 (1999) 1569-1614.

Olivier, Jake L. and Warren L. May. "A Discussion of Binomial Interval Estimators on the Boundary." *Journal of the Mississippi Academy of Sciences*. 52 2 (2007) 178-83.

Sheskin, David J. *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press, 1997.

Tajfel H. *Social Identity and intergroup relations*. Cambridge, England: Cambridge University Press, 1982.

Tajfel, H. and M. Billig. "Familiarity and categorization in intergroup behavior." *Journal of Experimental Social Psychology*. 10 159-170.

Townsend, Robert M. and Christian Ahlin. "Using Repayment Data to Test Across Models of Joint Liability Lending." *The Economic Journal*. 117 (2007) F11-F51.

Wilson, T.B. "Probable inference, the law of succession, and statistical inference." *Journal of the American Statistical Association*. 22 (1927) 209-212.

Filename: 2009 Senior Thesis-Shaner.doc
Directory: C:\Documents and Settings\zebroncm\Desktop\web-stuff\honors and
commencement\graduation 2009\honors-thesis
Template: C:\Documents and Settings\zebroncm\Application
Data\Microsoft\Templates\Normal.dot
Title: Houston Shaner
Subject:
Author: Houston Shaner
Keywords:
Comments:
Creation Date: 4/20/2009 5:34:00 PM
Change Number: 12
Last Saved On: 4/20/2009 5:39:00 PM
Last Saved By: Houston Shaner
Total Editing Time: 6 Minutes
Last Printed On: 7/10/2009 4:54:00 PM
As of Last Complete Printing
Number of Pages: 44
Number of Words: 12,310 (approx.)
Number of Characters: 70,173 (approx.)