Promoting Star Networks with Homogeneous Agents

Rong Rong and Daniel Houser

Abstract:

Information dispersion is key to many economic decisions, occurring through market systems as well as networks of agents (Jackson, 2009). Empirical and theoretical findings suggest that an efficient network for information dispersion takes the form of a star: small numbers of agents gather information and distribute it to a large group (Weimann, 1994; Rogers, 1995; Bala and Goyal, 2000; Galeotti and Goyal, 2010). However, testing of this theory in the lab has generally failed to find evidence that these networks emerge. An exception is Goeree et al (2008) that finds frequent star networks under the assumptions of ex ante agent heterogeneity and perfect information. Unfortunately, these conditions do not likely characterize natural environments (Feick and Price, 1987; Conley and Udry, 2010), and thus the conditions that promote star networks remain largely unknown. Drawing from institutions that exist in natural environments, we design experiments to shed light on this issue. We find that investment limits and the “right-of-first-refusal”, both of which are regular features of naturally occurring star networks, have a surprisingly strong ability to promote the formation of star networks even when agents are ex ante homogeneous. Using a cluster analysis, we are able to trace the large positive effects of these institutions to the impact they have on individual decision rules. In particular, we find that these institutions encourage individual rationality as well as positive habits that, it turns out, facilitate the formation and stability of star networks. Our results may have important implications for environments characterized by ex ante homogeneous agents, such as is found in the technology and agricultural sectors.

I. Introduction

Information dispersion is key to many economic decisions. In many relevant contexts, it occurs through networks of agents (Jackson, 2009). Empirical and theoretical findings suggest that efficient networks take the form of a star: small numbers of agents gather information and then distribute it to a larger group (Weimann, 1994; Rogers, 1995; Bala and Goyal, 2000; Galeotti and Goyal, 2010). Despite the theoretical advances, a persistent challenge has been to discover conditions under which star networks emerge within controlled laboratory environments. One study that successfully generated star networks imposed the requirements of ex ante agent heterogeneity and perfect
information (Goeree, et al, 2008). Unfortunately, these conditions rarely characterize natural environments (Feick and Price, 1987; Conley and Udry, 2010), leaving open the question of which naturally occurring conditions work to promote star network emergence.

The earliest works on star networks date to the 1950s. In their pioneering paper, Katz and Lazarsfeld (1955) coined the term “opinion leaders” to describe a small subset of highly connected people¹. Half a century later, studies continue to provide empirical support for the existence of opinion leaders (Weimann, 1994; Rogers, 1995; Katz and Lazarsfeld, 2006). Opinion leaders clearly make a difference. Empirical evidence has shown, for example, that words from opinion leaders boost sales of consumer products (Godes and Mayzlin, 2009), contribute to the prevention of AIDS (Kelly et al, 1992) and transmit political thought and ideas (Roch, 2007). Consequently, due to their importance in disseminating information, people in both the private and public sectors are eager to discover how to locate and influence the “opinion leaders” (Iyengar et al, 2008), an initiative that might be facilitated through a deeper understanding of the emergence and characteristics of star networks.

Theoretical studies of star networks have showed that under certain conditions star networks emerge endogenously and include efficient or stable equilibria (Bala and Goyal, 2000; Galeotti and Goyal, 2010²). Crucial conditions underlying the formation of equilibrium star networks are that (1) information can be shared (meaning it is non-rival) and (2) agents are able to form links unilaterally³. While strong, these conditions are attractive in that they are easily implementable in laboratory tests. The experiments we report in this study include both non-rival information and unilateral link formation protocol.

To investigate star-network formation, economists have begun to collect laboratory data from participants in various network environments (Callander and Plott, 2005; Falk and Kosfeld, 2003; Goeree et al, 2008). It turns out, however, that these studies have almost always failed to find the emergence of star networks⁴. The single exception that we are aware of is Goeree et al (2008), which found star networks to

¹ The concept of “influentials” can mostly be used interchangeably (Merton 1968)
² Some non-game-theoretical models of star network formation build upon preferential attachment and study the behavior of large networks (Barabasi and Albert, 1999; Jackson and Rogers, 2007). A direct test of those models needs validation of its behavioral assumptions, something hard to achieve with a lab experiment.
³ Jackson and Wolinsky (1996) discussed two cases in which those two conditions are lacking. They found that network efficiency and stability is hard to achieve under regular payoff functions.
⁴ Falk and Kosfeld (2010) found equilibrium “wheel” networks to emerge, but were not able to observe the formation of equilibrium star networks.
emerge only after imposing substantial agent heterogeneity among perfectly informed participants. The authors argue that this heterogeneity promotes star networks since it simplified the network coordination problem.

Despite its success, the conditions used by Goeree et al (2008) may not be satisfied outside the laboratory (Feick and Price, 1987; Conley and Udry, 2010). In this study, we investigate whether certain naturally occurring institutional features may promote star-network formation in the presence of ex ante homogeneous agents. In particular, we collect data from laboratory experiments to examine whether efficient network formation can be promoted by (1) limiting investment in acquiring information; (2) implementing sequential decisions in network investment; and (3) enforcing the “right of first refusal”, which ensures that an initial investor is able, should they desire to do so, to continue that investment.

We selected our treatment conditions because we believe each may bear relevance to the way networks emerge in natural settings. To begin, it is very common to create networks through a process of sequential decisions. For example, online networking sites, such as Facebook or Twitter, regularly ask current users to invite their friends to join the site, and then those friends ask their friends. As we detail below, it is straightforward to incorporate sequential decisions into our experiment design.

Investment limits are a regular feature of natural environments, as technologies are too expensive to allow repeated R&D investment (Dimasi et al, 2003; Dimasi and Grabowski, 2007). Further, when excessive investment causes large inefficiency, government policies may exist to monopolize “insider” information through procurement in order to effectively limiting aggregate investment (Tran, 2009). Moreover, when personal relationship is at stake, information investment may face natural binding constraint with regard to time or social distance (Marsden and Campbell, 1984). It is plausible that efficient star-networks might naturally emerge when resources are limited. We pursue this idea in the experiments we report below.

Beside common business contract practice, the “right of first refusal” (which we denote by RFR) emerges regularly whenever economic outcomes are in favor of persistent investments on one agent instead of spreading resources on multiple smaller investments. For example, low income families deciding whom to send to school, firms choosing which employees receive training, funding agencies selecting scholars to receive grants. In all these cases, efficient and stable star network configuration may benefit the group over all. Our study allows us to investigate the impact of the RFR on star network formation.
The key finding of our paper is that star networks can reliably emerge with ex ante homogeneous agents. Surprisingly, sequential decisions do not seem to promote equilibrium networks. Rather, we find that combining limits on over-investment with a right of first refusal generates robust star networks. The effect persists in either simultaneous or sequential decision environments. Moreover, investment limits alone generate only about half of the equilibrium outcomes compared to when they are combined with the RFR. We also examine individual decision rules using cluster analysis and find that the behavior of our lab subjects separates into clearly defined clusters. We investigate how different institutions may generate different behavior type. Our analysis demonstrates that investment limits in simultaneous decision environments promote positive habits to form which then translate into higher frequency of coordination on star networks. Investment limits in sequential environments generate frequent star networks, but the reason is that agents use more rationality to guide their choice.

This investigation makes both methodological and substantive contributions. Methodologically, we provide a network formation environment that allows participants to make simultaneous linking and information purchasing decisions. Perhaps surprisingly, and despite its prevalence in natural environments, to our knowledge we are the first to study this context experimentally. Substantively, the findings of our paper have direct implications for information dispersion, which are especially relevant to, for example, the agricultural sector. Our results suggest that (1) difficulties with coordination may lead to undesirable network outcome due to excessive purchases; and (2) rules that promote sustained investment by a single individual facilitates the formation of efficient networks.

The remainder of the paper is organized as follows: The next section briefly reviews the theoretical and experimental literature on network formation. Section 3 presents the experimental design and procedure, and sets up the hypothesis. Section 4 reports experimental results. Section 5 discusses cluster analysis. Section 6 concludes.

II. Literature Review
II.1 Theoretical works on star network formation

A great number of theoretical studies have been undertaken to shed light on the process of network formation (Jackson, 2003). The formation of star networks in particular has received substantial attention due to the asymmetric equilibrium prediction on investment and linking strategies (Bala and Goyal, 2000; Bramoulle et al, 2004; Galeotti and Goyal, 2010). Models to capture this asymmetry consider network
interactions as strategic substitutes\(^5\), where agents receive higher payoffs when some are investors and others link to those investors. This type of environment includes both anti-coordination games as well as games related to public goods provision\(^6\).

An early paper by Bala and Goyal (2000) studied an environment with non-rival network goods and the possibility of forming links unilaterally. They found that star networks emerge in equilibrium only if the benefit of information flows both directions on a link\(^7\). Their study was followed by Bramoulle et al (2004) who examined network formation in an anti-coordination game. They found that the shape of the equilibrium network need not be a star. The exact network shape depends on the cost of link formation. More recently, a study by Galeotti and Goyal (2010) extended the model of Bala and Goyal (2000) by endogenizing the choice to invest. Their study showed that star networks emerge in equilibrium as well\(^8\). Despite these theoretical successes, these results leave open the question of whether the conditions required by theory are sufficient to generate star networks empirically. Our paper studies this question based on the model of Galeotti and Goyal (2010). We discuss their paper in detail below.

The model of Galeotti and Goyal (2010) proceeds as follows: A group of identical rational agents face the choice of either purchasing information themselves or obtaining it less expensively by linking to one who has the information. The above two choices are made simultaneously and combined with other agents’ choices to determine the network structure. The key assumptions of the model are that information is non-rival and flows both ways on network links. These assumptions are important since they closely characterize certain situations of information dispersion in natural environments\(^9\).

Multiple equilibria exist in their environment. However, strict Nash equilibria are a much smaller set. In particular, when the cost of linking is sufficiently low, some agents will invest in becoming “opinion leaders” and the remaining agents will link to them\(^10\).

---

\(^5\) Theoretical works that studied network formation with strategic complement game include Jackson and Watts (2002) and Goyal and Vega-Redondo (2005). The environment of information dispersion we study here does not fit the nature of strategic complementarity.

\(^6\) Bramoulle and Kranton (2007) discussed public goods provision in network settings extensively. However, networks in their study are exogenous.

\(^7\) In their model, if benefit only flows along the direction of link sending, equilibrium network is a wheel.

\(^8\) Galeotti and Goyal (2010) predicts peripheral-sponsored stars, different from the center-sponsored stars as in Bala and Goyal (2000) and Bramoulle et al (2004).

\(^9\) For example, knowledge about agricultural technology is mostly non-rival and could be shared between personal connections of farmers regardless of the linking direction.

\(^10\) If opinion leaders deviate from investing, the group will end up with no provision of network goods, implying a lower payoff for the opinion leaders as well; if others deviate from linking to the opinion leaders, he/she will have to either pay a higher cost to self-invest or to end up with no information, implying a lower payoff for him/her. Therefore, no rational subject would want to deviate from the star network given others’ choices.
The equilibrium prediction, therefore, is a so-called core-peripheral network. A “star” network is the uniquely efficient core-peripheral network which includes a single agent in the core.

In light of their empirical relevance, several experimental studies have been conducted to shed light on the conditions under which star networks form in laboratory environments. Below we review in detail all of those studies of which we are aware.

II.2 Experiments on star network formation

Despite the abundance of empirical evidence related to star networks, we are aware of only three experimental studies on star network formation (Callander and Plott, 2005; Falk and Kosfeld, 2003; Goeree et al, 2008). Falk and Kosfeld (2003) tested the theory of Bala and Goyal (2000). In particular, they studied how equilibrium outcomes were impacted by one-way or two-way network information flow. They found, consistent with this theory, one-way flows lead to wheel networks. On the other hand, in contrast with the theoretical predictions, they found that under two-way information flows the network failed to converge to a star. They conclude that the need for asymmetric strategies combined with inequality aversion may contribute to the difficulty in realizing star networks.

Similar to Falk and Kosfeld (2003), Callander and Plott (2005) also tested Bala and Goyal (2000). They considered various conditions that differed in terms of the linking cost as well as the value of information. Then, going beyond the theory, they examine the impact of payoff heterogeneity among network agents. Their main finding was that star networks did not consistently emerge under theoretical conditions, and that even introducing payoff heterogeneity does not lead to systematic formation of star networks. Consequently, the report that “significant and persistent inefficiency” is a feature of all of their network environments.

In light of the complications with generating star networks, and following Callander and Plott (2005), Goeree et al (2008) explored whether common knowledge of agent heterogeneity combined with two-way information flows might promote star networks. They report that, in relation to conditions with homogeneous agents, significantly more stars are observed under heterogeneity, and that perfect information

---

12 Another study by Deck and Johnson (2004) tests the model of Droste et al (2000) where the predicted network is a lattice.
13 According to their experimental data, the heterogeneity on cost of linking doesn’t seem to promote star network formation significantly.
about the nature of heterogeneity plays an important role in facilitating the coordination required for network equilibrium.

Like the above studies, we explore conditions that may facilitate the emergence of star networks. But unlike these, our study emphasizes the importance of naturally occurring institutions by testing how those institutional characteristics may impact network formation.

While Callander and Plott (2005) and Goeree et al (2008) explore the impact of individual differences, there are not only theoretical but also strong empirical reasons to explore environments with ex ante homogeneous agents. One is that information about heterogeneity may not be always easily available in natural environments because it is unobserved. Indeed, substantial empirical research has found no differences between the observable characteristics of agents who play different roles in the network (Feick and Price, 1987; Geisser and Edison, 2005; Wiedman et al, 2001; Williams and Slama, 1995). Others have pointed out that obtaining information regarding the costs and benefits of other network agents may be difficult since people have an incentive to conceal their private information (Conley and Udry, 2010). Moreover, game theoretical models that predict star networks have always assumed ex ante agent homogeneity\textsuperscript{14}.

In view of the fact that individual-level information is costly and sometimes unfeasible to obtain, and to avoid introducing artificial focal points, we design our experiments to include ex ante homogeneity. We investigate conditions under which ex ante identical agents will make asymmetric equilibrium actions in order to establish efficient and stable star networks.

\textbf{III. Experiment design and hypothesis}

Recall that our goal with this experiment is to examine how naturally occurring institutions affect star network formation. Institutional characteristics such as sequential decisions, investment limits and the RFR often coexist with star networks. We conjecture that these institutional characteristics may be important conditions for star networks to form in undesigned environments. Our laboratory study brings those institutional features into a controlled laboratory setting and estimates the effect of each on star network formation.

\textsuperscript{14} Jackson and Lopez-Pintado (2011), Larrosa and Tahme (2011), Vandenbossche and Demuynck(2010) developed models with heterogeneous agents. However, none of these relate to incentives associated with information acquisition or diffusion, thus the predictions generated from those models are not star-shaped networks.
III.1 Experiment design

III.1.1. General environment

Our experiment design is based on Galeotti and Goyal (2010), but modified as described in Galeotti and Goyal (2007). This modification leads to the sharp prediction that star networks are the unique Nash equilibrium, and are also efficient. To our best knowledge, our study is the first to examine the network formation process where agents make simultaneous linking and information purchasing decisions.

Each experimental session included 16 subjects randomly divided into four groups. All subjects participated in three stage games. Each stage game consisted of a random number of rounds. Groups were fixed during each stage game, and each group member held a unique ID: J,K,L or M.

In each round, decision makers decide whom to link to among the other three group members and whether to purchase information. Table 1 details the costs and benefits associated with each action a player can take. If a participant purchases information, she pays a cost of E$0.9 and earns the value of information, E$3, with certainty. On the other hand, if a player decides to send a link to another player, she pays a cost of E$0.5 per link. When a link is sent to a person who has purchased information, the link sender also earns the value E$3 for the information; otherwise, the cost of linking has no return. Note that costs and values remain fixed throughout all three stage games and all treatments.

<table>
<thead>
<tr>
<th>Parameter value</th>
<th>Parameter value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of sending link</td>
<td>E$0.5</td>
</tr>
<tr>
<td>Cost of purchasing information</td>
<td>E$0.9</td>
</tr>
<tr>
<td>Value of information</td>
<td>E$3</td>
</tr>
<tr>
<td>Number of player in a group</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1. Costs and benefits associated with player’s actions

Subjects submitted their decisions using the decision screen (see Appendix A, Fig. 1). Then, a display screen informed all players of the current network outcome and each group member’s payoff (see Appendix A, Fig.2).

---

15 In this earlier version of the paper, Galeotti and Goyal (2007) described a best-shot game where investment decision is binary. Agent could choose either to invest on one unit of information or not to do so. The optimal level of investment for a group is also set to be one, so that any repeated investment is inefficient.

16 There are always at least 4 rounds in a stage. After round 4, the game has a random stopping probability of 0.04 at any given round. To keep control over the length of the real experiment, we use the predetermined length 16, 44 and 24 for experimental stage I, II and III respectively. Those numbers are generated using a random number device. The practice stage always lasts for 8 rounds.

17 We choose those parameters to ensure that the predicted equilibrium and efficient networks are star-shaped.
Within each of the three stage games, the payoff is determined by the accumulated earnings over all rounds. Players are informed about their own stage payoff at the end of each stage. They are also reminded that they will be re-matched with players with whom they had not played previously, and that their stage payoff would not be carried over to the new stage. Each subject’s earnings for the experiment are determined by one randomly determined stage game (each with equal probability).

III.1.2 Treatment Design

Within the general experimental environment described above, we study the effects of three institutional characteristics of network formation. We examine “sequential decisions” and “investment limits” on network formation both individually and jointly using a two-by-two treatment design. A fifth treatment then studies the effect of the “right of first refusal”.

In a two-by-two design, we vary the sequence of decisions in one dimension to be either sequential or simultaneous. In simultaneous treatments, subjects from the same group make their decisions at the same time not knowing what other subjects may choose. In sequential treatments, only one subject makes decision each round. Players make decisions according to the alphabetical order of their ID (First J, then K, L and finally M). Further, players earn money even on rounds when they do not make a decision, with their payoff determined by their most recent previous choice in combination with the choices of others.

The second dimension of our design varies whether there exists an investment limits. Absent an investment limits, players can purchase information and links at will, and independent of other players’ decisions. On the other hand, in treatments with investment limits the following three conditions hold: (i) in each round each player can either send a link or purchase information, but cannot do both; (ii) each player can send at most one link; and (iii) information can be held by at most one subject at any given time.

We refer to the treatment without limits as the “baseline” in both the sequential and simultaneous environments, and denote treatments with investment limits as “limits”.

Therefore, the four conditions we study are simultaneous-baseline (Sim_B), simultaneous-limits (Sim_L), sequential-baseline (Seq_B) and sequential-limits (Seq_L). Notice that Seq_L and Seq_B differ in two ways: while investment is limited in Seq_L, it also implies the RFR, by which we mean that a person who currently holds information has the right to continue to hold that information. The reason is that in Seq_L a subject who has purchased information will continue to hold it until their next decision, and nobody else will have been able to purchase additional information. Consequently,
the only way they can lose the information is if they give up the information. It follows that comparing Seq_L to Seq_B measures the total effect of the investment limits combined with the RFR.

While these two effects cannot be separated in our sequential environment, it is possible to achieve separation in a simultaneous setting. To do this, we construct a fifth treatment which builds on Sim_L but eliminates the RFR. At any given round, agents who choose to purchase information will have an equal chance to obtain the information regardless of whether she purchased the information in the previous round. This treatment is denoted as simultaneous-limits with no RFR (Sim_L_NoRFR).

In summary, we investigate network formation in five treatments that differ in terms of the sequence of moves, whether investment is limited and whether there exists a RFR. We list the properties of these five treatments in Table 2.

Table 2. Properties of Treatments

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Decision sequence</th>
<th>Investment limits?</th>
<th>RFR?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq_B</td>
<td>Sequential</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Seq_L</td>
<td>Sequential</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Sim_B</td>
<td>Simultaneous</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Sim_L</td>
<td>Simultaneous</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Sim_L_NoRFR</td>
<td>Simultaneous</td>
<td>Y</td>
<td>N</td>
</tr>
</tbody>
</table>

### III.2. Defense of Design

We selected our treatment conditions because we believe each may bear relevance to the way star networks emerge in natural settings.

- **Why sequential decisions?**

  It is very common to create networks through a process of sequential decisions. For example, online networking sites, such as Facebook or Twitter, regularly ask current users to invite their friends to join the site, and then those friends ask their friends. Similarly, since it is unfeasible to achieve face-to-face interaction with large numbers of people at the same time and location, social connections that exist “off-line” also mostly occur using a sequential process. For example, when farmers consult others before adopting certain new seeds, it is mostly likely that they observe what other farmers have done and turn to the ones that have experimented and have gathered enough experience with the seeds (Foster and Rozensweig, 1995; Conley and Udry, 2010). The emergence of “focal player”, in this case the center of the star network, is endogenously determined.

  We conjecture that there is a fundamental principle that underlies the casual observation about the connection between sequential decisions and star-shaped
information networks. We designed this sequential decision environment for the purpose of testing this specific connection in a more controlled manner.

- **Why investment limits?**
  Investment limits are a regular feature of natural environments. The limits can take two forms. First, repeated investment on information may not be feasible. One reason is that technologies are expensive to invent (Dimasi et al, 1991; Dimasi and Grabowski, 2007), which means it is impossible for multiple groups to obtain the capital needed to invest in developing the same idea. Further, policies may exist to prevent inefficient investment in information. For example, to reduce inefficiency caused by excessive rent seeking, many governments choose to monopolize “insider” information through procurement auctions, a method effective at limiting investment (Tran, 2009).
  Second, sending multiple links may not be feasible. Non-monetary constraints in terms of time or social distance become important in this regard. Their constraining effect can become especially significant in the context of developing personal relationships, where time spent together is an important factor and often a binding constraint (Marsden and Campbell, 1984).
  In natural environments, we often find that star networks emerge in the contexts where resources are limited for use of information purchasing or forming links. It is plausible that star networks might emerge naturally in the presence of investment limits to cope with the lack of resources. We pursue this possibility in our experiments by designing treatments with investment limits.
  The specific limits we consider in our experiment reduce the set of actions each player can choose, which consequently reduces the number of possible network outcomes in our experiment¹⁸. Note that these limits are a necessary but not sufficient for equilibrium star networks. Indeed, after imposing these restrictions there remains 521 unique network outcomes, of which only four represent equilibrium.

- **Why the “right of first refusal”?**
  While the RFR regularly appears in business contracts, it also shares a broader implication for any situation where long-term investment emerges naturally. An example is investment in personal education. Low income families in developing countries may only able to support formal education for one sibling (Song, Appleton and Knight, 2006). The RFR determines that the child who first begins education has the right to persist in

¹⁸ The conditions we impose reduce the number of players’ possible actions from 16 to 5, which then subsequently reduces the number of possible network outcomes from 65536 to 512.
his/her schooling. Other siblings would then expect to receive the return to education by connecting with their educated sibling, that is, the family forms a star network. Similar considerations may often be at play when deciding which person or group is to receive an advantage to access information and knowledge. For example, the choice of employees to receive special training or of scholars to receive research grants.

The above cases are examples of history-dependent investment decisions that arise naturally. And again, those naturally occurring RFR may be an important factor underlying the emergence of star-networks. Our experiment design allows us to investigate the impact of the RFR in a controlled environment using a design that has parallels to undesigned environments that exhibit star networks.

III.3 Hypothesis

Our above discussion suggests, in general, that network environments that include sequential decisions, investment limits and the “right of first refusal” are expected to facilitate star network formation. The effects could be demonstrated in multiple ways. We discuss three possible measures that capture the effects of these institutional characteristics: equilibrium frequency, network stability and individual rationality.

- **Frequency of Equilibrium**

  **Hypothesis 1.** Star networks emerge more frequently in environments characterized by sequential decisions, investment limits and the RFR. In light of the above discussion, we expect the frequency of realized star networks to follow the order below (where A>B denotes that the star networks are expected to occur more frequently in A than B):

  1) Effect of sequential decisions:
     Seq_B>Sim_B; Seq_L>Sim_L
  2) Combined effect of investment limits and the RFR:
     Sim_L>Sim_B; Seq_L>Seq_B
  3) Effect of RFR alone (only in simultaneous environments):
     Sim_L>Sim_L_NoRFR
  4) Effect of the investment limits alone (only in simultaneous environments):
     Sim_L_NoRFR>Sim_B

- **Network Stability**

  The network configuration changes over time, and network stability is key to many real world applications such as R&D in business firms (Dodd et al, 2003). We use two
measures to investigate the stability of networks among treatments: duration of equilibrium and duration of disequilibrium.

**Hypothesis 2.** This hypothesis again follows from our argument that star networks are promoted by sequential decisions, investment limits and the RFR. This implies that environments with these features should exhibit a longer mean duration of continuous star networks and, after falling out of equilibrium, a shorter mean duration of disequilibrium.

Specifically, we hypothesize the following regarding the mean duration of continuous star networks:

1) **Effect of sequential decisions:**  
   \[ \text{Seq}_B > \text{Sim}_B; \text{Seq}_L > \text{Sim}_L \]

2) **Combined effect of investment limits and the RFR:**  
   \[ \text{Sim}_L > \text{Sim}_B; \text{Seq}_L > \text{Seq}_B \]

3) **Effect of the RFR alone (only in simultaneous environments):**  
   \[ \text{Sim}_L > \text{Sim}_L_{\text{NoRFR}} \]

4) **Effect of the investment limits alone (only in simultaneous environments):**  
   \[ \text{Sim}_L_{\text{NoRFR}} > \text{Sim}_B \]

Similarly, we form hypotheses regarding the mean duration of disequilibrium after falling out of equilibrium. Note the effect is opposite of the first set of hypotheses, since the less time it takes to restore a star network the better an institution is at promoting stability.

5) **Effect of sequential decisions:**  
   \[ \text{Seq}_B < \text{Sim}_B; \text{Seq}_L < \text{Sim}_L \]

6) **Combined effect of investment limits and the RFR:**  
   \[ \text{Sim}_L < \text{Sim}_B; \text{Seq}_L < \text{Seq}_B \]

7) **Effect of the RFR alone (only in simultaneous environments):**  
   \[ \text{Sim}_L < \text{Sim}_L_{\text{NoRFR}} \]

8) **Effect of the investment limits alone (only in simultaneous environments):**  
   \[ \text{Sim}_L_{\text{NoRFR}} < \text{Sim}_B \]
**Individual Rationality**

Individual rationality measures whether each subject’s choice is rational. It is possible in principle to have high mean individual rationality and yet a low frequency of equilibrium (e.g., if three group members make systematically “good” decisions but one member makes systematically “bad” decisions”). In general, however, we would expect high individual rationality in network formation games to imply faster convergence to star networks as well as more stable network outcomes. Just as above, we expect sequential decisions making, investment limits and “right of first refusal” each to promote individual rationality.

**Hypothesis 3.** Individual rationality is greater in environments characterized by sequential decisions, investment limits and the RFR.

1) Effect of sequential decisions: $\text{Seq}_B > \text{Sim}_B$; $\text{Seq}_L > \text{Sim}_L$
2) Combined effect of investment limits and the RFR: $\text{Sim}_L > \text{Sim}_B$; $\text{Seq}_L > \text{Seq}_B$
3) Effect of the RFR alone (only in simultaneous environments): $\text{Sim}_L > \text{Sim}_L_{\text{NoRFR}}$
4) Effect of the investment limits alone (only in simultaneous environments): $\text{Sim}_L_{\text{NoRFR}} > \text{Sim}_B$

**III.4. Experimental Procedure**

The experiment sessions were conducted between Dec 2010 and March 2011 in the ICES laboratory at George Mason University. Subjects were recruited via email from registered students at George Mason University. Each subject participated in only one session and none had previously participated in a similar experiment.

In total 160 subjects participated in the computerized experiment programmed with z-Tree (Fischbacher 2007). Each experimental session lasted between 120 and 150 minutes. Subject’ total earnings were determined by the Experimental Dollars (E$) earned in the end of the experiment, which is then converted by a rate of E$3 per US dollar. The average earnings were $25.28, ranging from a maximum of $53 to a minimum of $8 across all session.

In all treatments, before a session starts, subjects were seated in separate cubicles to ensure anonymity. They were informed of the rules of conduct and provided with detailed instructions (see appendix B for the instruction of Sim_B as an example). The instructions were read aloud. In order to guard against confusion, after having finished the instructions subjects were asked to complete a quiz and an experimenter checked their answers. Then the experiment worked through the quiz questions on a white board in
front of the laboratory. The experiment began after all subjects confirmed they had no further questions.

We ran 2 sessions for each treatment condition. Thus, in the end, we obtained 672 network graphs for each treatment (excluding the practice stage). Most of our analysis assumes 24 observations (eight groups each of which plays three stage-games with perfect strangers) for each treatment.

IV. Results

Section IV.1 presents results from treatment comparisons at the aggregate level. Section IV.2 uses a cluster analysis approach to shed light on decisions at the individual level\(^{19}\).

IV.1 Aggregate Analysis

We present results in the order of hypotheses listed in section III.3. First we discuss results concerning the frequency of star networks. Then we investigate two network stability measures: the duration of equilibrium and disequilibrium. Finally, we discuss how the proportion of individually rational moves varies across treatments and over time.

IV.1.1 Frequency of Star Networks

We count a network graph as a star if and only if there is one member who chooses to invest in information and the other three agents send exactly one link to the sole investor. For each stage game, the frequency of equilibrium is found by dividing the total number of star networks by the total number of rounds in that stage game. Then, the mean frequency of star networks is the average over the 24 stage-game frequencies in each treatment. The mean frequency of star networks in each of our treatments is described by Figure 1. It is clear from this figure that star networks emerge at different rates, with baseline treatments displaying the lowest frequency of star networks. More formally, our findings are as follows:

\(^{19}\) All of the below excludes the non-incentivized practice stage. Our results remain qualitatively identical if practice stage data are included.
Figure 1. Mean frequency of star networks (in %) for all treatments

Note: standard deviation of each bar is shown with red mark

Result 1. (Test of hypothesis 1.1) Sequential decisions do not increase the frequency of star networks.

Support: We found star networks to emerge with frequency 12.6% and 12.7% in Sim_B and Seq_B, respectively (p=0.667\(^{20}\)). On the other hand, when investment limits and the RFR are both present, 53.5% of networks formed in Seq_L are star shaped in comparison to 65.3% in Sim_L, and this increase is also insignificant at standard levels (p=0.054).

Result 2. (Test of hypothesis 1.2) More star networks emerge when investment limits are combined with the RFR.

Support: Sim_L generates 65.3% of star networks, while only 12.6% of networks in Sim_B are star shaped. This difference is significant (p< 0.001). Agents in Seq_L formed star networks 53.5% of time. When compared the 12.6% in Seq_B, the difference is again significant (p< 0.001). Thus, our data provide clear evidence supporting the positive impact of investment limits and the RFR on star network formation.

Result 3. (Test of hypothesis 1.3) The RFR promotes star network formation in simultaneous decision environments.

Support: The final two bars in Figure 1 correspond to Sim_L and Sim_L_NoRFR. The only difference between these two treatments is that the RFR is present in the former but absent in the latter. In Sim_L star networks emerge at a rate of 65.3%\(^{21}\). The

\(^{20}\) Unless otherwise indicated, all p-value refers to two-tailed Mann-Whitney test

\(^{21}\) The frequency of star network we found here, 65.3%, is much higher than the best case in Goeree et al(2008) with agent heterogeneity condition.
frequency in Sim_L_NoRFR treatment (32.4%), is significantly lower than this (p =0.0016). This is evidence that the RFR promotes star networks.

**Result 4. (Test of hypothesis 1.4) In simultaneous environments star networks emerge more frequently with than without investment limits.**

Support: Sim_L_NoRFR generates star networks at a rate of 32.4%, while the frequency in Sim_B is 12.6% (p=0.0034). Thus, investment limits promote star networks under simultaneous decision making.

**Figure 2. Mean Frequency of Star Network (%) over Time**

**Result 5. (Time Trend) More star networks are formed in later rounds of the experiment in Sim_L, Seq_L and Sim_L_NoRFR.**

Support: Figure 2 depicts the mean frequency of star networks over time. In all treatments the frequency of equilibrium is low in early rounds and then increases. Despite this common increase, when considering only the final five rounds, we can reject the hypothesis that mean frequency of star network is the same in all treatments (Kruskal-Wallis test, p=0.0006).

To analyze differences in the rate at which equilibrium changes over time and between treatments, we use an OLS regression with fixed effects at the group level. The results are described by Table 3. The key finding is that Seq_L displays a statistically significantly higher rate of increase in star networks over time than any other treatment (see Table 4). In contrast, Sim_B and Seq_B have the lowest rate among all treatments (Chow test, maximum p-value less than 0.0028, See table 4).
Table 3. Linear Regression on the evolution of star network frequency

<table>
<thead>
<tr>
<th></th>
<th>Seq_B</th>
<th>Seq_L</th>
<th>Sim_B</th>
<th>Sim_L</th>
<th>Sim_L_NoRFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>period</td>
<td>0.0053 (0.000)</td>
<td>0.0053 (0.000)</td>
<td>0.0158 (0.000)</td>
<td>0.0158 (0.000)</td>
<td>0.0057 (0.000)</td>
</tr>
<tr>
<td>stage</td>
<td>-0.128 (0.005)</td>
<td>-0.128 (0.004)</td>
<td>-0.419 (0.000)</td>
<td>-0.419 (0.000)</td>
<td>-0.133 (0.006)</td>
</tr>
<tr>
<td>constant</td>
<td>0.1827 (0.001)</td>
<td>0.1826 (0.000)</td>
<td>0.7644 (0.000)</td>
<td>0.7644 (0.000)</td>
<td>0.1906 (0.001)</td>
</tr>
<tr>
<td>fe(group)</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>R²</td>
<td>0.0310</td>
<td>0.0981</td>
<td>0.1200</td>
<td>0.1528</td>
<td>0.0353</td>
</tr>
</tbody>
</table>

Table 4. Chow test for the same slope of regression in Table 4 (p-value)

<table>
<thead>
<tr>
<th></th>
<th>Seq_B</th>
<th>Seq_L</th>
<th>Sim_B</th>
<th>Sim_L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq_L</td>
<td>0.0007</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sim_B</td>
<td>0.6515</td>
<td>0.0017</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sim_L</td>
<td>0.0011</td>
<td>0.7743</td>
<td>0.0028</td>
<td>-</td>
</tr>
<tr>
<td>Sim_L_NoRFR</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0468</td>
</tr>
</tbody>
</table>

IV.1.2 How do treatments affect network centrality?

Networks are said to have high centrality when a small number of people have a large number of connections. So, for example, network centrality is highest in our environment when one person receives three links, and no others receive any links, as occurs in a star network. It is lowest, on the other hand, when everybody sends links to each other person. Here we take a similar approach to Georee et al (2008) and report the “degree centrality” per group. The following equation gives the formal definition of this measure:

\[
d - centrality(g) = \frac{1}{(n-1)(n-2)} \sum_{j \in N} \left[ \max_{j \in N} \text{deg}_j(g) - \text{deg}_i(g) \right]
\]  

(1)

We normalize the degree centrality of a star network to be one. Intuitively, centrality measures closer to one suggest the network is more similar to a star. Our goal is to determine whether agents form networks that are more star-like over time and whether this tendency differs between treatments.

Result 6. The degree centrality is higher in treatments with investment limits and the RFR.

Support: Figure 3 shows the mean degree centrality under each treatment conditions. After Bonferroni correction for multiple comparisons, the difference between Sim_B and
Sim_L, Seq_B and Seq_L remains statistically significant (p=0.000 and 0.078 respectively). Thus, under investment limits and the RFR, more star-like networks form in both sequential and simultaneous environments.

Figure 3. Mean degree centrality measure across treatments

IV.2 How stable are the star networks?

We consider two measures of stability: (1) once a star network forms, how many rounds does it last? (2) after falling out of equilibrium, how many round does it take to restore a star network? Figure 4 demonstrates the average length of continuous star networks across treatments.

Figure 4 The length of continuous equilibrium networks (%) across treatments
Result 7. (Test of hypothesis 2.1) Sequential decisions do not reduce the duration of star networks.

Support: In Sim_B, once a star network forms it will last on average for only 2.1 periods (0.075% of the total number of rounds in a stage game). While in Seq_B, star networks sustain for an average 2.24 rounds. The difference is not statistically significant (Mann_Whitney two tailed, p=0.6201). Similarly, when comparing results in Sim_L (0.298%) and Seq_L (0.445%), we cannot reject the null hypothesis that their equilibrium duration is identical (p=0.2742).

Result 8. (Tests of hypotheses 2.2, 2.3 and 2.4) Star networks last longer in treatments where there exist investment limits or the RFR.

Support: Observed duration for continuous star networks parallels the results found regarding the frequency of equilibrium.

In simultaneous environments, when investment limits are introduced (Sim_L_NoRFR), the average length of continuous star network increases to 5.8 rounds, a significant increase from Sim_B (p=0.0098). If we further add the RFR in Sim_L, the length of the equilibrium network increases to 12.46 rounds, a significant difference compared both to Sim_L_NoRFR and Sim_B (p=0.0032 and 0.0000 respectively).

In sequential environment, the same institutions of investment limits and the RFR are introduced. In Seq_L, star networks last for about 8.34 rounds in a 28-round stage game. Agents in Seq_L are not sustaining the same star as long as they did in Sim_L, but the difference is not significant (p=0.2742).

Result 9. (Test of hypothesis 2.5) Sequential decisions do not shorten the duration of disequilibrium in the Baseline condition, but do under investment limits.

Support: Figure 5 shows that the duration of disequilibrium before reestablishing a star network does not differ between Sim_B and Seq_B (p=0.7142), while it does between Sim_L and Seq_L (p=0.0371).

---

22 The longest lasting star network occurs in Seq_L treatment, a group of subjects kept on playing a single star network for 43 rounds in their 60-round stage game without disruption.
Figure 5. The length of Continuous Disequilibrium networks across treatments

![Graph showing the length of Continuous Disequilibrium networks across treatments.]

Result 10. (Test of hypothesis 2.6, 2.7, 2.8) Disequilibrium durations are shorter under investment limits or the RFR.

Support: Our data support hypothesis 2.6 and 2.7. Subjects in Seq_L take 7 rounds on average to return to equilibrium play, while in Sim_L, they take only 5.5 rounds. Comparing Sim_L_NoRFR to Sim_L, the disequilibrium length falls from 10.1 rounds to 5.5 rounds, a significant decrease (p=0.014). In Sim_B and Seq_B it takes on average 17.2 and 18.8 rounds to return to equilibrium after a disruption.

IV.3 Individual Rationality

We define a decision as rational if it satisfies both of the following conditions: (1) if any other agent has purchased information, one should link to her; (2) if no other agent has purchased information, one should purchase information. The percentage of decisions that satisfies these conditions are plotted in Figure 6 for each treatment condition.

Figure 6. Mean percentage of rational choices made under each treatment condition

![Graph showing the mean percentage of rational choices made under each treatment condition.]
Result 11. (Test of hypothesis 3.1) Sequential decisions do not improve individual rationality.

Support: We find no statistically significant differences in individual rationality between Sim_B and Seq_B (p=0.0512) or between Sim_L and Seq_L (p=0.2879).

Result 12. (Test of hypothesis 3.2) The combination of investment limits and the RFR promotes rational choice at individual level.

Support: In simultaneous environments, about 40.8% of choices made are rational in baseline condition, while in Sim_L, the number increases significantly to 79.3% (p<0.0000). In sequential environments, the effect is smaller (an increase from 52.2% to 85.5%) but remains highly significant (p<0.0000).

Result 13. (Test of hypothesis 3.3) The RFR alone increases individual rationality in the simultaneous environment.

Support: The last two bars in Figure 6 show that the existence of the RFR improves individual rationality. About 51.9% of choices made in Sim_L_NoRFR are individually rational, a significant decrease in comparison to Sim_L (p=0.0023).

Result 14. (Test of hypothesis 3.4) Investment limits promote rational choice at the individual level in simultaneous environments.

Support: The percentages of individual rational moves for Sim_L_NoRFR and Sim_B are 59.1% and 40.8% respectively (p=0.0010).

V. Cluster analysis

Cluster analysis, as a numerical method for classification, serves the function of organizing a large and complicated data set into a small number of groups of objects so that the original data set could be easily understood based on patterns of similarity. It is widely used in fields such as astronomy, biology and marketing, and increasingly in economics (Fisher, 1963; Hirschberg et al., 1991; Houser, et al., 2004).

Compared to regressions, cluster analysis can better explore possible patterns within a complex environment where the ideal classification may not be well defined. In our case, where investment decisions in networks are made, how to establish meaningful groups of decision rules may not be obvious. Cluster analysis allows us to explore the emergence of decision rules across individuals without the need to pre-define the nature or number of possible rules (see also Houser et al., 2004).
In this section, we first discuss methods for cluster analysis and then connect to the analysis of network data. Then, we describe how we implemented cluster analysis with data and report the results.

V.1. A cluster analysis method

With cluster analysis one develops indices and criteria to know in a mathematically precise way how “close” or “far apart” objects are to each other.

- Dissimilarity matrix and clustering criteria

The starting point of many clustering investigations is an \( n \times p \) multivariate matrix \( X \) with \( n \) observations each of which includes \( p \) distinct features. This matrix \( X \) could be used to generate an \( n \times n \) dissimilarity matrix that reflects a quantitative measure of closeness by assigning a distance between each pair of objects.

A variety of distance measures have been proposed for deriving a dissimilarity matrix from a set of continuous multivariate observations (see, e.g., Jajuga et al, 2003). Among them, the most commonly used is Euclidean distance:

\[
d_{ij} = \left[ \sum_{k=1}^{p} (x_{ik} - x_{jk})^2 \right]^{1/2}
\]

where \( x_{ik} \) and \( x_{jk} \) are, respectively, the \( k^{th} \) variable value of the \( p \)-dimensional observations for individual \( i \) and \( j \). This distance measure has the appealing property that the \( d_{ij} \) can be interpreted as physical distances between two \( p \)-dimensional points \( x_i = (x_{i1}, x_{i2}, \ldots, x_{ip}) \) and \( x_j = (x_{j1}, x_{j2}, \ldots, x_{jp}) \) in Euclidean space. Examples of other popular measures include city block distance

\[
d_{ij} = \sum_{k=1}^{p} |x_{ik} - x_{jk}|
\]

and Minkowski distance

\[
d_{ij} = \left( \sum_{k=1}^{p} |x_{ik} - x_{jk}|^r \right)^{1/r}
\]

An informative clustering includes groups such that the distance between objects in the same group is small, while the distance between groups is large. Based on this simple intuition, a variety of so-called “dissimilarity indices” (formed by taking
combinations of distance measures) have been suggested. The examples we provide below help to motivate the cluster analysis approach we pursue with our data.

With \( d_{ql,kv} \) defined as the dissimilarity between the \( l^{th} \) object in the \( q^{th} \) group and the \( v^{th} \) object in the \( k^{th} \) group, the following equations give one example of an index for heterogeneity within group \( m \):

\[
h_{1}(m) = \sum_{l=1}^{n_q} \sum_{v_{l}, v_{l} \neq l} d_{ml, mv}^2
\]

This index measures the sum of squared dissimilarities between two objects that belong to the same group \( m \).

Another commonly used similar index measures the sum of squared dissimilarities between an object in a cluster group \( m \) and the mean of objects in group \( m \). It is also known as the trace of within-group dispersion matrix\(^{23}\). The above index comprises the foundation for the \( k \)-means clustering algorithm, as discussed later.

\[
h_{2}(m) = \sum_{m=1}^{g} \frac{1}{2n_m} \sum_{l=1}^{n_q} \sum_{v=1}^{n_v} d_{ml, mv}^2
\]

The final index we note here is called the “star index”:

\[
h_{3}(m) = \min_{v_{l}, v_{l} \neq l} \left[ \sum_{l=1}^{n_q} d_{ml, mv}^2 \right]
\]

where a reference object \( v \) is connected with all other objects of the group to form a star, which then generates the sum of the distance measure for object \( v \). The smallest sum of the distance is achieved when the reference object \( v \) is at the center of the group. In contrast to the \( h_{2}(m) \) index used in \( k \)-means algorithm, the \( h_{3}(m) \) index corresponds to the equally widely used \( k \)-medians algorithm.

Ideally, one would be able to go through all combinations of objects and discover the right clustering by choosing the one that yields the lowest dissimilarity index within each group\(^{24}\). However, when the number of objects is not small, it becomes infeasible to do so. Indeed, Liu (1968) provides the exact number of possible partitions one must consider in order to cluster \( n \) objects into \( g \) groups:

---

\(^{23}\) The dispersion matrix is derived from multivariate matrix \( X \) directly without constructing the dissimilarity matrix \( D \). These two methods are mathematically equivalent, hence we omit the discussion of the other method.

\(^{24}\) Indices that measure the separation between groups are also used in many other methods. We refer interested readers to Everitt et al (2011).
\[ N(n,g) = \frac{1}{g!} \sum_{m=1}^{g} (-1)^{g-m} \binom{g}{m} m^n \]  

(8)

That is, in order to partition 100 network agents into five groups, the number of possible combinations to examine is about \(6.6 \times 10^6\). This excessive computational burden has let scholars to develop numerical search algorithms to approximate clustering solutions. Two numerical algorithms, k-means and k-median, are widely used.

- **K-means algorithm:**
  
The k-means algorithms (Ball and Hall, 1967; MacQueen, 1967; Steinley, 2006b) involve iteratively updating partitions by simultaneously relocating objects into the group whose mean is closest and then recalculating group means. All k-means algorithms attempt to minimize within-group sum of squared deviations from the (group) mean (see equation 6). However, numerical approximations can differ in many ways, and these differences can have substantial impacts on the results\(^{25}\).

- **K-median algorithm:**
  
In more recent years, the k-median algorithm has received increasing attention (Kaufman and Rousseeuw, 1990; Hansen and Jaumard, 1997; Kohn et al, 2010). This algorithm relocates an object to a group whose median is the closest to it according to a certain distance measure. The specific clustering criteria corresponds to equation 7. In our results section below we report the outcome of both k-means and k-median algorithms.

- **Determining the number of clusters**
  
A cluster analysis requires one to determine the number of clusters a dataset contains (see Milligan and Cooper (1985) for a comparison of 30 methods). In the Monte Carlo analysis they report, the top performing method was the one suggested by Calinski and Harabasz (1974) (which we denote by C-H)\(^{26}\). Because the measure we use is derived from C-H, it is worthwhile to describe it in some detail.

  C-H suggested that the optimal number of clusters, \(g^*\), should maximize the following value \(C(g)\), where \(C(g)\) is given by:

---

\(^{25}\) We have found substantial differences in K-means clustering results produced by the standard packages in Stata, R and Matlab. We traced it to differences in the specific numerical algorithms used by each package.

\(^{26}\) Another successful technique developed by Duda and Hart (1973) works with hierarchical cluster methods. The network data do not fit these types of cluster analysis.
\[ C(g) = \frac{\text{trace}(B)}{g-1} \left/ \frac{\text{trace}(W)}{n-g} \right. \]  \hspace{1cm} (9)

where

\[ B = \sum_{m=1}^{g} n_m (\bar{x}_m - \bar{x})(\bar{x}_m - \bar{x})' \]  \hspace{1cm} (10)

representing the between-group dispersion matrix, and

\[ W = \sum_{m=1}^{g} \sum_{l=1}^{n_m} (x_{ml} - \bar{x}_m)(x_{ml} - \bar{x}_m)' \]  \hspace{1cm} (11)

representing the within-group dispersion matrix, both of which derive from the original multivariate matrix \( X \).

We are interested in tuning the C-H index to ensure it performs well with our particular type of data. To do this, we introduce a parameter \( \alpha \) that influences the penalty for including larger numbers of clusters:

\[ C'(g) = \frac{\text{trace}(B)}{g^\alpha-1} \left/ \frac{\text{trace}(W)}{n-g^\alpha} \right. \]  \hspace{1cm} (12)

We conduct a Monte Carlo analysis to discover which \( \alpha \) level performs best within the context of our data.

V.2. Implementing cluster analysis on network behavior data

In this section, we discuss the implementation of cluster analysis on data generated in our experiments. The key steps are as follows:

a) Estimate decision rules using choice data from the experiment

We assume subjects make decisions based on information regarding their rational move, their own past choice or other group member’s past choice. In particular, we look at how the information purchase decision is made depending on whether it is rational for the subject to purchase, the subject’s immediately past purchase decision and the past purchase decision of other group members. We estimate the impact of each information variable using a linear regression (see also Kurzban and Houser, 2005):

\[ purchase_{i,t} = \beta_0 + \beta_1 \text{rationality}_{i,t} + \beta_2 \sum_{j\neq i} \sum_{s=1}^{8} purchase_{j,t-s} + \beta_3 \sum_{s=1}^{8} purchase_{i,t-s} + \varepsilon_{i,t} \]  \hspace{1cm} (13)

where we define the variables in the regression as follows:

\[ purchase_{i,t} = \begin{cases} 1, & \text{if subject i purchased information at round t} \\ 0, & \text{otherwise} \end{cases} \]
\[ r_{\text{rationality}}_{i,t} = \begin{cases} 1, & \text{if subject } i \text{ should have purchased information at round } t \\ 0, & \text{otherwise} \end{cases} \]

\[ p_{\text{purchase}}_{i,t-s} = \begin{cases} 1, & \text{if subject } i \text{ purchased information in round } t-s \\ 0, & \text{otherwise} \end{cases} \]

We implement the above regression for each individual separately. The vector of estimates for each individual can then be represented by a single point in 3-space. Thus, we obtain 142 points in 3-space (See Figure 7, where each point represents the estimates \{\beta_1, \beta_2, \beta_3\} for one subject).

**Figure 7. Plot of individuals’ estimates, color coded for treatment**

![3D plot of beta](image)

**b) Cluster individual estimates using a k-means algorithm with an arbitrarily determined number of clusters g*:**

We cluster our estimates using a fixed \(g^* = 2, 3\) and 4. Greater numbers of clusters are not feasible to consider in light of the number of data points in our sample. We use the `kmeans` package in Matlab to implement clustering. Figure 8 plots Figure 7 but with objects color coded according to their cluster assignment. The three subfigures correspond to \(g^* = 2, 3\) and 4.
Figure 8. Plot of estimates, color coded by cluster

$g^* = 4$

3D plot of beta (dependent: purchasing decision, color coded for cluster)

$g^* = 3$

3D plot of beta (dependent: purchasing decision, color coded for cluster)
Our Monte Carlo analysis requires us to calculate the center of each cluster as well as the dispersion of its data. The cluster center is calculated by taking the mean over each dimension for all objects that belong to the same cluster. We also calculate the standard deviation and size (number of members) for each cluster.

c) **Simulate new data sets using the estimated decision rule coefficients.**

   For each level of $g^*$, we simulate new data sets based on the mean and standard deviations of clusters calculated in step (b). We simulate the $j^{th}$ dimension of the $i^{th}$ cluster by repeatedly drawing from a normal distribution with mean and standard deviation corresponding to step (b) calculations. We repeat the process for each cluster and each dimension. The simulated data therefore contains exactly the same number of objects as our experimental data, which are generated following a process of random perturbation around its mean. We simulated 100 datasets to be used in step (d).

d) **Calculate the optimal number of clusters $g$ for different level of $\alpha$ and compare it with the true number of cluster $g^*$. Find the optimal level of $\alpha^*$ that would perform best at recovering the true number of clusters from the simulated dataset.**

   We run k-means cluster analysis on each of the 100 simulated datasets allowing from two to four clusters. We use the modified C-H index discussed above to determine the optimal number of clusters for the simulated dataset. For 100 datasets, therefore, we...
obtain a distribution of the optimal number of clusters under a fixed level of the penalty parameter $\alpha$. We compared this distribution with the “true” number of simulated clusters.

To determine the “best” penalty parameter $\alpha$ we proceed as follows. First, we examine whether the modal result coincides with the true number of simulated clusters. Second, when the algorithm uncovers the correct number of clusters, we proceed to determine how well the specific location (means and standard deviations) of the cluster matches the true values used in the data generating process. Third, we determine whether pairs of objects that truly belong to the same cluster were also clustered into the same group by the algorithm under penalty parameter $\alpha$. To do this, we use the index proposed by Rand (1971). This index is constructed as follows.

Given a set of $n$ elements $A = \{a_1, \ldots, a_n\}$, its true partition $X = \{x_1, \ldots, x_s\}$ and another partition $y = \{y_1, \ldots, y_r\}$, the Rand index is calculated as follows:

$$
\text{Rand} = \frac{a + b}{a + b + c + d}
$$

(14)

where $a,b,c$ and $d$ is defined as in the following table:

<table>
<thead>
<tr>
<th>The count of pairs of objects that...</th>
<th>belong to the same set in $X$</th>
<th>belong to different sets in $X$</th>
</tr>
</thead>
<tbody>
<tr>
<td>belong to the same set in $Y$</td>
<td>$a$</td>
<td>$d$</td>
</tr>
<tr>
<td>belong to different sets in $Y$</td>
<td>$c$</td>
<td>$b$</td>
</tr>
</tbody>
</table>

A higher Rand index implies a more “accurate” clustering, in the sense that a higher percentage of pairs of objects were either clustered together when they originated in the same group, or were separated when they originated in different groups.

We list the results of all three methods mentioned above in Appendix C. Our general finding is that the penalty parameter $\alpha$ substantially impacts results. For the cases we considered, we found alpha =1.1 to be optimal. In particular, when $\alpha = 1.1$, (1) it is most likely for k-means algorithm to recover the correct number of clusters, (2) the distance

---

27 To implement this idea, we calculated the Euclidean distance of the cluster center to the center of the clusters used to generate the simulated data. Since cluster ID may not be kept the same with each cluster analysis, we use the minimum distance across all possible combinations between newly generated cluster centers and original cluster centers. The distance measure was generated for each simulated data. In the end, we obtain the mean and standard deviation of this distance measure averaged over 100 simulations. The lower the distance measure is the better the matches between new clustering results and the true clusters are.
between new cluster center and the true cluster center is closest and (3) the Rand index is largest.

e) Adopt $\alpha^*$ as the optimal level of penalty. Determine the number of cluster $\tilde{g}$ in our network data. Cluster the data into $\tilde{g}$ clusters using k-means algorithm.

Using $\alpha^* = 1.1$ and the modified C-H index, we discover three clusters in our data. The three panels of Figure 9 plot the projection of each 3-vector estimate into corresponding 2-space. Points of the same color belong to the same cluster.

Figure 9. Projections of estimates onto planes

![Figure 9. Projections of estimates onto planes](image)
The estimates display clear separation of two of the information variables we investigate in this study: individual rationality to purchase and one’s own past purchasing choice. We summarize the characteristics of the three groups as follows:

1) “Rational” type: people who are highly rational but not persistent at investing (or not investing)
   \[ \beta_1 = 0.6663, \beta_2 = -0.0092, \beta_3 = -0.0868, \text{size of group=58} \]

2) “Habit” type: people who are less rational but highly persistent at investing (or not investing)
   \[ \beta_1 = 0.3163, \beta_2 = -0.0437, \beta_3 = 0.5540, \text{size of group=42} \]

3) “Overinvestor”: people who tend to over-invest
   \[ \beta_1 = -0.1097, \beta_2 = -0.0015, \beta_3 = 0.1164, \text{size of group=42} \]

Table 5 reports types by treatments. Recall that the frequency of stars in Seq_B was 12.7%. This treatment included 3.45% of Habit types, 44.83% Overinvestors and 51.72% Rational types. The reason that stars tended not to form in this treatment is that participants tended to overinvest in information, and were unable to coordinate on a single investor.

The frequency of star in Sim_B was 12.6%. This treatment includes 10% of Habit type, 86.67% of Overinvestors and 3.33% of Rational type. The reason for low frequency
of star in Sim_B is that majority of participants overinvest. They also frequently change their linking and investing decisions, which leaves coordination very difficult to achieve.

The frequency of star in Seq_L was 53.5%. The treatment includes 92% of Rational type, 4% Habit type and 4% Overinvestor. The reason for high frequency of star in Seq_L is that large majority of subjects are making choices rationally. However, they switch their decisions frequently so that we do not observe positive habit formation.

The highest frequency of stars across all treatments occurs in Sim_L (65.3%). This effective institution generates 89% of Habit type, 3.7% of Overinvestor and 7.41% of Rational type. Different from Seq_L, the high frequency of star networks in Sim_L is due to the fact that majority of subjects belongs to Habit type. Their strategy relies less on individual rationality, but rather they form positive habits in the sense that once a star emerges, agents tend to sustain the same choice for an extensive period of time. Positive habit is the key to the success of this institutional condition.

Sim_L_NoRFR generates 32.4% of star networks. It includes 41.94% of Habit type, 3.23% of Overinvestors and 54.84% of Rational type. In this treatment, Investment limits promoted individual rationality, but were unable to facilitate the formation of positive habits due to the absence of the RFR. The consequence is an improved frequency of star networks in relation to the baseline, yet a frequency that falls far short of that achieved in environments that include RFR.

| Table 5. Number of individuals that belong to certain type and treatment |
|---------------|---------------|---------------|---------------|---------------|---------------|
|               | Seq_B 12.6% of star | Seq_L 53.5% of star | Sim_B 12.7% of star | Sim_L 65.3% of star | Sim_L_NoRFR 32.4% of star |
| Habit         | 1 (3.45)         | 1 (4.00)       | 3 (10.00)      | 24 (88.89)      | 13 (41.94)     |
| Over-investor | 13 (44.83)       | 1 (4.00)       | 26 (86.67)     | 1 (3.7)         | 1 (3.23)       |
| Rational      | 15 (51.72)       | 23 (92.00)     | 1 (3.33)       | 2 (7.41)        | 17 (54.84)     |
| Total         | 29 (100.00)      | 25 (100.00)    | 30 (100.00)    | 27 (100.00)     | 31 (100.00)    |

Note: percentage in parenthesis

VI. Conclusion

Star networks emerge naturally in many social environments, and theory indicates star network equilibria are efficient. Based on a theory suggested by Galeotti and Goyal (2007, 2010), we study star network formation using laboratory experiments. Previous experimental studies of network formation suggest that persistent star networks could
emerge in the lab, but only under ex ante agent heterogeneity (Goeree et al, 2008). This contrasts with natural environments where star networks frequently emerge even when agents are ex ante homogeneous (Feick and Price, 1987; Conley and Udry, 2010). We conjectured that certain naturally occurring institutional characteristics, such as sequential decisions, investment limits and the “right of first refusal” may help to explain why star networks form in these natural (undesigned) environments. Our main finding is that investment limits and the “right of first refusal” promote star network formation. In comparison to baseline treatments, we find that environments with those features realize increased star-network frequency, improved network stability and higher levels of individually rational decision outcomes. Using a cluster analysis to recover similarities in individual decision rules, we further find that institutions matter for decisions, in the sense that some institutions promote rationality and positive habits better than others. In particular, environments with investment limits tended to promote rationality, and environments with the right of first refusal tended to promote habits that were conducive to the formation of stable star networks.

To our knowledge, our paper reports the first experimental investigation of the impact of institutions on star network formation. Results from this study suggest how one might design institutions to promote the flow of information through social networks relevant to political campaigns, franchise training or agricultural innovation.28

Finally, this paper offers a direct test of network formation theory (Galeotti and Goyal, 2010). Our findings suggest that the institutional environment can substantially promote the formation of star networks. It would be profitable for future theory research to analyze the specific role of investment limits and the right of first refusal within an environment characterized by multiple agent behavioral types.

28 For example, our study suggests that when an important agricultural innovation occurs, designing a network formation process that incorporates investment limits or the RFR may lead to more knowledge transmission, therefore higher adoption rate, with lower cost.
References


Jackson, M. an D. Lopez-Pintado (2011): Diffusion in Networks with Heterogeneous Agents and Homophily', working paper


Song, L., S. Appleton and J. Knight (2006): Why do girls in rural China have lower school enrollment? World Development. 34(9), 1639-1653


Appendix A. Z-tree experimental interface

Fig. A An example of the decision screen
Fig B. An example of the display screen

<table>
<thead>
<tr>
<th></th>
<th>Player J</th>
<th>Player K</th>
<th>Player M</th>
<th>Yourself</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Links sent to...</strong></td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>-</td>
</tr>
<tr>
<td><strong>connections formed with...</strong></td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>-</td>
</tr>
<tr>
<td><strong>Number of items obtained</strong></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Earnings (E$) for this round</strong></td>
<td>3.1</td>
<td>3.5</td>
<td>1.0</td>
<td>3.5</td>
</tr>
<tr>
<td><strong>Accumulated earnings (E$) for this stage</strong></td>
<td>9.3</td>
<td>8.0</td>
<td>3.0</td>
<td>5.5</td>
</tr>
</tbody>
</table>
Appendix B. Multiple Comparison Correction:
We find most of the results in section IV.1 hold when we correct the p-value for multiple comparisons. Table A-D report pair-wise comparisons between treatments with a Bonferroni correction for equilibrium frequency, equilibrium duration, disequilibrium duration and individual rationality. Table entries show row mean-column mean (p-value in parenthesis).

### Table A. Differences in the mean frequency of star networks by treatments

<table>
<thead>
<tr>
<th></th>
<th>Seq_B</th>
<th>Seq_L</th>
<th>Sim_B</th>
<th>Sim_L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq_L</td>
<td>0.4078 (0.000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim_B</td>
<td>-0.0011 (1.000)</td>
<td>-0.4089 (0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim_L</td>
<td>0.5260 (0.000)</td>
<td>0.1182 (1.000)</td>
<td>0.5271 (0.000)</td>
<td></td>
</tr>
<tr>
<td>Sim_L_noRFR</td>
<td>0.1963 (0.080)</td>
<td>-0.2115 (0.044)</td>
<td>0.1974 (0.076)</td>
<td>-0.3297 (0.000)</td>
</tr>
</tbody>
</table>

### Table B. Difference of the mean equilibrium duration by treatments

<table>
<thead>
<tr>
<th></th>
<th>Seq_B</th>
<th>Seq_L</th>
<th>Sim_B</th>
<th>Sim_L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq_L</td>
<td>0.223 (0.022)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim_B</td>
<td>0.052 (1.000)</td>
<td>-0.217 (0.028)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim_L</td>
<td>0.369 (0.000)</td>
<td>0.147 (0.409)</td>
<td>0.364 (0.000)</td>
<td></td>
</tr>
<tr>
<td>Sim_L_NoRFR</td>
<td>0.132 (0.664)</td>
<td>-0.091 (1.000)</td>
<td>0.126 (0.778)</td>
<td>-0.238 (0.011)</td>
</tr>
</tbody>
</table>

### Table C. Difference of the mean disequilibrium duration by treatments

<table>
<thead>
<tr>
<th></th>
<th>Seq_B</th>
<th>Seq_L</th>
<th>Sim_B</th>
<th>Sim_L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq_L</td>
<td>-0.361 (0.001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim_B</td>
<td>0.058 (1.000)</td>
<td>0.419 (0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim_L</td>
<td>-0.416 (0.000)</td>
<td>-0.055 (1.000)</td>
<td>-0.474 (0.000)</td>
<td></td>
</tr>
<tr>
<td>Sim_L_NoRFR</td>
<td>-0.253 (0.048)</td>
<td>0.108 (1.000)</td>
<td>-0.310 (0.006)</td>
<td>0.164 (0.650)</td>
</tr>
</tbody>
</table>

### Table D. Difference of the mean disequilibrium duration by treatments

<table>
<thead>
<tr>
<th></th>
<th>Seq_B</th>
<th>Seq_L</th>
<th>Sim_B</th>
<th>Sim_L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq_L</td>
<td>0.333 (0.000)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim_B</td>
<td>-0.114 (0.314)</td>
<td>-0.447 (0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sim_L</td>
<td>0.270 (0.000)</td>
<td>-0.062 (1.000)</td>
<td>0.385 (0.000)</td>
<td></td>
</tr>
<tr>
<td>Sim_L_NoRFR</td>
<td>0.069 (1.000)</td>
<td>-0.264 (0.000)</td>
<td>0.183 (0.007)</td>
<td>-0.202 (0.002)</td>
</tr>
</tbody>
</table>
Appendix C. Simulated clustering result

Real Data: 3D Contribution regression data, g*=4, kmeans method
# of simulation:100 per parameter change
Searching range: g=2:4
Changing parameter: the alpha as the power over g in kmeans-Calinski-Harabasz method, Seed(4321)

<table>
<thead>
<tr>
<th>alpha</th>
<th>Distribution of g</th>
<th>Mean of g</th>
<th>Sd of g</th>
<th>Mean distance of cluster center if g=g*</th>
<th>Sd of distance of cluster center if g=g*</th>
<th>Mean Rand statistics</th>
<th>Sd of Rand statistics</th>
<th>Mean Rand statistics if g=g*</th>
<th>Sd of Rand statistics if g=g*</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>100% g=4</td>
<td>4</td>
<td>0</td>
<td>0.025</td>
<td>0.017</td>
<td>0.898</td>
<td>0.029</td>
<td>0.898</td>
<td>0.029</td>
</tr>
<tr>
<td>1</td>
<td>5% g=2</td>
<td>3.82</td>
<td>0.500</td>
<td>0.025</td>
<td>0.017</td>
<td>0.883</td>
<td>0.060</td>
<td>0.902</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>8% g=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>87% g=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>11% g=2</td>
<td>3.62</td>
<td>0.678</td>
<td>0.024</td>
<td>0.016</td>
<td>0.864</td>
<td>0.080</td>
<td>0.906</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>16% g=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>73% g=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.25</td>
<td>31% g=2</td>
<td>3.06</td>
<td>0.827</td>
<td>0.026</td>
<td>0.020</td>
<td>0.799</td>
<td>0.103</td>
<td>0.909</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>32% g=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>37% g=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>63% g=2</td>
<td>2.43</td>
<td>0.607</td>
<td>0.025</td>
<td>0.014</td>
<td>0.720</td>
<td>0.083</td>
<td>0.904</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>31% g=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6% g=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Real Data: 3D Contribution regression data, $g^*=3$, kmeans method
# of simulation: 100 per parameter change
Searching range: $g=2:4$
Changing parameter: the alpha as the power over $g$ in kmeans-Calinski-Harabasz method, Seed(4321)

<table>
<thead>
<tr>
<th>alpha</th>
<th>Distribution of $g$</th>
<th>Mean of $g$</th>
<th>Sd of $g$</th>
<th>Mean distance of cluster center if $g=g^*$</th>
<th>Sd of distance of cluster center if $g=g^*$</th>
<th>Mean Rand statistics</th>
<th>Sd of Rand statistics</th>
<th>Mean Rand statistics if $g=g^*$</th>
<th>Sd of Rand statistics if $g=g^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>94% $g=3$ 6% $g=4$</td>
<td>3.06</td>
<td>0.239</td>
<td>0.037</td>
<td>0.014</td>
<td>0.850</td>
<td>0.033</td>
<td>0.853</td>
<td>0.031</td>
</tr>
<tr>
<td>1</td>
<td>100% $g=3$</td>
<td>3</td>
<td>0</td>
<td>0.037</td>
<td>0.015</td>
<td>0.854</td>
<td>0.032</td>
<td>0.854</td>
<td>0.032</td>
</tr>
<tr>
<td>1.1</td>
<td>100% $g=3$</td>
<td>3</td>
<td>0</td>
<td>0.037</td>
<td>0.015</td>
<td>0.854</td>
<td>0.032</td>
<td>0.854</td>
<td>0.032</td>
</tr>
<tr>
<td>1.25</td>
<td>1% $g=2$ 99% $g=3$</td>
<td>2.99</td>
<td>0.1</td>
<td>0.037</td>
<td>0.015</td>
<td>0.852</td>
<td>0.035</td>
<td>0.854</td>
<td>0.032</td>
</tr>
<tr>
<td>1.5</td>
<td>15% $g=2$ 85% $g=3$</td>
<td>2.85</td>
<td>0.359</td>
<td>0.035</td>
<td>0.014</td>
<td>0.836</td>
<td>0.055</td>
<td>0.856</td>
<td>0.032</td>
</tr>
</tbody>
</table>
Real Data: 3D Contribution regression data, $g^*=2$, kmeans method
# of simulation: 100 per parameter change
Searching range: $g=2:4$
Changing parameter: the alpha as the power over $g$ in kmeans-Calinski-Harabasz method, Seed(4321)

<table>
<thead>
<tr>
<th>alpha</th>
<th>Distribution of g</th>
<th>Mean of g</th>
<th>Sd of g</th>
<th>Mean distance of cluster center if $g=g^*$</th>
<th>Sd of distance of cluster center if $g=g^*$</th>
<th>Mean Rand statistics</th>
<th>Sd of Rand statistics</th>
<th>Mean Rand statistics if $g=g^*$</th>
<th>Sd of Rand statistics if $g=g^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>67% g=2</td>
<td>2.54</td>
<td>0.822</td>
<td>0.153</td>
<td>0.010</td>
<td>0.720</td>
<td>0.069</td>
<td>0.760</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>12% g=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21% g=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>93% g=2</td>
<td>2.07</td>
<td>0.256</td>
<td>0.152</td>
<td>0.011</td>
<td>0.749</td>
<td>0.048</td>
<td>0.756</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>7% g=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>96% g=2</td>
<td>2.04</td>
<td>0.197</td>
<td>0.152</td>
<td>0.011</td>
<td>0.752</td>
<td>0.046</td>
<td>0.756</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>4% g=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.25</td>
<td>99% g=2</td>
<td>2.01</td>
<td>0.1</td>
<td>0.152</td>
<td>0.011</td>
<td>0.754</td>
<td>0.043</td>
<td>0.755</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>1% g=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>100% g=2</td>
<td>2</td>
<td>0</td>
<td>0.152</td>
<td>0.011</td>
<td>0.755</td>
<td>0.042</td>
<td>0.755</td>
<td>0.042</td>
</tr>
</tbody>
</table>