Vying for Dominance: An Experiment in Dynamic Network Formation

Nate Neligh

November 15, 2017

Abstract

Centrality in a network is highly valuable. This paper investigates the idea that the timing of entry into the network is a crucial determinant of a node's final centrality. We propose a model of strategic network growth which makes novel predictions about the forward-looking behaviors of players. In particular, the model predicts that agents entering the network at specific times will "vie for dominance"; that is, they will make more connections than is myopically optimal in hopes of receiving additional connections from future players and thereby becoming dominant. The occurrence of these opportunities varies non-monotonically with the parameters of the game. In a laboratory experiment, we find that players do exhibit "vying for dominance" behavior, but do not always do so at the predicted critical times. We find that a model of heterogeneous risk aversion best fits the observed deviations from initial predictions. Timing determines whether players have the opportunity to become dominant, but individual characteristics determine whether players exploit that opportunity.

1 Introduction

Network structures play a vital role in determining the behavior of many economic systems.[1] This makes it important to understand how networks form and, in particular, how some nodes end up as central (i.e. how a node ends up being "close" to many other nodes). In many settings, being close to other nodes is profitable–it means having more information, more opportunities for exchange, or more power[2]–but why are some nodes (firms, individuals, politicians) more central than others?

We hypothesize that the *timing of entry into a network* plays a critical role in determining which nodes are the most central in the eventual network. It is common wisdom in the technology industry that startup timing, that is when a new firm joins the market, is critical to eventual success. In general, it is neither the first firm to enter an industry nor the last that ends up being the most successful.[3] It is not just fundamental differences between nodes or equilibrium selection that determines whether a node becomes well connected; when the node joins the network can also have a large impact.

The following example illustrates how the order of entry may impact which nodes become central. Figure 1 shows the formation of a network where connections represent compatibilities between pieces of animation software. At the start the network contained Photoshop and Poser, which were not connected. Maya entered the network and connected to both of the existing softwares. Later, additional animation softwares joined the network, and they all connected to Maya. By 2017, Maya has become the dominant firm in the animation software market, with more compatibilities than any other software in the network.

One possible hypothesis is that Maya became dominant in the network because it joined at a critical time. By connecting to both Photoshop and Poser, Maya became central, meaning that subsequent

[1] For theoretical evidence see: Corominas-Bosch (2004); Kranton and Minehart (2001); Allouch (2015); Apt et al. (2016); McCubbins and Weller (2012); Carpenter et al. (2012). For reviews of the empirical evidence see: Bala and Goyal (2000); Jackson (2003); Jackson and Wolinsky (1996); Carrillo and Gaduh (2012). For experimental evidence see: Charness et al. (2007); Charness et al. (2014); Kittel and Luhan (2013); McCubbins and Weller (2012); Kosfield (2003).

[2] Theoretical Evidence: Kranton and Minehart (2001); Blume et al. (2009); Apt et al. (2016); Chen and Teng (2016) Empirical Evidence: Pollack et al. (2015); Sarigöl et al. (2014); Powell et al. (1996); Rossi et al. (2015)

[3] See Lilien and Yoon (1990).

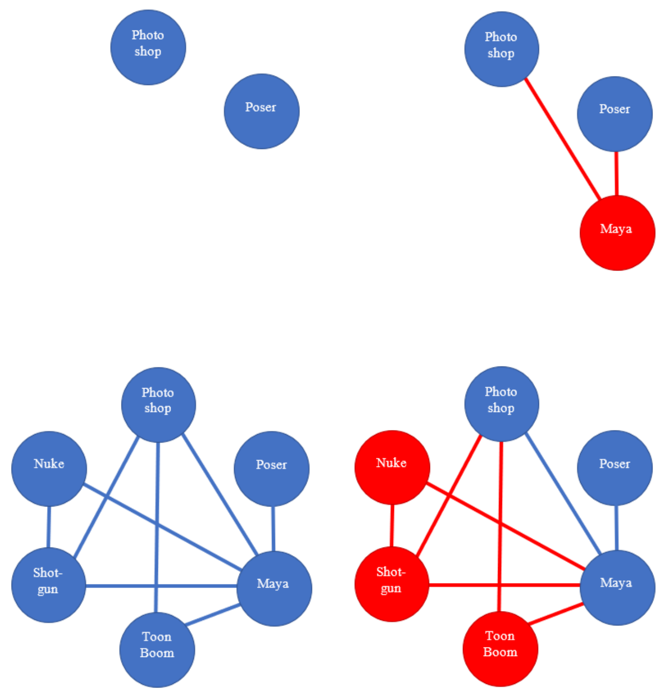


Figure 1: The formation of a network of compatibilities in animation software starting at the top left and going clock-wise.

players wanted to connect to it. Earlier players could not achieve high enough centrality to dominate the network because there were not enough players to connect to. For later players it was too expensive to challenge Maya's position of dominance.

In order to better understand how opportunities to vie for dominance arise in this type of system, we construct a dynamic model of network formation with forward looking strategic agents. In the model, players form a network by joining one at a time. As they join, players unilaterally decide which existing nodes to connect with. Centrality is beneficial, but connections are costly.[4]

Having a dynamic model of this type is essential in exploring the impact of entry timing on centrality, because entry timing is an inherently dynamic feature: players make decision taking into account their expectation of future moves. According to our hypothesis, Maya is willing to sustain the cost of connecting to both Photoshop and Poser, because it expects that this action will generate connections from other softwares in the future. As we discuss in Section 1.1, previous models of network formation have either ignored dynamics entirely, had agents who are not strategic and forward looking, or are constructed in such a way that equilibrium depends only on static features of the network.

In Section 3 we discuss the basic features of the model. When the cost of connections is high relative to the benefits of centrality, the minimally connected network is efficient, and players will form the minimally connected network in equilibrium. When the reverse is true, the maximally connected network is efficient and players will form the maximally connected network. However, we find a potentially large intermediate parameter region where players form non-degenerate networks, and behavior can be strategically rich. Two intuitively plausible moves that occur in many equilibria are a *myopic* action or *vying for dominance.*

Definition: A *myopic* move is a move which would be optimal if the game ended immediately after that move. In the games we discuss in this paper, this means making one connection to one of the most central (or *dominant*) nodes.

[4] See Neligh (2017) for more detail on the theory and an exploration of what happens when we relax some of these assumptions.

Definition: *Vying for dominance* is a move causing the player to become one of the dominant nodes immediately after his move.

Solving the unconstrained version of the game can be difficult for large networks. We focus on two approaches which simplify the game so that equilibrium behavior can be exhaustively characterized as either vying for dominance or taking a myopic action.[5]

One approach–summarized here, and covered in more depth in Neligh (2017)–involves restricting the game so that each player must connect to one of the most central nodes. Surprisingly, this simple, plausible restriction is powerful enough to limit the set of possible networks and thus keep the game solvable. In this game we find that vying for dominance happens periodically; the time between vying moves increasing exponentially as the game progresses in response to the increasing cost of vying. This result highlights a general property: players who vie for dominance do so because they expect connections from future players taking myopic actions. As the network becomes larger, and vying for dominance increasingly costly, more myopic players are needed to justify the investment made by each new vying player. The timing between vying moves must therefore increase.

In this paper we focus on a scenario that is both theoretically tractable and easily amenable to experimental testing: we restrict the maximal size of the network by limiting the number of nodes. If the network is small enough, then the only equilibrium moves are myopic or vying for dominance. This restriction, is natural for our study, since it would be impractical to study much larger network in an experimental laboratory.

We begin by solving a game with five players, the setting used in the lab. Player 5's move is simple; he should play a myopic move. The interesting behavior is that of the intermediate players, 3 and 4. Their incentives to vie for dominance naturally depend on the cost of connections, essentially creating four parameter regions. Player 4 can profitably vie for dominance in the two lower cost regions but not in the higher two. The logic is not trivial, however: a player's incentives depend also on the number of competing central nodes they face when they enter the network and, crucially, on their expectation of future nodes' behavior. There is a parameter region where Player 4 will not vie for dominance if there are too many competing central nodes. Player 3 wants Player 4 to vie, however, so he chooses not to vie in that cost region when he otherwise might have. This leads to an interesting non-monotonic relationship between the cost of connections and vying behavior for Player 3.

Section 4 describes the experiment we used to test the model of the five-node game. We ran two treatments with different costs for connections, corresponding to the second lowest and the second highest of the parameter regions previously mentioned. Using our solution to the five-node game, we predict that in the low cost treatment Player 3 should choose a myopic move and Player 4 should vie for dominance, and in the high cost treatment we should see the reverse.

In Section 5 we present the data, and compare it to the predictions of the model. In general, most players either vied for dominance or chose myopic moves as predicted. However, players often did not always vie for dominance at the predicted critical times. Broadly speaking, the model does better in predicting the play for later movers. The prediction that player 5 always plays myopically is well supported. For player 4, the comparative statics were right, but the levels were sometimes wrong. Player 4 is predicted to vie most often when costs are low, and when Player 3 did not vie, and this is the case in our data. However, Player 4s only vied 22% of the time in that scenario, far different from the 100% predicted by the theory. For Player 3 even the comparative static was wrong: Player 3s were found to vie more when costs were low.

Overall, players vied for dominance less often than expected, but when they did vie, they did so more often in conditions where the average gain from vying was higher. One possible explanation for the difference from the equilibrium prediction is that players have some aversion to vying for dominance which could be overcome with high enough payoffs. Because vying is a risky option relative to playing myopically, risk aversion is a natural candidate. In Section 6 we explore the possibility that risk aversion is influencing subjects' choices. We elicit risk preferences and find that they have significant power in predicting when players vie for dominance. On the basis of this finding we develop a version of the model with heterogeneous risk preferences. We find that this model does a good job of matching the aggregate moments in the data.

[5] With some caveats regarding the exact definition of myopic in the case of the first approach. See Neligh (2017) for more details.

Our conclusion then is that timing does play a strong role in whether nodes have the opportunity to become central However, whether nodes (individuals, managers, firms) are able to exploit the opportunity that the right timing gives them depends on individual characteristics as well.

1.1 Literature: Theory and Experiments on Network Formation

Before we present the model, we review the literature concerning theories and experiments on networks and network formation.

Previous models of network formation are not well suited to studying the role of vying for dominance and entry timing in determining network structure. These models generally attribute node centrality to either luck[6] or some combination of fundamentals and equilibrium selection[7] rather than the timing of node entry. For example, because any pairwise stable network can be a solution in Jackson and Wolinsky (1996), whether node is central depends only on whether there exists a pairwise stable network where that node is central (fundamentals) and whether that pairwise stable network happens to be the one that is chosen (equilibrium selection). Existing models of network formation usually lack at least one of the critical elements for exploring the impact of entry timing on centrality: either the agents are not forward looking, the nodes do not enter the network sequentially, or the set of solutions depends only on static features of the network.

The earliest models of network formation, such as the preferential attachment model of Yule (1925) and the small world model of Erdös and Rényi (1960) did not include any optimizing agents or strategic behavior.

Network formation models were introduced to economics by Jackson and Wolinsky (1996) and their model of cooperative network formation. In this model, a network is stable if no two unconnected players want to form a connection, and no player who is party to a connection wishes to break that connection. This network formation process is called cooperative, because two players must agree on a connection for it to persist. This model has no dynamic aspect.

Bala and Goyal (2000) propose a similar stability based network formation model, but they allowed players to generate connections unilaterally. As such, their model is referred to as non-cooperative network formation. Bala and Goyal (2000) also introduces dynamics to their models, but as in many dynamic models of network formation players were not strategically forward looking.[8] Players are assumed to best respond to the strategies of players from the previous period. In a similar vein, the model of Watts (2001) assumes that players myopically update their connections, and in the model of Kim and Jo (2009) connections only provide an immediate benefit.

There are several papers that do include dynamics as well as forward looking strategic agents. In many of these models, payoffs or game structures are chosen such that the set of possible outcomes depends on some static feature of the network. For example Currarini and Morelli (2000) and Mutuswami and Winter (2002) both find that only efficient networks can be supported in equilibria of their game. Song and van der Schaar (2015) find that the dynamic network formation process can converge to any network which satisfies a static individual rationality constraint requiring that each player make a payoff of at least zero. These papers either lack strong history dependence or use specialized payoff functions which simplify strategic considerations.

While dynamics do not drop out of the network formation model of Aumann and Myerson (1988), the payoff function used guarantees that only complete connected components can form. In other words, all nodes in a "group" must be connected to all other nodes in that group. Only the number of nodes in a particular group matters, because only one structure is possible for a given group size. This allows the network formation model to be reduced to a more standard model of dynamic coalition formation where players are picking their groups.

The model of Chowdhury (2008) is one of the most similar to our own. Both models include sequential link formation and forward-looking strategic agents. In addition, there is the possibility in Chowdhury (2008) for early movers to make myopically sub-optimal moves in hopes of gaining future connections, which can be thought of as loosely similar to the vying for dominance behavior of our

[6] Kim and Jo (2009)

[7] Watts (2001); Currarini and Morelli (2000); Jackson and Wolinsky (1996); Bala and Goyal (2000)

[8] Watts (2001) and Kim and Jo (2009)

model. However, Chowdhury (2008) assumes that each node can only sponsor one connection, and thus rules out by assumption the possibility of competing for centrality by making multiple connections that is the center of the present paper.

Our experiment is designed with a more defined temporal structure than previous network formation experiments.[9] It is this rigid time structure that allows us to carefully study how entry order and move order relate to the eventual centrality of nodes.

To our knowledge, there is only one other paper in the economics literature which has examined network growth in with a similar structure. Celen and Hyndman (2006) have players form small three-person networks using a fairly similar sequential process to the one used in this experiment. In their experiment, new players can pay to gain information about the state of the world from older nodes. The informational flows form a directed network.

Our experiment differs from Celen and Hyndman (2006) in that players care about the behavior of later nodes and the network is larger, which makes the space of possible behavior much richer. In Celen and Hyndman (2006), the behavior of future players is irrelevant. Players are instead concerned with inferring the behavior of previous players. As such, we need to conduct a new experiment to test the vying for dominance prediction of Neligh (2017) and to examine the importance of entry timing in determining node centrality.

Several experiments have found that network structure can have large impacts on behavior. Experimentalists have studied the impact of network structure on trading games,[10] public goods games,[11] and group decision making games.[12] A good review of older experiments examining the role of network structure in determining economic outcomes can be found in Kosfield (2003).

Experiments have been conducted testing various network formation models. In examining these studies, we find two consistent important findings which are potentially relevant to our own experiment. First, is the competition for centrality. Players want to be central, because it is often beneficial to be so in many experimental setups. As such, there are often several players all attempting to become central in these experiments. Second, is the role of heterogeneity. Player differences, whether inherent or exogenously given, have a large impact on the behavior of players in network formation games.

Kearns et al. (2012), van Leeuwen et al. (2013), and Goeree et al. (2007) all find evidence that competition for centrality plays a role in determining whether and how players converge to a stable solution. In all experiments, players were slow in converging to stable networks, at least in part because multiple players consistently tried to become the most central in the network. Players can be heterogeneous in how much they compete for centrality. Kearns et al. (2012) found a very bimodal distribution of connections made. Players either made a lot of connections or very few. In Section 6 we examine how heterogeneity in players can influence the competition for centrality in our experiment.

Competition for centrality is a very important feature of our model as well. Players vie for dominance by making multiple connections in hopes of being highly central and receiving many connections as the game progresses. However, while this competition for centrality has been a confounding factor in previous studies, it is a direct prediction in our model. As such, it will allow us to discuss the phenomenon with more rigor and detail than in previous studies.

2 The Game

We now present the general concept of the network formation model. There is a set of players, each one represented by a node. New players/nodes join the network one at a time. As players join the network, they choose which existing nodes to connect to. They must connect to at least one existing node. Once the last player has joined the network and made their choice the game ends, and players receive points based on the number of connections they made and their position in the final network. Centrality is beneficial but making connections is costly.

[9] Carrillo and Gaduh (2012); Bernasconi and Galizzi (2005); Carrillo and Gaduh (2012); Kearns et al. (2012); van Leeuwen et al. (2013)

[10] Charness et al. (2007)

[11] Charness et al. (2014); Carpenter et al. (2012)

[12] Kittel and Luhan (2013);McCubbins and Weller (2012)

We now present the model formally. There is a set of players represented by nodes indexed $j \in \{1, ..., J\}$. Networks are represented as $G = \{\mathbf{n}(G); \mathbf{x}(G)\}$ where $\mathbf{n}(G)$ is a set of nodes, and $\mathbf{x}(G)$ is a set of edges represents by pairs of nodes. The networks are also indexed by time as G_t where $t \in \{1, 2, ...J\}$. Note that there is one time period for every player/node, so indices are largely interchangeable. The game begins with the initial network containing only Node 1 $G_1 = \{1; \emptyset\}$.

A strategy for player j maps every possible network state they can face, G_{t-1}, to a distribution over sets of connections. Each set of connections $\mathbf{h}_t$ must be non-empty and contain only connections between Node t and existing nodes in G_{t-1}. Player t is choosing which existing nodes to connect to.

After player t makes their move, the network evolves according to the following rule:

$$G_t = G_{t-1} \cup \{t; \mathbf{h}_t\}$$

In other words, the new network is created by adding a node representing the new player and all of the connections made by that player to the existing network.

The game concludes after Player J makes his choice, generating the final network G_J.

Once the game has concluded, each player gets a payoff according to the following utility function.

$$u_i(\mathbf{h}_i, G_J) = Y - C|\mathbf{h}_i| + B\zeta_i(G_J, \delta) \tag{1}$$

$Y \in \mathbb{R}$ is a constant base payoff. $C|\mathbf{h}_i|$ is the cost of connections by individual i who purchased the set of connections $\mathbf{h}_i$. $C \in \mathbb{R}^+$ is the constant cost of connections. $B\zeta_i(G_J, \delta)$ is the benefit from centrality. $B \in \mathbb{R}^+$ is a constant multiplier, and $\zeta(G_J, \delta) = \sum_{j \neq i} \delta^{d_{ij}(G_J)-1}$ is a standard measure of closeness centrality. Decomposing $\sum_{j \neq i} \delta^{d_{ij}(G_J)-1}$, $\delta \in (0, 1)$ is a geometric discount factor. $d_{ij}(G_t)$ is the minimum distance between Node i and Node j in edges under network G_t. The minus one in the exponent adjusts the term such that we do not have to normalize B and C with respect to δ.

This type of payoff function is common in the network formation literature. It is very similar to the payoff function used in Watts (2001) and Jackson and Wolinsky (1996).[13] This type of network payoff is most relevant for systems in which some beneficial opportunity or information lands at a random node and then disseminates throughout the network with value decaying over time. It can, however, be applied as a useful approximation in any system where more central nodes gain more benefits, as this measure of centrality is highly correlated with other measures of centrality, especially in networks with low diameter.[14]

2.1 Solutions

We take Subgame Perfect Equilibria (SPE) as our solution concept of choice, because it captures the idea of fully forwards looking strategic agents. A SPE is defined in the standard manner, as a strategy profile in which players only choose moves after a given action history which are optimal for the subgame resulting from that action history. Existence is guaranteed by the fact that we are considering a finite game of perfect information. The solution to the game is not always unique. Because this is a finite game of perfect information, multiplicity of equilibria derives from the manner in which players resolve indifferences. As such it is useful to address the way players resolve indifferences in a systematic manner.

Definition. Tie-Breaking Rule: a tie-breaking rule refers to some rule by which players resolve indifferences in the construction of a SPE.

A player's tie-breaking rule can be thought of as a mapping from action histories to strict orderings over moves. Whichever strict ordering is drawn after a given action history is used to transform the current actor's weak preference ordering on moves into a strict one (thereby determining that player's move). Note that the indifferences in this game are due to structural symmetries and similarities

[13] Their payoff function is has $Y = 0$ and $B = \delta$, but otherwise is identical.

[14] For an examination of correlation in measures of centrality in real world networks, see Valente et al. (2008)

inherent in network formation and are not related to off path behavior. As such the indifferences cannot be easily dealt with using equilibrium refinements or payoff perturbations. [15]

Note that because indifference resolution is the only source of multiplicity in this game, a solution to the game can be fully characterized by the set of parameters and the tie-breaking rules employed by all players.

3 Results

3.1 Efficiency Results

We first explore the structure of the efficient network. This depends on the parameters of the problem and in particular the ratio $\frac{C}{B}$. When $\frac{C}{B} > (1-\delta)$ most of the outcomes are not Pareto ranked, but we can productively consider whether networks are efficient in the following sense.

Definition: We say an outcome network G_J is efficient if it generates the highest possible sum of utilities of all feasible outcome networks for given parameters. This is equivalent to the "strong efficiency" of Jackson and Wolinsky (1996).

The following proposition characterizes efficient networks for several parameter regions.

Proposition 1:

- If $\frac{C}{B} < 2(1-\delta)$, then the efficient network is the complete network.
- If $\frac{C}{B} > 2(1-\delta)$, then the efficient network is the star network (on Node 1 or Node 2)
- If $\frac{C}{B} = 2(1-\delta)$, then all feasible networks which contain stars are efficient

For proofs, see Neligh (2017).

This result is similar to that of Jackson and Wolinsky (1996) with a few key differences. First, the empty network is never efficient, because it is never feasible in this game. Second, the threshold below which the complete network is efficient in Jackson and Wolinsky (1996) is $C/B = 1-\delta$. This difference comes from the fact that, in our model, the cost of each connection is only paid once by the player who makes it. In the cooperative game, the connection is costly to both parties. As such connections must be twice as costly in our game before they become socially inefficient.

Note that, while the sequential nature of the game does impose limits on the set of feasible networks, it does not impose strong limits on the structure of the networks other than connectedness. Given any connected network of un-indexed nodes, we can find a sequence of actions which generates a network of that same shape.

3.2 Subgame Perfect Equilibria

Having established efficiency, we now examine the types of networks that can form in different parameter regions.

Proposition 2:

- If $\frac{C}{B} < (1-\delta)$ then the complete network is the unique network which can form in SPE's of the game.
- If $\frac{C}{B} > (1-\delta)$, then the complete network is no longer a possible outcome of any SPE at all. To see this consider the move of Player J.
- If $\frac{C}{B} > (J-1) - \frac{1-\delta^{J-3}}{1-\delta}$, then the star networks centered on Node 1 and Node 2 are the only networks which can be formed in SPE's of the game.

[15] In face, tie-breaking rules behave much like small move dependent bonus payment perturbations which are drawn randomly from a distribution which depends on the move history of the game, assuming that the payments are small enough never to change the relationship of two moves between which the player is not indifferent.

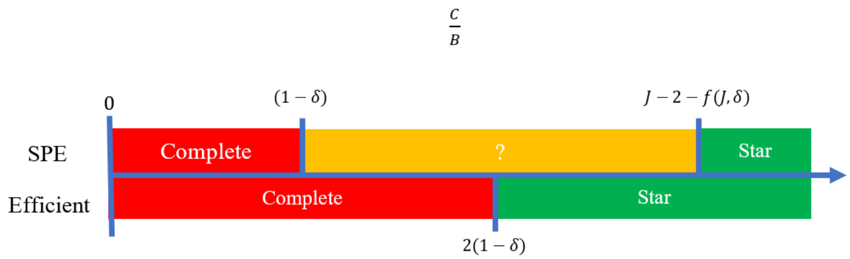


Figure 2: Visualization of parameter regions of interest. Threshold locations imply low δ. $f(J,\delta) = \frac{\delta-\delta^{J-3}}{1-\delta}$.

For proofs, see Neligh (2017).

As in Jackson and Wolinsky (1996), when $C < 1-\delta$ the complete network forms as the possible only outcome of subgame perfect equilibria, and when $C > 1-\delta$. That is where the similarities end, however. The ability of earlier moves to affect the incentives of later players means that the potential benefits of additional connections are much higher than in a one shot model; recall the vying for dominance discussed in the introduction. As such, we cannot guarantee a minimally connected network unless costs are relatively very high.

Also, note that the right hand side of the condition, $\frac{C}{B} > (J-1) - \frac{1-\delta^{J-3}}{1-\delta}$, is increasing in J, so the condition is more restrictive in large networks. Intuitively, this means that it is easier to generate non-star networks when the number of players is large and when the geometric discount factor is large.

Proposition SPE 2 is tight as long as δ is small in a weaker sense than for Proposition SPE 1. If $\frac{C}{B} < J-2$ and δ is sufficiently small such that there exist a SPE of the game parameterized by $\frac{C}{B}$ and δ which does not always generate a star network.[16]

3.3 Summary of Results

These results are intuitive and are generally quite robust to small changes in the assumptions of the model.[17] These results also provide a basis for some of the more novel results such as those discussed in Section3.5 and tested in the experiment. The results of the previous sections are summarized in Figure 2.

There are parameter regions where the star network and complete network are formed as the unique SPE outcome and regions where they efficient. There are also regions where both networks are efficient. In addition, there is an interesting region, where we cannot guarantee either the star or the complete network. In the yellow region, the complete network cannot form. The star network can form, but it is not guaranteed to be a solution, and it is never the unique SPE outcome as long as δ is small. Instead, we often see more complex strategic behaviors in the yellow region, like vying for dominance.

The above results are comparable to those found in Jackson and Wolinsky (1996) and Watts (2001) with several major differences. First, the fact that we are using non-cooperative network formation shifts the efficiency threshold, Since the region where the complete network is guaranteed does not shift, this change allows for the possibility of inefficient under-connection.

Second, the nature of the non-degenerate networks that can form is very different. The stable networks in Jackson and Wolinsky (1996) are always locally efficient in the sense that changing them by adding or removing one connection will decrease the overall sum of payoffs generated by the network. Inefficiency in their model is driven by the difference between global and local optimum. In our model, on the other hand, local optimality is not guaranteed. Inefficiency is instead driven by the existence of positive and negative externalities. The positive externalities are easy to see, because players are benefiting from connections that other players pay for, but the negative externalities are somewhat subtler. Players can vie for dominance in order to gain connections from future players, but vying

[16]See Neligh (2017) for proof

[17]See Neligh (2017) for examples of ways that the results can be generalized. In that paper we loosen assumption on the timing of entry, the homogeneity of nodes, and when nodes can make connection. We also examine what happens when players can own multiple nodes and when the end of the game is not deterministic.

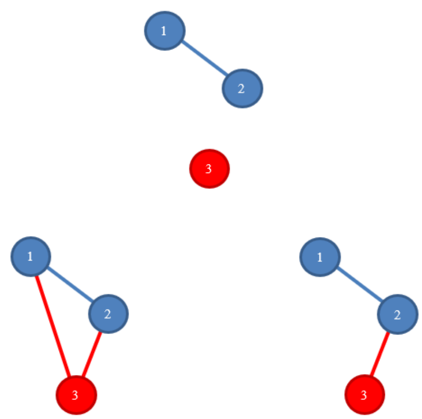


Figure 3: An unattached Player 3 (top) may choose to make a myopic move (lower right) or vie for dominance (lower left).

for dominance can produce negative externalities by crowding out players who would otherwise be dominant. An example of this is shown in Section 3.5.

It is interesting to note that, unlike in previous literature, when J is high and δ is low it is possible to generate both inefficiently under-connected networks and inefficiently over-connected networks through SPE's. Inefficient under-connection arises when $(1-\delta) < \frac{C}{B} < 2(1-\delta)$, and inefficient over-connection can arise when $(J-1) - \frac{1-\delta^{J-3}}{1-\delta} > \frac{C}{B} > 2(1-\delta)$.

3.4 Vying for Dominance

The question naturally arises of what happens in the non-degenerate region (yellow) region. Behavior in this region is rich and can include a lot of strategically interesting move such as vying for dominance and taking myopic actions. As discussed in the introduction, a *myopic* move is a move which would be optimal if the game ended immediately after that move

In all of the cases we discuss in this paper, a myopic move involves making a single connection to one of the most central (dominant) nodes, and vying for dominance corresponds to connecting to all existing nodes. See Figure 3 for examples of these behaviors. It is important to note that when we say players take a myopic action, we do not mean that they are not forward looking and strategic. As we will show, the myopic action is often optimal in the subgame perfect sense. Players can also vie for dominance.

In general, even more complex types of behavior exist in this parameter region than just these two, making it hard to solve see Neligh (2017) for an example of a six node network in which a player makes a more strategically sophisticated move: setting up a later player to vie for dominance by making a sub-optimal one connection move. As we increase the size of the network, the possible strategic complexity increases further. In addition, brute force backwards induction rapidly becomes unfeasible. In a J node network would require looking at the payoffs associated with $(J-1)!2^{J^2-J}$ possible networks.

In order to make things solvable we focus on cases in which vying and myopia are the only moves. [18] In one case we restrict the game such that players are always required to connect to at least one of the most central nodes in the network. This is called the Dominant Node Restricted Game, and it is tractable even for large networks. We can also focus on smaller simple networks such as when $J = 5$ and $C/B > 1$. We cover the Dominant Node Restricted Game in detail in Neligh (2017), while in this paper we focus on the five-node game.

[18] With some caveats discussed in detail in Neligh (2017)

3.4.1 Preview of Results from the Dominant Node Restricted Game

Before moving on to discuss the five-node game, we preview two of the results from the Dominant Node Restricted Game. First, in all Markov Perfect Equilibria of the Dominant Node Restricted Game every player will either connect to all nodes or connect to a subset of the most connected existing nodes.[19] This is a very useful result, as it allows us to eliminate a large number of possible moves for each player. If there are n players tied for most central, then Player t has $n2^{t-2}$ possible moves. This result eliminates all but 2^n of them (unless all nodes are tied for most central in which case it eliminates no moves). This result is also interesting because it predicts a very two-tiered distribution of centralities. There are players who just do what is myopically optimal without expecting many future connections, and there are players who vie for dominance in order to get connections from those players who make myopic moves.

The second result is: when players break ties in favor of connecting to newer nodes, the solution is characterized by periods of players playing myopic actions punctuated by individuals vying for dominance. Furthermore, the amount of time between players who vie for dominance increases exponentially as the game progresses. This result highlights two general the properties of vying for dominance: expected myopic moves increase the appeal of vying and expected vying moves decrease the appeal.

Players who vie for dominance do so because they want connections from the players taking myopic actions, but these connections are scarce. Only one vying player will receive the connection from a myopic player. Once another player vies for dominance, that new dominant node will be receiving those connections, due the assumed tie-breaking rule which favors newer nodes. The tie-breaking rule exaggerates the crowding out, but it is still a concern in other cases. Players who vie for dominance, therefore, need a period of non-vying nodes after their move to make up for the cost of vying. Note that this cost is increasing over time, because the network is getting larger; Vying for dominance requires connecting to all existing nodes. As such, the number of dedicated myopic actions needed to make a vying move profitable is increasing over time.

It is also interesting that very small differences in when players join the network can have very large impacts on the eventual centrality of nodes. The last node to vie for dominance ends up with a connection to every node in the network. The node who joins immediately after ends up with one connection only. Timing is critical in determining which nodes are dominant.

3.5 The Five-Node Game

In this section we will solve the game with five nodes and $C/B > 1$. [20] Recall that solutions to the game are characterized by the parameters and the tie-breaking rule. We will use the random tie-breaking rule in which each optimal move is picked with equal chance. This tie-breaking rule is chosen for two reasons. First, it doesn't ex-ante favor any node. Second, it matches well with experimental data as we show in Section 5.1. We discuss other tie breaking approaches in the Appendix Section 8.5.

We solve the game by backwards induction.

3.5.1 Player 5

Regardless of network configuration, Player 5 will always connect to a single dominant node. Player 5 will take an action that is myopically optimal: connecting to a single dominant node. This is the best one connection move, and the benefit of multiple connection moves is never worth the additional cost.

When there are multiple dominant nodes, Player 5 will connect to one at random due to the assumed tie-breaking rule. For examples of possible moves from Player 5 see Figure 4.

[19] Note that, connecting to a non-singleton subset of the current most connected nodes is not myopically optimal. Generally, however, players will connect to a singleton subset of the most connected node, which is myopically optimal, so it is best to think of this move type as myopic for comparison with later results.

[20] The game is solvable when $C/B < 1$, but the behaviors are very different from that used for other regions, so that case is covered in the Appendix Section 8.1.

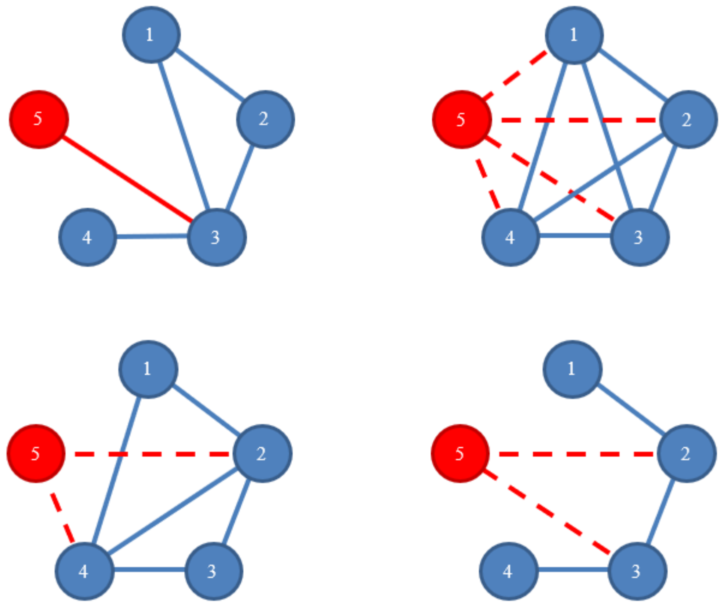


Figure 4: Several examples of possible move by Player 5. In cases with dotted lines, Player 5 picks one connection at random.

3.5.2 Player 4

Player 4's move can depend on the type of network he is facing. He can face two networks, ignoring the symmetric case. We say that two networks are symmetric if the nodes of one network can be relabeled to create the other. Because our tie-breaking rule does not depend on node labels, symmetric networks generate effectively identical behavior. The network after Player 3 connects only to Node 1 is symmetric to the network after Player 3 connects only to Node 2 If Player 3 made one connection, Player 4 faces a network with one dominant node. If Player 3 made two connections, Player 4 faces a network with three dominant nodes. See Figure 5 for examples of the networks Player 4 can face.

Player 4 Facing One Dominant Node: We first consider the case where Player 4 is facing one dominant node. There are a number of moves that Player 4 can make, but we will focus on two moves which will serve as bases for comparison: the myopic move (connecting to a single dominant node) and vying for dominance by connecting to all existing nodes.

When Player 4 plays a myopic move, he connects to a single dominant node. His single connection

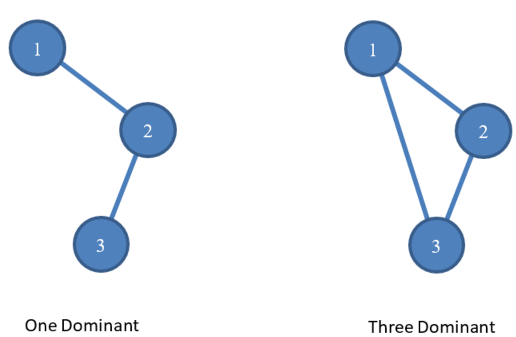


Figure 5: Possible networks faced by Player 4

Move	Cost	Immediate Benefit	Prob P5 Connection	Expected Connection Benefit	Expected Utility
Myopic	C	$B + 2\delta B$	0	δB	$B + 3\delta B - C$
Vie (3 Con)	3C	$3B$	0.5	$0.5B + 0.5\delta B$	$3.5B + 0.5\delta B - 3C$
Vie (2 Con)	2C	$2B + \delta B$	0.25	$0.25B + 0.5\delta B + 0.25\delta^2 B$	$2.25B + 1.5\delta B + 0.25\delta^2 B - 2C$
Other (2 con)	2C	$2B + \delta B$	0	δB	$2B + 2\delta B - C$
Other (1 Con)	C	$B + \delta B + \delta^2 B$	0	$0.5\delta B + 0.5\delta^2$	$B + 1.5\delta B + 1.5\delta^2 B - C$

Table 1: Move costs and benefits for Player 4 facing three dominant nodes. Moves listed: Myopic–as defined previously; Vie (3 Con)–Becoming dominant by making three connections; Vie (2 Con)–Becoming dominant by making two connections to the non-dominant nodes; Other (2 Con)–Making two connections without becoming dominant; Other (1 Con)–Making one connection to a non-dominant node.

has a cost of C, and he makes an immediate benefit of $B + 2\delta B$ (B from one directly connected node and $2\delta B$ from two second degree connected nodes). Because player 4 is not a core there is no chance that player 5 will connect to them. Player 4 will then only expect to gain δB from his second degree connection with Player 5.

Player 4 can also vie for dominance by connecting to all existing nodes. Making 3 connections incurs a cost of $3C$, and gains Player 4 an immediate benefit of $3B$. After this move, Player 4 becomes one of two dominant nodes, so he will have a 50% chance of receiving a connection from Player 5. Therefore, he will have an expected benefit of $0.5B + 0.5\delta B$ from his connection with Player 5.

Table 1 summarizes the costs and benefits for these moves and others available to Player 4 facing one dominant node.

We can immediately see that the one connection non-myopic move is dominated by the myopic move. In addition, vying for dominance by making two connections is always worse than either vying for dominance by making three connections or worse than the myopic move.[21] The remaining three possible optimal moves are Myopic, Three Connections Vie, and Two Connections Not Vie. The payoffs from all three of these options can be normalized by as functions of one summary parameter $\frac{C}{B(1-\delta)}$.

Figure 6 shows the normalized payoffs as a function of $\frac{C}{B(1-\delta)}$. We can see from the figure that vying for dominance is optimal up until some threshold, after which the myopic move is optimal which gives us the following

Lemma 1: In the five node game with random tie-breaking and $\frac{C}{B} > 1$

- If $\frac{C}{B(1-\delta)} < 1.25$, Player 4 facing one dominant node will connect to all existing nodes (Vying)
- If $\frac{C}{B(1-\delta)} > 1.25$, Player 4 facing one dominant node will connect to one dominant node (Myopic)

See Figure 7 for visualizations of the optimal moves of Player 4 facing one dominant node.

This gives us a threshold of interest, whether $\frac{C}{B(1-\delta)}$ is above or below 1.25. If we divide our regions graph from earlier through by $1 - \delta$.

Player 4 Facing Three Dominant Nodes: We now consider Player 4's move when facing three dominant nodes. In this situation, Player 4 only has three real moves due to symmetry. He can make one connect (myopic), two connections, or three connections (vying). The results of each move are reported in Table 2.

The payoffs are very similar to the one dominant node case, but now Player 4 only gets a one quarter chance of receiving a connection from Player 5 if he vies for dominance, because there will be four dominant nodes. See Figure 8 for the new normalized payoffs as functions of $\frac{C}{B(1-\delta)}$.

Vying and myopic moves are again the only optimal moves with vying being optimal up to some threshold, but the threshold decreased, because the payoffs from vying for dominance have decreased.

Lemma 2: In the five-node game with random tie-breaking and $\frac{C}{B} > 1$

[21] To see this, note two connection vie better than three connection vie implies $C > 1.25B - 1\delta B - 0.25\delta^2 B$. Two connection vie better than myopic implies $1.25B - 1.5\delta B + 0.25\delta^2 B > C$. It is impossible for both statements to be true at the same time. Therefore, vying by making two connections can never be optimal.

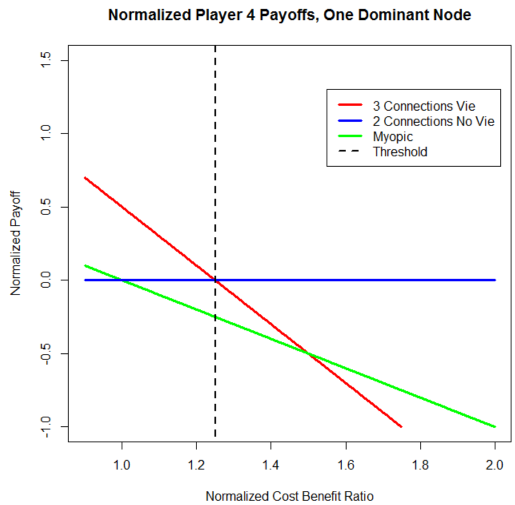


Figure 6: Normalized payoffs for Player 4 facing one dominant node. Payoffs normalized by subtracting $B + 3\delta B - C$ and dividing by $B(1-\delta)$. Due to the different structure of the payoffs, some moves (two connection vying and one connections other) cannot be normalized this way. These moves are always worse than three connection vying or the myopic move.

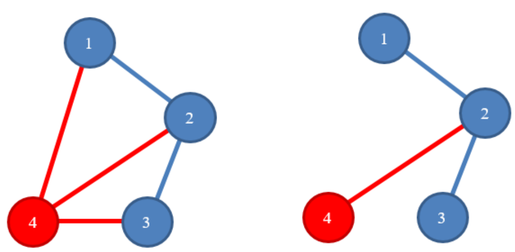


Figure 7: Visual representations of the possible moves of Player 4 facing one dominant node: vying for dominance (left) and myopic (right).

Move	Cost	Direct Benefit	Prob P5 Connection	Expected Benefit	Expected Utility
Myopic	C	$B + 2\delta B$	0	δB	$B + 3\delta B - C$
Vie	3C	$3B$	0.25	$0.25B + 0.75\delta B$	$3.25B + 0.75\delta B - 3C$
Two Connections	2C	$2B + \delta B$	0	δB	$2B + 2\delta B - C$

Table 2: Move results for Player 4 facing one dominant node

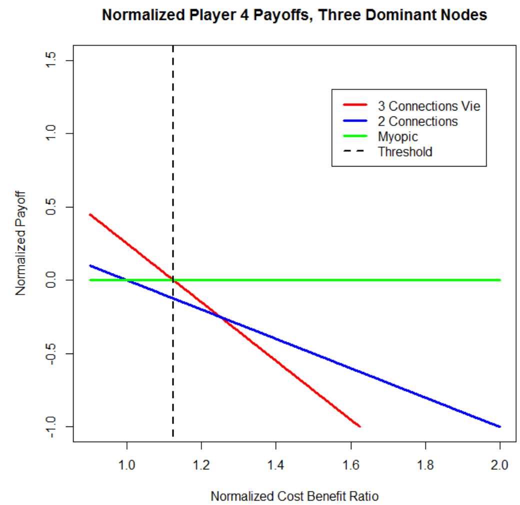


Figure 8: Normalized payoffs for Player 4 facing three dominant nodes. Payoffs normalized by subtracting $B + 3\delta B - C$ and dividing by $B(1-\delta)$

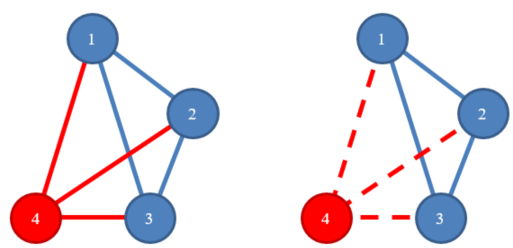


Figure 9: Visual representations of the possible moves of Player 4 facing three dominant nodes

- If $\frac{C}{B(1-\delta)} < 1.125$, Player 4 will connect to all existing nodes when facing three dominant nodes (Vying)
- If $\frac{C}{B(1-\delta)} > 1.125$, Player 4 will connect to one dominant node when facing three dominant nodes (Myopic)

See Figure 9 for visualizations of the optimal moves of Player 4 facing three dominant nodes. This adds a second threshold of interest, whether $\frac{C}{B(1-\delta)}$ is above or below 1.125.

3.5.3 Player 3

Player 3 always faces the same network: Nodes 1 and 2 connected. As such, he does not have to condition his move on network faced.

Player 3 also only has two moves (disregarding the symmetric case). He can make one connection (myopic) or two connections (vying for dominance). Player 3 does have to consider, however, how his own choice effect Player 4's choice to vie for dominance. Table 3 reflects the choices and trade-offs faced by Player 3.

When $\frac{C}{B(1-\delta)} > 1\frac{2}{3}$, vying is not worth the cost for Player 3. When $\frac{C}{B(1-\delta)}$ drops to the $(1.25, 1\frac{2}{3})$ range, vying becomes profitable. If $\frac{C}{B(1-\delta)}$ drops further into the $(1.125, 1.25)$ range, something interesting happens. When $\frac{C}{B(1-\delta)} \in (1.125, 1.25)$, whether Player 4 vies for dominance depends on

$\frac{C}{B(1-\delta)}$	P4 Vie Given P3 Vie	P4 Vie Given P3 Myopic	Vie Expected Connections	Myopic Expected Connections	Pref
$(1-\delta, 1.125)$	Yes	Yes	1.25	1	Vie
$(1.125, 1.25)$	No	Yes	2/3	1	Myopic
$(1.25, 1\frac{2}{3})$	No	No	2/3	0	Vie
$(1\frac{2}{3}, \infty)$	No	No	2/3	0	Myopic

Table 3: Player 3 trade-offs

Figure 10: Possible moves for Player 3. Vying (left) and myopic(right). In the case with the dotted line, one connection is picked at random.

whether Player 3 vies for dominance. Player 3 wants Player 4 to vie, because Player 4 vying provides a connection to Player 3 with probability one.

When $\frac{C}{B(1-\delta)} \in (1-\delta, 1.125)$, Player 4's decision no longer depends on behavior from Player 3, and vying for dominance is again profitable. We can summarize results in the following lemma.

Lemma 3: In the five-node game with random tie-breaking and $\frac{C}{B} > 1$

- $\frac{C}{B(1-\delta)} \in (1-\delta, 1.125)$, Player 3 makes two connections
- $\frac{C}{B(1-\delta)} > (1.125, 1.25)$, Player 3 makes one connection randomly
- $\frac{C}{B(1-\delta)} \in (1.25, \frac{5}{3})$, Player 3 makes two connections
- $\frac{C}{B(1-\delta)} > \frac{5}{3}$, Player 3 makes one connection randomly

See Figure 10 for a visualization of the moves Player 3 makes

Notice that there is now one more threshold we need to keep track of, whether $\frac{C}{B(1-\delta)}$ is above or below $1\frac{2}{3}$. Players 1 and 2 have no decisions to make.

3.5.4 Summary

We can then compile the behaviors into solutions for each parameter region discussed.

Proposition 3: The solution to the five-node game with $C/B > 1$ can be characterized by the following table:

$\frac{C}{B(1-\delta)}$ Range	Player 3	Player 4	Player 5
$(\frac{1}{1-\delta}, 1.125)$	Vie	Vie (Three Connections)	Myopic
$(1.125, 1.25)$	Myopic	Vie (Three Connections)	Myopic
$(1.25, 1\frac{2}{3})$	Vie	Myopic	Myopic
$(1\frac{2}{3}, \infty)$	Myopic	Myopic	Myopic

For visualizations of what typical networks in each region look like see Figure 11.

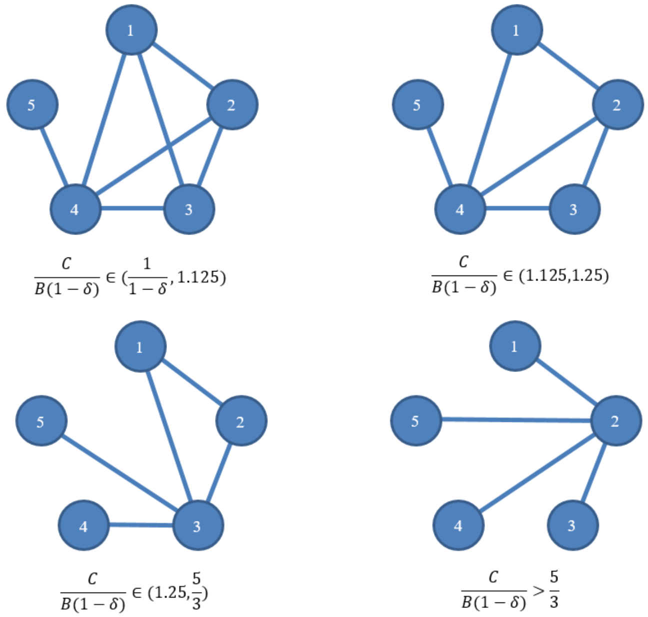


Figure 11: Typical outcome networks by region.

4 The Experiment

4.1 The Experimental Game

The experimental game uses a slightly modified payoff in order to simplify the problem for participants:

$$\pi_i(\mathbf{h}_i, G_J) = Y - C|\mathbf{h}_i| + B\sum_{j\neq i}(d_{ij} = 1) + b\sum_{j\neq i}(d_{ij} = 2)$$

Players receive B points for each directly connected node in the final network and b points for every node at a distance of two.[22] Nodes which are farther than two away provide no benefit. We conducted experiments using the following parameters:$J = 5$, $Y = 160$, $B = 100$, $b = 10$. The value of C varied between treatments with high cost treatments using $C = 140$ and low cost treatments using $C = 110$

The modified game with these parameters produces the same solution as the base game with $J = 5$, $Y = 160$, $B = 100$, $\delta = 0.1$, $C = 110, 140$. The cost levels 110 and 140 correspond to solution regions 2 ($\frac{C}{B(1-\delta)} \in (1.125, 1.25)$) and 3 discussed above ($\frac{C}{B(1-\delta)} \in (1.25, \frac{5}{3})$) with typical networks like those represented in Figure 12.

4.2 Setup

In this experiment, each round corresponded to the creation of one network for each group. Players were grouped in sets of three representing Nodes 3 through 5[23] and were randomly regrouped and reordered for each round. Each session had 28 rounds, including three practice rounds and 25 paying rounds. There was only one cost level (110 or 140) per treatment. At the end of the experiment players were paid \$1 for every 200 points earned.

Figure 13 presents an example of what players might see during the network formation process. The viewer's own node always appears in blue when it is present in the network. Potential connections appear in blue. They can be created or destroyed by clicking on existing nodes in the network. Once

[22] Note that d_{ij} in the above refers to $d_{ij}(G_J)$

[23] Nodes 1 and 2 have no choices and so no players were assigned to those roles.

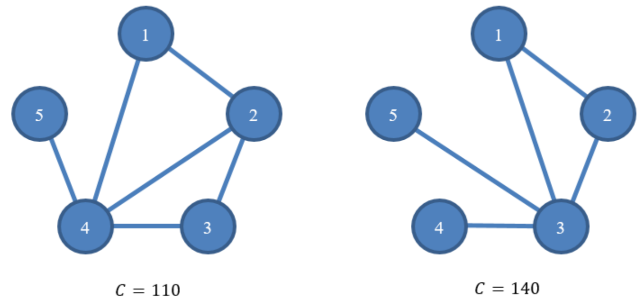


Figure 12: Typical networks formed in the experimental treatments

Session	C	Questions	Subject	Male	Female	Engineer	Econ	Grad
1	110	No	15	10	5	7	2	2
2	140	No	15	7	8	3	1	4
3	140	No	12	4	8	3	1	0
4	110	No	15	5	9	3	1	1
5	140	No	12	5	6	4	1	0
6	110	No	15	7	8	3	2	0
7	110	Yes	15	9	5	3	3	0
8	110	Yes	15	8	7	5	1	0
Total			114	55	56	31	12	7

Table 4: Sessions Summary

a player is satisfied with the set of connections they have chosen, they can click "Confirm" to finalize their connections, adding them to the network. The next player then uses the newly expanded network as the basis for their own move.

Several of the later $C = 110$ treatments included question batteries to the end of the experiment. We included sets of questions designed to risk references, elicit beliefs, and personality characteristics.

The first battery of questions consisted of a series of binary choices between gambles in a multiple price list, a la Holt and Laury (2002). The second battery elicited beliefs about the moves of successive players, incentivized via a quadratic probability scoring rule. Finally, the third battery consisted of non-incentivized Big Five personality questions, aimed at detecting subject's entrepreneurial inclinations.[24] We describe the three series of questions in more detail later in the paper. The relevant screenshots and additional details are reported in Appendix Section 10. Screenshots of the BFI questions were not included as the questions were identical to those in John and Srivastava (1999).

4.3 Sessions

We conducted all sessions at the Columbia Experimental Laboratory for the Social Sciences (CELSS). The experiment was implemented using the zTree experimental platform,[25] and participants were recruited from the CELSS subject pool which is managed using the Online Recruitment System for Economic Experiments (ORSEE).[26]

Table 4 summarizes information about the sessions that were run. The questions column refers to whether or not the question batteries were included at the end of the session. The C column refers to the cost level for the session. All other columns report demographic information of potential interest.

[24]Zhao and Seibert (2006)
[25]Fischbacher (2007)
[26]Greiner (2015)

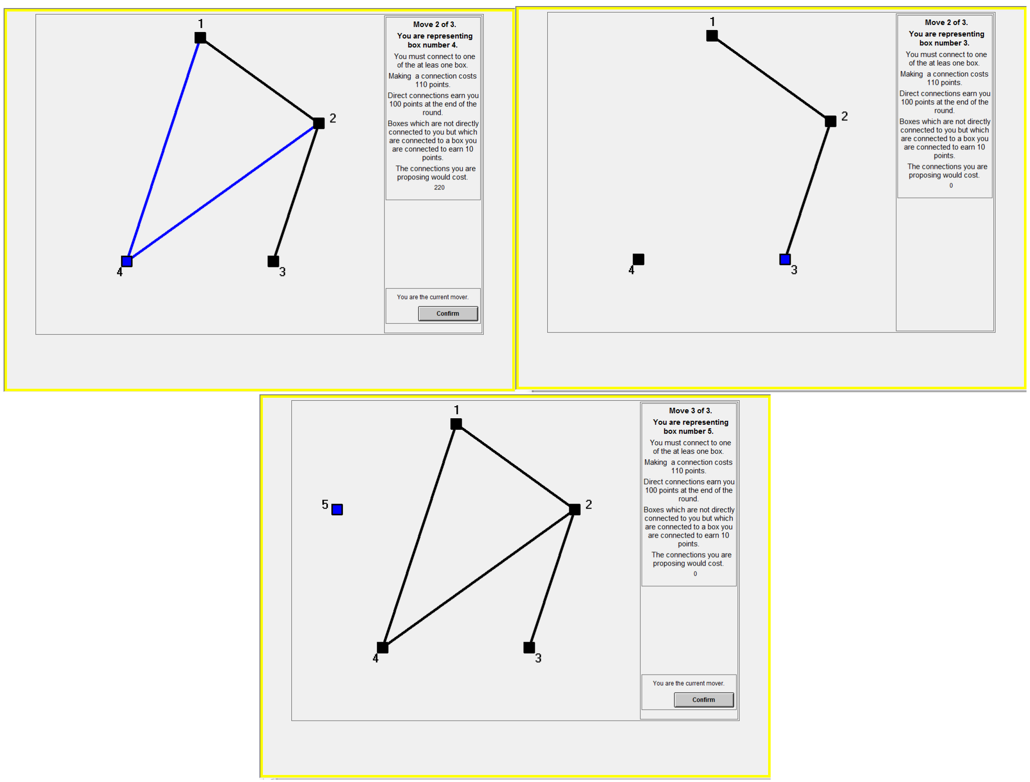


Figure 13: Examples of what players see during the network formation process. Beginning from the top left we have: (1) what Player 4 sees while making a decision; (2) what Player 3 sees while Player 4 makes a decision; and (3) what Player 5 sees after Player 4 made a decision.

Cost	Player 3	Player 4 (Facing Three Dominant Nodes)	Player 4 (Facing One Dominant Node)	Player5
110	No	No	Vie	No
140	Vie	No	No	No

Table 5: Node Move Predictions

Cost\Player	1	2	3	4	5	Total
110	447.5	337.5	270	185	180	1420
140	440	300	160	150	150	1200

Table 6: Player Payoffs: Predictions

4.4 Predictions

4.4.1 Move Predictions

Using the results from Section 3.5 we can make several predictions about the behavior of players in the experimental game.

- **Prediction 1:** All players in both treatments will either play a myopic move or vie for dominance.
- **Prediction 2: Player 5**
  - Player 5 will play a myopic move
- **Prediction 3: Player 4**
  - **3a:** Player 4 will vie more often when there is one dominant node and $C = 110$ than in any other cost/state combination
  - **3b:** Player 4 will play a myopic move when $C = 140$ or when $C = 110$ facing three dominant nodes
  - **3c:** Player 4 will vie when there is one dominant node and $C = 110$
- **Prediction 4: Player 3**
  - **4a:** Player 3 will vie more in the $C = 140$ treatment than in the $C = 110$ treatment
  - **4b:** Player 3 will vie in the $C = 140$ treatment
  - **4c:** Player 3 will Play a myopic move in the $C = 110$ treatment

Table 5 summarizes the predictions on when players should vie for dominance

4.4.2 Payoffs

Table 6 summarizes the expected number of points made by each player

At both cost levels there is a substantial early mover advantage when it comes to profit. Predictable there is a drop in payoff for all players as we move from low to high cost. This difference tends to hit the middle players harder, because they either lose their future connections or have to pay large costs if they wish to become dominant. Early players do not need to pay high costs to become dominant, and later players will not receive future connections regardless of the cost.

There are two main effects from the increase in connection costs. First, the total number of connections goes down, decreasing the efficiency of the outcome network. The high cost treatment will generate five connections while the low cost treatment will generate six. Second, the cost of each connection increases. This means that the social welfare decrease can be attributed partly to the increase in the cost of existing connections and partly to the shift to a less efficient configuration.

Cost\Player	1	2	3	4	5
110	0.25	0.25	0	0.5	0
140	0.33	0.33	0.33	0	0

Table 7: Probability of Final Dominant Node: Predictions

Because the resulting networks always have a largest minimum distance (diameter) of two in the cases we are examining, each new connection should increase welfare by $2(B-b)-C$ points.

It should be noted that the difference in configuration efficiency coupled with the predictability of connections number provide an opportunity for a welfare improving taxation/subsidy scheme. A planner could impose a flat tax mirrored by a subsidy on connections, thereby enforcing the network structure associated with any cost level. Subsidizing connections in this manner in order to generate a complete network is welfare increasing as long as $\frac{C}{B} < 2(1-\delta)$.

For example, consider what would happen if $C = 140$ and a planner were to impose a flat tax of 36 points on each player and then subsidize connections by 30 points. The tax and subsidy cancel, leaving a balanced budget, and the effective cost of connections becomes 110 points. Welfare would then go from 1200 points to $1420 - 180 = 1240$ points, a gain of 40 points.

Due to Proposition EF 1, as long as $C < 180$ it will always be optimal for the planner to impose a subsidy on connections such that the effective C is less than 90, since the complete network is the most efficient possible network in this case. Whether this theoretical gain from planner intervention can actually be practically achieved depends on whether the actual networks are responsive to changes in connection cost.

4.4.3 Dominant Nodes

The theory also has predictions about which node will be dominant at the end of the game. Table 7 provides the probability that each player will end the game as the dominant node.

In the lower cost treatment, weight is shifted toward the later nodes being the final dominant node more often. Lower costs mean more opportunities for later players to profitably vie for dominance. More vying later in the game means more efficient networks and a higher chance of newer nodes being dominant at the end.

5 Results

In this section we go over the results use them to explore the relationship between centrality and entry timing. We also discuss the relationship between costs and efficiency. The primary unit of observation is the network, although analysis is also performed at the move and subject level. In sessions with $C = 110$, we observed a total of 1875 moves by 75 subjects, or 635 networks; in sessions with C=140 we observed a total of 975 moves by 39 subjects or 325 networks.

5.1 Justification of Random Tie-Breaking Rule

Before we examine player behavior in depth, it is important to establish that the random tie-breaking rule is a credible model for the way players resolve indifferences.

The tie-breaking rule is a complex multidimensional object which would be difficult to estimate, but only some features of the tie-breaking rule actually affect optimal equilibrium play. The critical feature of the random tie-breaking rule for determining incentives is the fact that a Player 3 or 4 who vies for dominance will receive a connection from the next player with a probability equal to one over the number of resulting dominant nodes.

Table 8 compares the observed frequency of receiving a connection after vying for dominance to the probability predicted by random tie-breaking in different conditions.

While there are many significant deviations from the predicted values, they are not large enough to alter the predicted SPE strategies of the players in any situation. If we were to substitute those

Player	C	Predicted	Data	Pval	Obs
Player 3	110	0.333	0.554	0	372
Player 4, 1 Dominant	110	0.500	0.486	0	37
Player 4, 3 Dominant	110	0.250	0.312	0.003	32
Player 3	140	0.333	0.468	0	111
Player 4, 1 Dominant	140	0.500	1	1	2
Player 4, 3 Dominant	140	0.250	1	0.564	1

Table 8: Probability of Receiving a Connection from the Next Node After Vying for Dominance

Treatment	Vie	Myopic	Total	Random Benchmark Total	One Connections Other	Multiple Connections Other
$C = 110$	0.25	0.61	0.86	0.54	0.05	0.09
$C = 140$	0.13	0.78	0.91	0.54	0.07	0.03

Table 9: Aggregate Move Proportions

observed frequencies which are significantly different from the predicted values at the 5% level into the model, it would not change the predicted moves of any player in any situation.

For example, if Player 4 facing one dominant node in the $C = 110$ treatment had a 48.6% chance of receiving a connection from Player 5 after vying for dominance rather than the 50% predicted by uniform random tie-breaking, it would still be optimal for him to do so. Similarly, if Player 4 facing three dominant nodes in the $C = 110$ treatment had a 31.2% chance of receiving a connection after vying, it would still be optimal for him to choose a myopic move.

5.2 Aggregate Data and Prediction 1

In this section we examine the aggregate data and consider whether the data is consistent with prediction 1: all moves should be either myopic or vying for dominance. Table 9 reports the aggregate proportion of each move type.

The second and third column report the fractions of *myopic* and *vying* moves respectively. *Total* is the sum of the two. The random benchmark shows what proportion of moves would be myopic actions or vying for dominance if all players were to mix uniformly randomly over moves.[27] The *Other* category refers to all moves which are neither myopic actions nor vying for dominance. This category is subdivided into moves which involve one connection to a non-dominant node and moves which involve multiple connections without vying for dominance. As we can see, the proportion of vying and myopic moves is quite large, and very few players are choosing moves of other types. We can then say that Prediction 1 seems to generally hold in the data.

5.3 Player 5 and Prediction 2

In this section we look at the behavior of Player 5. Prediction 2 states that Player 5 should always connect to a single dominant node, which is the myopic move. Figure 14 shows the move proportions of Player 5.

Moves are categorized as myopic (as defined above), multiple connections, or one connection not myopic. One connection not myopic moves involve making one connection to a non-dominant node. These moves can be thought of as small errors, because they are not optimal, but the amount of points lost by choosing them is fairly small. Multiple connection moves include any move that corresponds

[27] Note, we mean each of a Player i's $(i-1) * 2^{i-2} = tot_i$ feasible moves gets equal probability. As a consequence, the probability of Player i making j connections is $\frac{BinomCoef(i-1,j)}{tot_i}$ for $j \in \{1, 2, ..., i-1\}$.

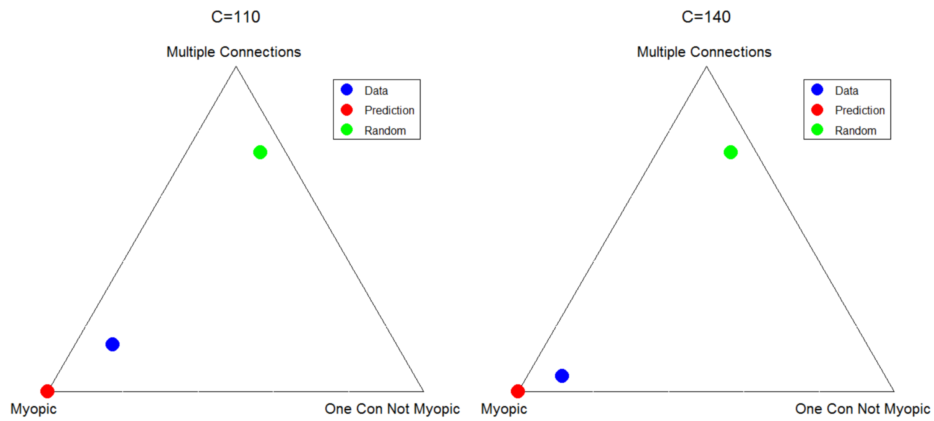


Figure 14: Player 5 move proportions

	Estimate	Std. Error	Pr(>|t|)
(Intercept)	0.220	0.043	0
C=110 3 Dominant	-0.128	0.040	0.001
C=140 1 Dominant	-0.173	0.046	0.0002
C=140 3 Dominant	-0.211	0.044	0

Table 10: Regression of Player 4 Vie Dummy on Cost/State Combination. Intercept Corresponds to C=110, One Dominant Node. Errors are clustered at the subject level.

to making more than one connection. These moves can be thought of as larger errors, because making multiple connections can be fairly costly relative to the optimal move.

The three colored dots represent the predicted move proportions (in red), the observed move proportions (in blue), and the move proportions of a hypothetical player who chose their move in a uniform random way (in green).[28] As we can see, the data is fairly close to the theoretical predictions, with most players making myopic moves. The data is also quite different from the random benchmark, so we can conclude that players are generally adhering to Prediction 2.

5.4 Player 4 and Prediction 3

We now look at the behavior of Player 4. Recall that Player 4's move can depend on the network he faces. Prediction 3 has three components. Prediction 3a is comparative, stating that Player 4 should vie for dominance more often when facing one dominant node in the low cost treatment than in any other cost/state combination. Prediction 3b states that Player 4 should play myopic moves when $C = 140$ or when $C = 110$ and facing three dominant nodes. Prediction 3a states that Player 4 should vie for dominance when facing one dominant node if $C = 110$.

Figure 15 shows the move proportions for Player 4 under different conditions. Moves are categorized differently here than in the discussion of Player 5 actions, because Player 4 may sometimes optimally choose another move: three connection vying for dominance. We must specify three connection vying, because when facing two one dominant node, Player 4 can vie for dominance by making two connections to the non-dominant nodes. In equilibrium, however, vying by making two connections is always worse than vying or making a myopic move. As before, dots of different colors are included on the figure representing the predicted move proportions, the observed move proportions, and the move proportions of a hypothetical player who chose their move in a uniform random way.

[28]Given the empirical distribution of network states faced by Player 5.

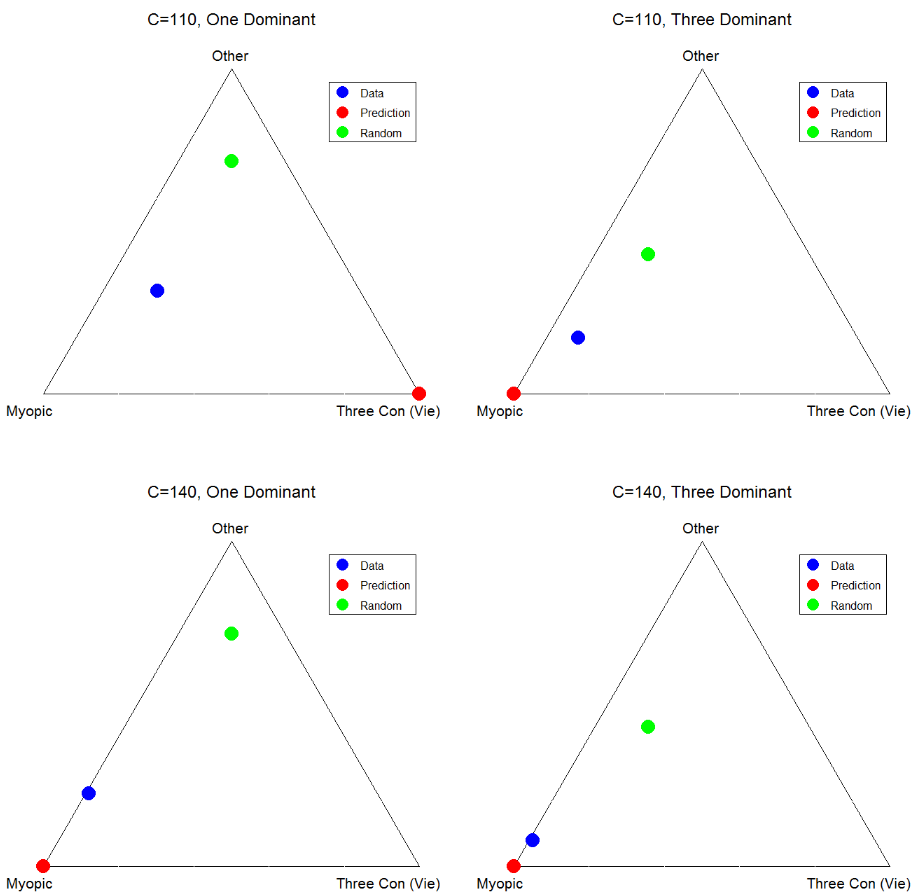


Figure 15: Player 4 move proportions

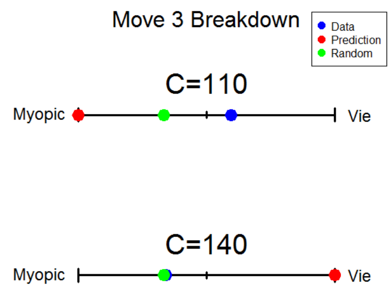


Figure 16: Move proportions for Player 3

As we can see in Figure 15 , the proportion of three connection vies is higher for Player 4 facing one dominant node in the $C = 110$ treatment than for any other situation Player 4 can face. Table 10 tests whether this difference is significant. We regressed a dummy for vying for dominance on each state/cost combination using only Player 4 data. The default situation is $C = 110$, facing one dominant node. The coefficients on every other state/cost combination are significant and negative, meaning that Player 4 vied for dominance more when $C = 110$, facing one dominant node than in any other cost/state combination. As such we can say that the comparative static of Prediction 3a is supported by the data.

We can also see by inspecting Figure 15 that Player 4 data matches the theoretical predictions fairly well in the cases where myopic actions are predicted, so we can say that prediction 3b is also generally supported. However, the data clearly contradict prediction 3c. When facing C=110, one dominant node, we see more myopic actions and more "Other" than vying, although it is vying the theory predicts in equilibrium.

One question that immediately springs to mind is whether, given the small deviations we saw earlier on the part of Player 5, it is no longer optimal for Player 4 to vie for dominance in the $C = 110$ one dominant node condition. However, Player 4's who make the three-connection vying move make an average of 1.4 more points than those making myopic moves.[29] The difference is small and not significant at the 5% level, but we can at least say that myopic moves are not better on average than vying for dominance. As such, the high number of myopic moves is still unexplained.

5.5 Player 3 and Prediction 4

Consider now the behavior of Player 3. Player 3 only has two possible actions:[30] they can either choose the myopic move or they can vie for dominance. Recall, Prediction 4a states that Player 4 should vie for dominance more in the $C = 110$ treatment than in the $C = 140$ treatment. Prediction 4b states that Player 3 should vie in the high cost treatment, and Prediction 4c states that he should play myopic moves in the low cost treatment.

Figure 16 reports the move proportions for Player 3. As before, dots represent the predicted move proportion, the actual move proportion, and the move proportion of a hypothetical player choosing an action in a uniform random manner. We can immediately see that Player 3 is actually vying more in the low cost treatment than in the high cost treatment. Furthermore, Player 3 is vying the majority of the time when $C = 110$ and not vying the majority of the time when $C = 140$. Predictions 4a, 4b, and 4c do not hold well in the data.

[29] Note that when there are multiple dominant nodes, there are multiple possible myopic moves, one corresponding to each dominant node.

[30] Ignoring the symmetric case.

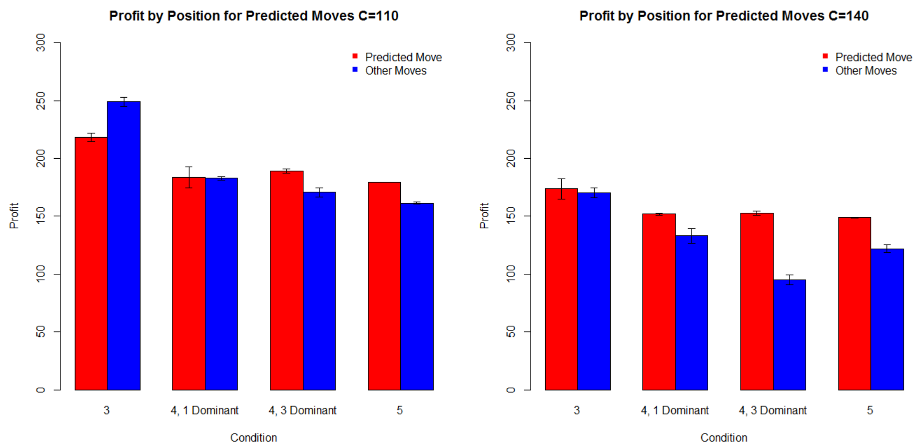


Figure 17: Observed payoffs by situation for players making the predicted move and players not making the predicted move

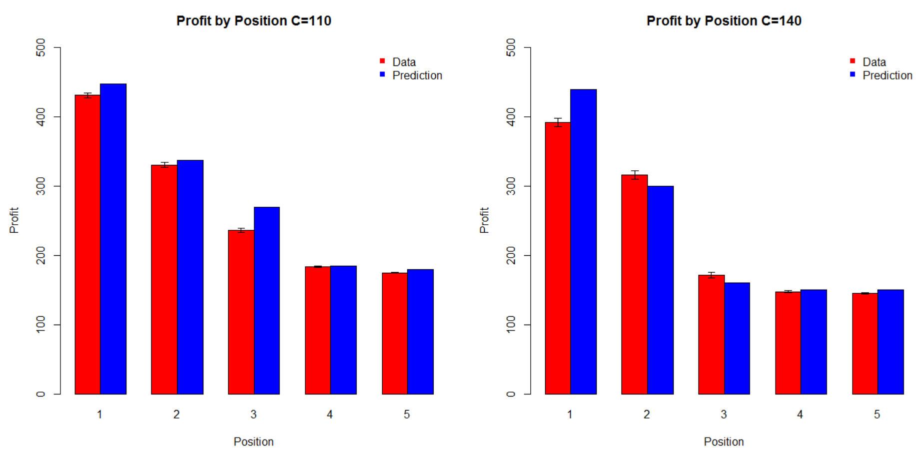


Figure 18: Average payoff by position, data and prediction. Note that payoffs for position 1 and 2 are hypothetical as no real players were associated.

Examining the payoffs of players helps reveal a possible reason for Player 3's deviation from predictions, at least in the $C = 110$ case.

Figure 17 shows the average observed payoff made by players in each position when they made the move predicted by theory vs average payoff when making all other moves. As we would expect, players generally make fewer points when making moves other than the predicted ones. Those moves should be sub-optimal in equilibrium. The one exception is Player 3 for in the $C = 110$ treatment. Because Player 4 is not vying for dominance when predicted, Player 3 is actually receiving a lower payoff for playing myopically, making vying more appealing. Player 4's deviation has changed Player 3's optimal move.

5.6 Efficiency and Dominance

This subsection examines the consequences of the observed behaviors in terms of efficiency and dominant nodes.

Figure 18 shows the predicted and actual payoffs for players in each position in both treatments. In general, payoffs were quite close to the prediction. Overall, players earned fewer points than predicted

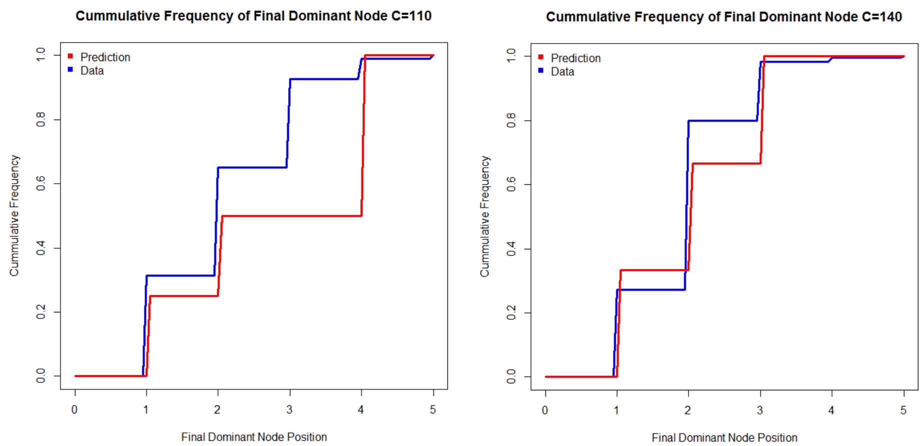


Figure 19: Cumulative frequency of final most dominant node.

with the exception of Players 2 and 3 in the high cost treatment who likely benefited from a small bias away from connecting to Node 1. Overall each network generated an average of 1357 points in the $C = 110$ treatment and 1171 points in the $C = 140$ treatment, a loss of 4.4% compared to the predicted 1420 points with C=110, and of 2.4% percent compared to the predicted 1200 points with C=140.

On average players made 5.15 connections per network in the low cost treatment vs 4.47 connections per network in the high cost treatment, suggesting that a tax/subsidy scheme may still be welfare improving, though less than predicted.

Figure 19 shows the cumulative frequencies of each node being the dominant node in the final network.[31] In the $C = 110$ case there is a definite shift back relative to predictions with earlier nodes being dominant more often than the theory would suggest. Joining the network as the third player is more conducive to becoming dominant than joining as the fourth player. The frequencies are more similar to predictions in the $C = 140$ case, except Node 2 is the final dominant node much more often than predicted, taking weight from both Player 3 and Player 1. We examine the data on final dominant nodes in more detail in Appendix Section 9.1

5.7 Summary of Results

We can summarize the results of the experiment with regards to the predictions as follows with **(T)** indicating mostly true or true and **(F)** indicating mostly false or false.

- **(T)** **Prediction 1:** All players play myopic moves or vie for dominance.
- **(T)** **Prediction 2: Player 5**
  - Player 5 will play myopic moves
- **Prediction 3: Player 4**
  - **(T)** **3a:** Vie more often when one dominant node and $C = 110$
  - **(T)** **3b:** Play myopic moves when $C = 140$ or when $C = 110$ and three dominant nodes
  - **(F)** **3c:** Vie when one dominant node and $C = 110$
- **Prediction 4: Player 3**
  - **(F)** **4a:** Vie more when $C = 140$ than when $C = 110$

[31] In the rare occurrence when multiple nodes are dominant in the final network it counts as one observation for each of them.

Treatment	Prediction	Observed
$C = 110$	0.333	0.248
$C = 140$	0.333	0.126

Table 11: Fraction of Players Vying for Dominance

	Estimate	Pr(>|z|)
(Intercept)	-1.234	0.000
Expected Gain from Vying (Vs Myopic)	0.053	0.000
AIC	1808.96	

Table 12: Logit regression of vying by connecting to all nodes against expected gain vs myopic. Errors clustered at the subject level.

- **(F) 4b:** Vie when $C = 140$
- **(F) 4b:** Myopic move when $C = 110$

In order to better understand how the reasons behind the behavioral deviations we can highlight two noticeable features of the experimental results. First as summarized in Table 11, players vie for dominance less often than predicted.

Second, players vie for dominance more in situations where the reward is higher. For each of eight possible role/conditions,[32] we calculated the average observed payoff for playing myopic moves and the average observed payoff for vying for dominance by connecting to all existing nodes.[33] These values were used to calculate the empirical average gain from vying, relative to the myopic moves. In Table 12 we show the results of a logit regression of vying for dominance against this average gain from vying. As expected, the effect is both positive and significant.

It seems plausible that players might have some aversion to vying for dominance which can be overcome by higher expected gains.

6 Risk Aversion

Vying for dominance is an inherently risky move, because players must invest in connections now in hopes of receiving future connections which, may not arrive. As such, it is plausible that risk aversion might be driving the deviations that we see in the data. We explore that possibility in the next section.

6.1 Risk Aversion Data

As described earlier, we concluded two of the session with additional questions. The first battery of questions was used to elicit risk aversion via multiple price lists (MPLs) in the spirit of Holt and Laury (2002). Each list had ten questions, with each question comparing a risky option on the left (option A) and a safe option or less risky option on the right (option B). Option A was a fixed gamble, while option B improved moving down the page. Figure 20 shows screenshots of the MPLs.

Each of the MPLs was designed to mimic the trade-off between vying for dominance and playing myopic moves in one of the situations where players frequently vied for dominance in earlier sessions. The left hand side of each list was a gamble mimicking the equilibrium payoffs from vying for dominance

[32] Player 3, Player 4 facing one dominant node, Player 4 facing three dominant nodes, and Player 5 for both cost levels

[33] Regressions showing the average impact of vying for dominance on payoffs are located in Appendix Section ??.

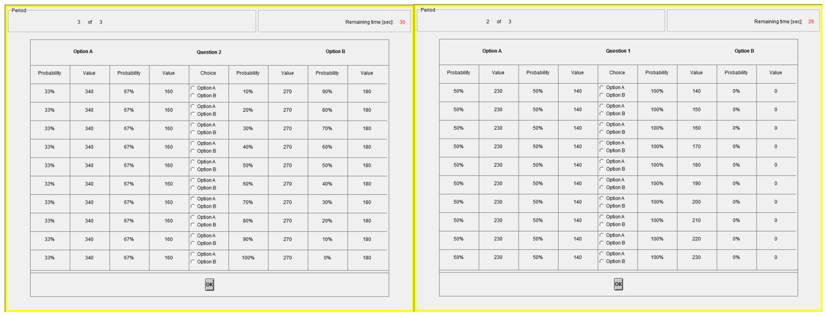


Figure 20: Screenshots of the multiple price lists used in eliciting risk preferences

in a particular situation, while the right hand contained gambles which improve going from top to bottom, including one gamble that provides the same payoffs as a myopic action in equilibrium and gambles which more closely match the observed payoffs from myopic actions The point at which the player switches gives us information about whether that player is likely to choose a myopic move or vie for dominance in each situation.

MPL 1 mimics the trade-off faced by Player 4 facing one dominant node in the $C = 110$ treatment. MPL 2 mimics the trade-off faced by Player 3 in the $C = 110$ treatment.[34] Note that while MPL 1 is comparing risky options with certain options, MPL 2 is comparing two risky options. The myopic outcome for Player 3 in the $C = 110$ treatment has an uncertain outcome, because Player 4 might vie for dominance afterwards. Appendix Section 10 contains screenshots and a more detailed discussion of list construction.

Previous experiments suggest that risk aversion measured by comparing safe options to risky options will often be very different from the risk aversion measured by comparing risky options to risky options.[35] This could lead to some differences in observed risk preferences between MPL 2 and MPL 1.

The data from the MPLs was used in two ways. First, the data was used to categorize players based on their risk preference type in each list. Based on whether each player's switch point was above or below the risk neutral switch point, the players were categorized as risk loving, risk neutral, or risk averse.[36] Players who switched the wrong way, switched multiple times, or chose first order stochastically dominated options were categorized as "undefined." We use these characterizations to provide a sene of the distribution of risk aversion and the regularity of player choices, but they are not used in further analysis.

Second, the data was used to estimate a risk aversion parameter from a CRRA model using the same stochastic model as Holt and Laury (2002).[37] Note that because we are only using the CRRA, so we assume

$$u_i(\pi) = \begin{cases} \dfrac{\pi^{1-\eta_i}}{1-\eta_i} & \eta_i \neq 1 \\ ln(\pi) & \eta_i = 1 \end{cases}$$

In their paper, Holt and Laury (2002) use a power utility function with two parameters which nests both CARA and CRRA, but we are using situation specific measure of risk aversion estimated from one list, so only need a one-dimensional measure of risk aversion. Let's focus on subjects that

[34]There was also an MPL 3 which mimics the trade-off faced by Player 3 in the $C = 140$ treatment, but that data was not used, because only sessions with $C = 110$ included question batteries. See 10 for details.

[35]This difference can be attributed to the certainty effect found by Kahneman and Tversky (1979) or to the implicit framing of gamble as buying or selling gambles as found by Hershey and Schoemaker (1985) and Sprenger (2015).

[36]We were fairly generous with our definition of risk neutral, classifying players with switch points immediately on either side of the risk neutral switch point as risk neutral. In the case of MPL 2 this means three Switch points were classified as risk neutral, because the risk neutral switch point fell exactly on one option. In MPLs 1 and 3, two switch points were classified as risk neutral.

[37]defined as $Pr(ChooseOptionA) = \frac{U(OptionA)^{1/\mu}}{U(OptionA)^{1/\mu}+U(OptionB)^{1/\mu}}$ where μ is a responsiveness parameter.

	Mean η	SDev η
MPL 1	0.401	0.429
MPL 2	1.314	1.847

Table 13: Summary Statistics Estimated η_is

	Risk Averse	Risk Neutral	Risk Loving	Undef
Panel 1	8	7	3	12
Panel 3	1	12	3	14

Table 14: Estimated Risk Preference Type

are consistent - then the only information in MPL is the switch point. Any standard measure of risk aversion should be able to fit the data perfectly, with the estimated risk aversion parameter being some monotone transformation of the switch point.

There is no theoretical reason to believe one utility function is better or to believe that given the correct specification there should be a linear relationship between the estimated risk parameter and vying behavior. As such, to show that the results are not dependent on functional form, we will use the CRRA form in the main text and repeat the analysis with the CARA form in Appendix Section 9.2. Note, there are some empirical reasons we might prefer CRRA to CARA, which are discussed in that section.

Tables 13 and 14 include information about the estimated risk aversion parameters and risk preferences types for players in each MPL.

In general, players look very different in MPL 1 with MPL 2 having substantially different risk aversion estimates. This is not surprising given the difference in the type of choices made. A large number of people have undefined risk types, and those with undefined risk types generally have multiple switch points. As such, the risk aversion measure may also contain some information about how much effort and attention players are devoting to the game.

6.2 Other Elicited Values

We also elicited beliefs and personality characteristics for each subject. This allows us to determine whether deviations from baseline predictions are due primarily to non-equilibrium beliefs or non standard preferences. the personality characteristics were included as a potential control for preference heterogeneity not captured by risk aversion. There is evidence that the Big Five personality characteristics are related to entrepreneurial activity, which can be thought analogous to vying for dominance in our game.[38]

Belief elicitation took the form of hypothetical questions placing the player in positions 3 or 4, facing specific hypothetical networks and asking the player to estimate the probability that they would receive a connection from a later node. See Appendix Section 10 for screenshots of belief elicitation questions. The questions were rewarded in a manner that made revealing one's true predicted average frequency incentive compatible for expected utility maximizing players (following Schotter and Trevino (2014)). The elicited beliefs were used to construct an expected number of connections gained from vying. Table 15 reports the additional number of connections expected after vying relative to the myopic move (on average).

We also elicited personality characteristics using the Big Five Inventory of John and Srivastava (1999). This test asks people to score their agreement with various statements on a scale of 1 to 5. These responses are then summed to create metrics of Conscientiousness, Agreeableness, Neuroticism,

[38]See Zhao and Seibert (2006)

	Avg Expected Connections Gain from Vying	SDev
Player 3	0.337	0.439
Player 4: One Dominant	0.594	0.252
Player 4: Three Dominant	0.276	0.177

Table 15: Beliefs About the Expected Number of Connections Gained of Vying for Dominance

	Openness	Neuroticism	Conscientiousness	Agreeableness	Extroversion
Mean	37	23.9	31.1	34.2	24.5
Out of	50	40	45	45	40
St Dev	5.2	5.5	6.2	5.5	6.1

Table 16: Summary of Big Five Personality Metrics. Each measure is the sum of scores from related responses (with scores reversed where appropriate). .

Openness, and Extroversion. Table 16 summarizes the mean and the standard deviation for each category among our subjects.

6.3 Risk Preference and Vying

We now look at whether risk aversion can help to explain vying behavior when controlling for other elicited characteristics in two situations: Player 3 in the $C = 110$ treatment and Player 4 facing one dominant node in the $C = 110$ treatment. We look at these scenarios in particular, because these are the only two scenarios in which players can potentially vie for dominance in an equilibrium with heterogeneous risk aversion. In all other situations, players should play myopic moves regardless of risk aversion (See next section).

In Tables 17 we regress a dummy for vying for dominance against player characteristics using data from Player 4 facing one dominant node. In all specifications, the coefficient for MPL 1 η was negative and significant. In addition, no other coefficients other than the intercept are significant in any specification.

Table 17 also shows results of similar regressions, this time predicting the vying behavior of Player 3. The Player 3 regression looks similar, although here we find that the coefficient on neuroticism is also significant and negative. It is possible that this is an artifact due to the large number of variables we are considering, but it is also possible that the neuroticism measure is capturing some features related to entrepreneurial tendency such as the ability to easily deal with new situations. Zhao and Seibert (2006) found that neuroticism is negatively correlated with entrepreneurial tendency.

In general, risk aversion appears to be influencing vying for dominance behavior. For contrast in Appendix Section9.4, we report the results of similar looking at the relationship between risk aversion as estimated from MPL 1 and the vying behavior of Player 4 facing three dominant nodes when $C = 110$. As predicted, there is not a significant impact. In addition, the coefficients on the risk aversion are substantially smaller. Some readers may be concerned that our risk aversion measures are picking up confusion or attention due to the large number of undefined subject. As we show in Appendix Section9.3, the players with undefined risk types do not have a significant impact on results.

6.4 Equilibrium

So far we have only looked at the individual choice data without considering the equilibrium effects of introducing heterogeneous risk aversion into the model. In this section we introduce a model for risk aversion and compare the equilibrium predictions of that model to the data.

The model we use is a very general model of heterogeneous risk aversion:

Variable	Player 4 Facing One Dominant Node			Player 3		
Corresponding η	-0.280^{**}	-0.286^{**}	-0.259^{**}	-0.060^{**}	-0.071^{*}	-0.077^{**}
	(0.019)	(0.035)	(0.040)	(0.033)	(0.082)	(0.032)
Intercept	1.340	0.376^{**}	0.308^{***}	1.233^{*}	0.685^{***}	0.709^{***}
	(0.279)	(0.016)	(0.000)	(0.067)	(0.000)	(0.000)
Expected Vie Gain	-0.019	-0.099		0.141	0.050	
	(0.928)	(0.599)		(0.313)	(0.743)	
Openness	-0.010			0.011		
	(0.296)			(0.278)		
Extroversion	0.001			0.001		
	(0.945)			(0.904)		
Conscientiousness	-0.002			-0.007		
	(0.867)			(0.424)		
Agreeableness	-0.018^{*}			-0.007		
	(0.100)			(0.422)		
Neuroticism	-0.001			-0.023^{***}		
	(0.946)			(0.008)		
Adj R^2	0.058	0.055	0.062	0.153	0.063	0.065
Obs	94	94	94	250	250	250

Table 17: Predicting Vying for Dominance for Player 4 Facing One Dominant Node and Player 3 $C = 110$. The Player 4 regressions use ηs estimated from MPL 1, while the Player 3 regressions use η estimated from MPL 2. Errors clustered at the individual level. ($\leq 0.1^{*}, \leq 0.1^{**}, \leq 0.01^{***}$)

Cost	3	4 (Three Dominant)	4 (One Dominant)	5
110	Yes (β)	No	Yes (α)	No
140	Yes (κ)	No	No	No

Table 18: Potential for Vying in the General Model with Risk Aversion

$$U_i(\mathbf{h}_i, G_J) = g_i(x(\mathbf{h}_i, G_J))$$

where $g(x)$ is a concave function of x. We assume that players know the population distribution of g_i but not precise g_i of other players. Further, we assume that the population is sufficiently large that seeing an action does not influence a players beliefs about conditional distribution of g_i in the remaining players. We are ruling out are risk seeking behavior and behavior inconsistent with expected utility maximization.

While the model is very general, it has fairly specific predictions for behavior in this game using backwards induction. Under this model, players can only vie in a subset of situations. In all other situations they play myopic moves.

Risk Aversion Myopic Prediction 110: When $C = 110$, Player 5 will play a myopic move, and Player 4 will play a myopic move when facing three dominant nodes. Player 3 can vie, as can Player 4 when facing one dominant node.

Risk Aversion Myopic Prediction 140: When $C = 140$, Player 5 will play a myopic move, and Player 4 will play a myopic move. Player 3 can vie.

Proof in Appendix Section 8.3. Table 18 summarizes in which situations vying may be possible.

6.4.1 Aggregate Evidence for the Risk Aversion Model

We now examine how well these aggregate predictions match the data. In Table 19 we show the vying probability in each situation, ordered from most observed vying to least. The three highest proportions

	Vie Prob	RA Prediction	RN Prediction
C=110, Node 3	0.576	> 0	0
C=140, Node 3	0.342	> 0	1
C=110, Node 4: 1 Dominant	0.151	> 0	1
C=110, Node 4: 3 Dominant	0.093	0	0
C=110, Node 5	0.011	0	0
C=140, Node 4: 1 Dominant	0.009	0	0
C=140, Node 4: 3 Dominant	0.009	0	0
C=140, Node 5	0.003	0	0

Table 19: Vying Proportions: Data and predictions from the Risk Averse Model

of vying for dominance behavior occur in the three situations where the risk averse model predicts vying might occur.

We refer to Player 4's probability of vying for dominance facing one dominant node when $C = 110$ as $\alpha \in [0, 1]$, Player 3's probability of vying for dominance when $C = 110$ as $\beta \in [0, 1]$), and Player 3's probability of vying for dominance when $C = 140$ as $\kappa \in [0, 1]$. The risk neutral version of the model corresponds to $\alpha = 1$, $\beta = 0$, and $\kappa = 1$.

Not all possible combinations of α, β, and κ can be supported by subgame perfect equilibria of given some population of utility functions. For example if $\alpha = 1$, then we must have $\beta = 0$, because then the myopic move for Player 3 in the $C = 110$ treatment would second order stochastically dominate vying.

From the data, we estimate very different values: $\tilde{\alpha} = 0.151$, $\tilde{\beta} = 0.576$, and $\tilde{\kappa} = 0.342$. While it is in general difficult to know exactly which combinations can be supported in equilibrium by some population of utility functions, we do find that these moments can be supported by a population of Modified CRRA utility functions. See Appendix Section8.4 for details and construction.
[39]

The predictions of the risk averse model are able to fit all the major moments of the data to a first approximation. Risk aversion helps explain the data both on an individual choice level and in the aggregate data through equilibrium predictions. We do consider alternative behavioral models in Appendix Section 9.5, but we do not find any of them to be promising as alternative explanations of the data.

7 Conclusion

In this paper, we presented a new network formation model that uses dynamics and forward looking strategic agents to explore novel behaviors such as vying for dominance. Here vying for dominance refers to a behavior whereby players make many connections in the present, even when doing so is myopically detrimental, in order to potentially gain connections from nodes joining in the future. The model predicts that the dominant players in the market are not dominant only because of a difference in fundamentals or equilibrium selection; the timing of a node's entry into the network is a critical factor.

We previewed some general theoretical results from the model, finding parameter regions where the star network and complete network are efficient and regions where these networks form with certainty. We also found parameter regions where such outcomes cannot be guaranteed. In general, finding solutions for the game can be difficult, but we find two ways of simplifying the game for greater tractability: restricting player moves and focusing on small networks.

In the Dominant Node Restricted Game, players are required to connect to at least one dominant

[39] Note that the shape of the utility function does matter a great deal. For example, if we only allowed unmodified CRRA utility functions in the population, we would only be able to fit data where $\beta > \alpha > \kappa$. As a consequence, we would never be able to fit our data using our estimated risk parameters, regardless of what those parameters were. This would not be useful to us, because in our data $\beta > \kappa > \alpha$.

node as they enter the network. Under this restricted model, we predict that players should vie for dominance periodically with the time intervals between vying players increasing exponentially over time. See Neligh (2017) for more details regarding the restricted game.

In the latter part of the paper we focus on the unrestricted game with five nodes. We solve the game and use that solution to make predictions about player behavior in an experimental test of the model. In equilibrium, Player 4 should vie for dominance in the low cost treatment and take a myopic action in the high cost treatment. Player 3, on the other hand, should take a myopic action in the low cost treatment and vie for dominance in the high cost treatment.

In the experiment, the predictions of subgame perfect equilibrium generally fit best for later moves. Player 5 chose myopic moves most of the time, as predicted. The comparative statics held for Player 4, but vying did not occur in large amounts even when predicted to do so. The predictions for Player 3 were very far off in both the comparative and absolute sense.

In order to explain the observed behavioral deviations from theory, we examined the possibility that players may have been risk averse. We elicited player risk preferences (as well as player beliefs and personality characteristics). Risk preferences were found to have a significant relationship with vying behavior while other characteristics did not, in general, have such a relationship in general. We also examined the equilibrium predictions of a general model of risk averse expected utility maximization. This flexible risk aversion model predicts that vying for dominance should occur in only three situations. In the data, players vied for dominance in those three predicted situations more than in any others.

Other explanations for the behavioral deviations were considered. QRE and Level-K models were discussed qualitatively and found to miss certain key features of the data. Learning appears to have very little impact on outcomes with no discernible time trends in behavior.

The results suggest that success and dominance in many systems can result from the combination of entry timing and decision making characteristics. Entry timing can provide a player an opportunity. Risk aversion determines whether the player takes it.

This conclusion suggests a natural next step: explore the role of entry timing and risk aversion in a context of immediate economic interest. In both social networks within the firm and production networks between firms we should see that the combination of good timing and low risk aversion is essential to achieving network dominance.

In addition, there are several questions about this network formation model left to explore in the lab. The applicability of the heterogeneous risk aversion model to larger networks remains to be tested. While testing the baseline model in larger networks may prove computationally difficult. The Dominant Node Restricted Game provides an appealing way to apply the lessons from this model to larger groups.

Theoretically, Neligh (2017) covers many extensions and modifications to the baseline game but sometimes not in great depth. Additional work may uncover surprising and useful findings in these different contexts.

Overall, the study of network growth with history dependence and forward looking strategic agents provides a rich avenue for research with the potential to better understand and predict many economic structures of importance. This paper serves to provide an example for future work, showing how by taking the dynamic and strategic elements of network formation serious, we can generate intuitive and novel predictions about the evolution of important systems in our world.

References

Allouch, N. (2015). On the private provision of public goods in networks. *Journal of Economic Theory*, 157:527–552.

Apt, K. R., de Keijzer, B., Rahn, M., Schaefer, G., and Simon, S. (2016). Coordination games on graphs. *International Journal of Game Theory*. Forthcoming.

Aumann, R. and Myerson, R. (1988). Endogenous formation of links between players and coalitions: An application of the shapley value. In Roth, A., editor, *The Shapley Value*. Cambridge University Press, Cambridge.

Bala, V. and Goyal, S. (2000). A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229.

Bernasconi, M. and Galizzi, M. (2005). Coordination in networks formation: Experimental evidence on learning and salience. *Nota Di Lavoro, Fondazione Eni Enrico Mattei*, (107).

Block, H. D. and Marschak, J. (1960). *Contributions to Probability and Statistics*, volume 2, chapter Random orderings and stochastic theories of responses, pages 97–132. Stanford University Press.

Blume, L., Easly, D., Kleinberg, J., and Tardos, E. (2009). Trading networks with price-setting agents. *Games and Economic Behavior*, 67(3):36–50.

Carpenter, J., Kariv, S., and Schotter, A. (2012). Network architecture, cooperation and punishment in public good experiments. *Games and Economic Behavior*, 73(1):76–90.

Carrillo, J. D. and Gaduh, A. (2012). The strategic formation of networks: Experimental evidence. *CEPR discussion paper.*

Celen, B. and Hyndman, K. (2006). Endogenous network formation in the laboratory. *Southern Methodist University, Department of Economics in its series Departmental Working Papers*, (0701).

Charness, G., Corominas-Bosch, M., and Frechette, G. (2007). Bargaining and network structure: An experiment. *Journal of Economic Theory*, 136:28–65.

Charness, G., Feri, F., Jimenez, M. M., and Sutter, M. (2014). Equilibrium selection in experimental games on network. *Econometrica*, 82:1615–1670.

Chen, W. and Teng, S. H. (2016). Interplay between social influence and network centrality: A comparative study on shapley centrality and single-node-influence centrality. *CoRR*, abs/1602.03780.

Chowdhury, A. (2008). Essays in network formation (doctoral dissertation).

Corominas-Bosch, M. (2004). Bargaining in a network of buyers and sellers. *Journal of Economic Theory*, 115(1):35–77.

Currarini, S. and Morelli, M. (2000). Network formation with sequential demands. *Royal Holloway, University of London: Discussion Papers in Economics*, (99/2).

Erdös, P. and Rényi, A. (1960). On the evolution of random graphs. In *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, pages 17–61.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Goeree, J., Riedl, A., and Ule, A. (2007). In search of stars: Network formation among heterogeneous agents. *Games and Economic Behavior*, 67(2):445–466.

Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125. http://link.springer.com/article/10.1007/s40881-015-0004-4.

Hershey, J. C. and Schoemaker, P. J. H. (1985). Probability versus certainty equivalence methods in utility measurement: Are they equivalent? *Management Science*, 31(10):1213–1231.

Holt, C. and Laury, S. (2002). Risk aversion and incentive effects. *The American Economic Review*, 92(5):1644=1655.

Jackson, M. (2003). A survey of models of network formation: Stability and efficiency. *Game Theory and Information*, (0303011).

Jackson, M. and Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, 71:44–74.

John, O. P. and Srivastava, S. (1999). *Handbook of Personality: Theory and Research*, volume 2, chapter The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives, pages 102–138. New York Guilford Press.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of choice under risk. *Econometrica*, 47(2):263–292.

Kearns, M., Judd, S., and Vorobeychik, Y. (2012). Behavioral experiments in a network formation game. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 690–704.

Kim, J.-Y. and Jo, H.-H. (2009). A signaling explanation for preferential attachment in the evolution of social networks. *Journal of Evolutionary Economics*, 20:375–393.

Kittel, B. and Luhan, W. (2013). Decision making in networks: an experiment on structure effects in a group dictator game. *Social Choice and Welfare*, 40:141–154.

Kosfield, M. (2003). Network experiments. *IEW - Working Papers*, (152). Institute for Empirical Research in Economics - University of Zurich.

Kranton, R. and Minehart, D. (2001). A theory of buyer-seller networks. *American Economic Review*, 91:485–508.

Lilien, G. and Yoon, E. (1990). The timing of competitive market entry: An exploratory study of new industrial products. *Management Science*, 36(5):568–585.

McCubbins, M. and Weller, N. (2012). Effects of network structure on costly coordination. *Association for the Advancement of Artificial Intelligence Technical Reports*, (FS-12-08).

McKelvey, R. and Palfrey, T. (1998). Quantal response equilibria for extensive form games. *Experimental Economics*, 1(1):9–41.

Mutuswami, S. and Winter, E. (2002). Subscription mechanisms for network formation. *CORE Discussion Papers*.

Neligh, N. (2017). Vying for dominance in dynamic network formation. Working Paper.

Pollack, J., Forster, W., Johnson, P., Coy, A., and Molden, D. (2015). Promotion- and prevention-focused networking and its consequences for entrepreneurial success. *Social Psychological and Personality Science*, 6(1):3–12.

Powell, W., Koput, K., and Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administratvie Science Quarterly*, 41(1):116–145.

Rossi, A., Blake, D., Timmermann, A., Tonks, I., and Wermers, R. (2015). Network centrality and pension fund performance. *CFR Wokring Papers*. University of Cologne, Centre for Financial Research.

Sarigöl, E., Pfitzner, R., Scholtes, I., Garas, A., and Schweitzer, F. (2014). Predicting scientific success based on coauthorship networks. *EPJ Data Science*, 3(9).

Schotter, A. and Trevino, I. (2014). Belief elicitation in the lab. *Annual Review of Economics*, 6:103–128.

Song, Y. and van der Schaar, M. (2015). Dynamic network formation with foresighted agents. Available online: arXiv:1509.00126v2.

Sprenger, C. (2015). An endowment effect for risk: Experimental tests of stochastic reference points. *Journal of Political Economy*, 123(6).

Stahl, D. and Wilson, P. (1994). Experimental evidence on players' models of other players. *Jourrnal of Economic Behavior and Organization*, 25(3):309–327.

Valente, T., Coronges, K., Lakon, C., and Costenbader, E. (2008). How correlated are network centrality measures? *Connect (Tor)*, 28(1):16–26.

van Leeuwen, B., Offerman, T., and Schram, A. (2013). Superstars need social benefits: An experiment on network formation. *Working Papers at the Department of Business*.

Watts, A. (2001). A dynamic model of network formation. *Games and Economic Behavior*, 34:331–341.

Yule, G. U. (1925). A mathematical theory of evolution, based on the conclusions of dr. j. c. willis, f.r.s. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 213(402-410):21–87.

Zhao, H. and Seibert, S. E. (2006). The big five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology*, 91(2):259–271.

8 Appendix A: Additional Mathematical Details

8.1 5 Node Game When $C/B < 1$

This section assumes you have already read Section 3.5. Note that when $C/B < 1-\delta$ we know the complete network forms by Proposition SPE 1. However, we still have not covered the solution to the 5 node game when $C/B \in (1-\delta, 1)$. In the region, the behavior of Player 5 changes. While Player 5 will still connect to a single node when facing most configurations, when facing a chain,[40] he will connect to any pair of nodes other than the first two nodes or the last two nodes on the chain.

This means that Player 4 when faced by a chain can either choose to vie for dominance by making 3 connections (payoff $3.5B + 0.5\delta B - 3C$), play myopic moves (payoff $B + 3\delta B - C$), or connect to one end and lengthen the chain (payoff $1.5B + 1.5\delta B + \delta^2 B - C$). No other moves can be optimal

Lengthening the chain is preferred to a myopic move when $0.5 + \delta^2 > 1.5\delta$ or when $\delta \in (0, 0.5)$. Lengthening the chain is preferred to three connection vying when $C/B > 1 - \delta/2 - \delta^2/2$. Therefore Player 4 will choose to lengthen the chain if $\delta \in (0, 0.5)$ and $\frac{C}{B(1-\delta)} > 1.25$ or when $1.25 > \frac{C}{B(1-\delta)} > 1 + \frac{\delta-\delta^2}{2(1-\delta)}$.

If Player 4 chooses to extend the chain, after Player 3 plays a myopic move, Player 3 receives $2B + \frac{13}{8}\delta B + \frac{3}{8}\delta^2 B - C$ which is preferred to vying when $C/B > \frac{2}{3} - \frac{7}{24}\delta - \frac{3}{8}\delta^2$.

Therefore, if either $\delta \in (0, 0.5)$ and $\frac{C}{B(1-\delta)} > 1.25$ or $1.25 > \frac{C}{B(1-\delta)} > 1 + \frac{\delta-\delta^2}{2(1-\delta)}$, then if $C/B > \frac{2}{3} - \frac{7}{24}\delta - \frac{3}{8}\delta^2$, Player 3 will play a myopic move, and $C/B < \frac{2}{3} - \frac{7}{24}\delta - \frac{3}{8}\delta^2$, he will vie. Player 4 will then lengthen the chain, and Player 5 will connect to a random pair of nodes that is not the first or last two nodes on the chain. If neither $\delta \in (0, 0.5)$ and $\frac{C}{B(1-\delta)} > 1.25$ nor $1.25 > \frac{C}{B(1-\delta)} > 1 + \frac{\delta-\delta^2}{2(1-\delta)}$ holds, then outcomes are as reported in Section 3.5.

8.2 Prediction 1 Robustness

Robustness Result: Prediction 1 is robust to the inclusion of behavioral noise in the following sense. Assume $C > B$, $J < 5$, and players have a utility function $u(x, \eta)$.We assume that $u(x, \eta)$ satisfies we could be using any functional form which satisfies (1) $u(..., \eta)$ is strictly more risk averse than $u(..., \eta')$ as long as $\eta > \eta'$ (2) $u(x, 0)$ is linear in x, and (3) $u(..., \eta)$ becomes arbitrarily risk averse as η goes to infinity. We also assume that $\eta \geq 0$, so players are risk neutral or risk averse, not risk seeking.

In addition, assume players have a $\rho \in [0, 1]$ probability of randomly mixing over moves and a $1-\rho$ probability of selecting their optimally in a subgame perfect manner, then it still always optimal for each player to either connect to a single dominant node or vie for dominance.

This is primarily a statement about Player 4 as Player 5's optimal choice does not depend on his beliefs about the behavior of others, and Player 3 must satisfy the prediction trivially.

[40] A chain is a network wherein all but two nodes in the are connected to precisely two other nodes with no other connections. The last two nodes are each connected to precisely one node. This type of network is the most "spread out" a network can be while remaining connected.

Proof of Robustness: To start, note that Player 5's optimal move is risk free in the no-risk aversion model, so introducing risk aversion will not change it.

In this proof all of the conditional and unconditional probabilities of move types by Player 5 were found by brute force counting. If Player 4 does not vie for dominance then his maximum chance of receiving a direct connection from Player 5 is $\rho\frac{8}{15}$. Assuming that he made k connections, his maximum chance of receiving a second degree benefit from Player 5 is $\rho\frac{7}{15}f(k) + (1-\rho)$. Here $f(k)$ is the probability that Player 5 will connect to at least one of the k nodes that Player 4 is connected to without connecting directly to Player 4, conditional on Player 5 acting randomly.

Note that $f(1) = \frac{4}{7}, f(2) = \frac{6}{7}, f(3) = \frac{7}{7}$. Therefore, if Player 4 connects to a single dominant node, he gets a utility of $\rho\big(\frac{8}{15}u(2B+2b-C,\eta) + \frac{7}{15}f(1)u(B+3b-C,\eta) + \frac{7}{15}(1-f(1))u(B+2b-C,\eta)\big) + (1-\rho)u(B+3b-C,\eta)$

. Connecting to a single non-dominant node will get him strictly less. By connecting to three nodes, Player 4 will automatically vie for dominance, so it suffices to show that connecting to two nodes is always worse than connecting to a single dominant node. By connecting to two nodes without vying for dominance, Player 4 makes at most $\rho\big(\frac{8}{15}u(3B+b-2C,\eta) + \frac{7}{15}f(2)u(2B+2b-2C,\eta) + \frac{7}{15}(1-f(2))u(2B+b-2C,\eta)\big) + (1-\rho)u(2B+2b-2C,\eta)$.

The gain from the second connection is maximized when $\rho = 1$, since $u(B+3b-C,\eta) > u(2B+2b-2C,\eta)$.

We then compare $\frac{8}{15}u(2B+2b-C,\eta) + \frac{7}{15}f(1)u(B+3b-C,\eta) + \frac{7}{15}(1-f(1))u(B+2b-C,\eta)$ to $\frac{8}{15}u(3B+b-2C,\eta)+\frac{7}{15}f(2)u(2B+2b-2C,\eta)+\frac{7}{15}(1-f(2))u(2B+b-2C,\eta)$. Note that$u(2B+2b-C,\eta) > u(3B+b-2C,\eta)$, so if we can show that $f(1)u(B+3b-C,\eta) + (1-f(1))u(B+2b-C,\eta) > f(2)u(2B+2b-2C,\eta) + (1-f(2))u(2B+b-2C,\eta)$. This must be the case, because $B+3b-C > B+2b-C > 2B+2b-2C > 2B+b-2C$. $\square$

8.3 Risk Aversion Solution

In this section we provide detailed predictions of the 5 node game when players have risk aversion as described in Section 6.4.

Risk Aversion Prediction 1: In the SPE of the game with heterogeneous risk averse utilities in which players adopt a random tie-breaking approach, if $C = 110$, then the equilibrium has the following features: Player 5 will connect to a single dominant node. Player 4 may. Otherwise, Player 4 will connect to a single dominant node. Player 3 may vie for dominance or choose a myopic move depending on the behavior of Player 4 and his own utility function.

Proof of Risk Aversion Prediction 1: Player 5's behavior follows from the proof of SPE Prediction 1. Player 4, knowing this, is willing to become a dominant node in order to gain a potential connection from Player 5 as long as the cost is low and the probability of the connection is high. Consider first what happens if Player 4 is facing a chain (which implies one dominant node).

If Player 4 faces a complete network he will either connect to one node or all nodes. Connecting to two nodes is strictly worse than connecting to one, because it has higher costs with no benefit. Connecting to one node he again receives $g_i(Y+B-C+3b) = g_i(180)$ points. Connecting to all nodes gives $0.25g_i(Y+4B-3C) + 0.75g_i(Y+3B-3C+b) = 0.25g_i(230) + 0.75g_i(140)$ points. In this case, Player 4 will connect to one node, because that is the dominant option.

If Player 4 faces a chain, he can either connect to the dominant node earning $g_i(Y+B+3b-C) = g_i(180)$;[41] connect to the ends of the chain becoming one of four dominant nodes and earning $0.25g_i(Y+3B-2C+b)+0.5g_i(Y+2B-2C+2b)0.25g_i(Y+2B-2C+b) = 0.25g_i(250)+0.5g_i(160)+0.25g_i(150)$; or connect to all three nodes nodes becoming one of two dominant nodes earning $0.5g_i(Y+4B-3C)+0.5g_i(Y+3B-3C+b) = 0.5g_i(230) + 0.5g_i(140)$, where the 0.5's come from a one half chance that Player 5 will connect to node 4 and a one half chance Player 5 will connect to the other dominant node.

Making one connection second order stochastically dominates making two, so Player 4 will always either make one connection to a dominant node or three connections when facing a chain, depending on the shape of $u_i()$. Say that the probability of Player 4 having a $g_i()$ such that vying for dominance is optimal is α.

[41] connecting to one non-dominant node is strictly worse

Player 3 can either connect to one node or two. By connecting to a single node, Player 3 gets

$$\alpha g_i(Y + 2B + 2b - C) + (1 - \alpha)g_i(Y + B + 3b - C) = \alpha g_i(270) + (1 - \alpha)g_i(180)$$

By making two connections, Player 3 gets

$$\tfrac{1}{3}g_i(Y + 4B - 2C) + \tfrac{2}{3}g_i(Y + 2B + 2b - 2C) = \tfrac{1}{3}g_i(340) + \tfrac{2}{3}g_i(160)$$

.

Note that if $\alpha > 5/9$ making one connection second order stochastically dominates making two connections. This can be shown by examining the area under the cumulative distribution functions for each distribution. One can also see that when Player 3 is sufficiently risk averse, he will make one connection, because the minimum payment for that choice is higher. If Player 3 is risk neutral, he will prefer to make two connections if $\alpha < 4/9$ and he will make one connection if the reverse is true. □

Now we consider the equilibrium when $C = 140$.

Risk Aversion Prediction 2: In the SPE of the game with heterogeneous risk averse utilities in which players adopt a random tie-breaking approach, if $C = 140$, then the equilibrium has the following features: Players 4 and 5 will choose myopic moves. Player 3 may choose a myopic move or may vie for dominance depending on the shape of his utility function.

Proof of Risk Aversion Prediction 2: Player 5 will connect only to a single dominant node as before for the reasons given in previous proofs. Additional connections cannot provide Player 5 enough benefits to make up the costs.

If Player 4 faces a complete network he will either connect to one node or all nodes. Connecting to two nodes is strictly worse than connecting to one, because it has higher costs with no benefit. Connecting to one node he again receives $g_i(Y + B - C + 3b) = g_i(150)$ points. Connecting to all nodes gives $0.25g_i(Y + 4B - 3C) + 0.75g_i(Y + 3B - 3C + b) = 0.25g_i(140) + 0.75g_i(90)$ points. In this case, Player 4 will connect to one node, because that is the dominant option.

If Player 4 faces a chain, he can either connect to the dominant node earning $g_i(Y + B + 3b - C) = g_i(150)$;[42] connect to the ends of the chain becoming one of four dominant nodes and earning $0.25g_i(Y + 3B - 2C + b) + 0.5g_i(Y + 2B - 2C + 2b)0.25g_i(Y + 2B - 2C + b) = 0.25g_i(190) + 0.5g_i(100) + 0.25g_i(90)$; or connect to all three nodes nodes becoming one of two dominant nodes earning $0.5g_i(Y + 4B - 3C) + 0.5g_i(Y + 3B - 3C + b) = 0.5g_i(140) + 0.5g_i(50)$, where the 0.5's come from a one half chance that Player 5 will connect to node 4 and a one half chance Player 5 will connect to the other dominant node. Connecting to one dominant node second order stochastically dominates all other options.

Player 3 can then make one connection receiving $g_i(Y + B - C + 3b) = u(150)$ or make two connections earning

$$\tfrac{1}{3}g_i(Y + 4B - 2C) + \tfrac{2}{3}g_i(Y + 2B + 2b - 2C) = \tfrac{1}{3}g_i(280) + \tfrac{2}{3}g_i(100)$$

If $\frac{1}{3}g_i(280) + \frac{2}{3}u_i(100) > u_i(150)$ Player 3 will make two connections. If the reverse is true, he will make one connection. □

8.4 Finding a Population of Utility Functions that Support the Data

Here we examine whether the moments $\tilde{\alpha} = 0.151$, $\tilde{\beta} = 0.576$, and $\tilde{\kappa} = 0.342$ can be supported be a subgame perfect equilibrium of the game with heterogeneous risk aversion.

Empirical Risk Aversion Proposition: There exists a population of utility function which can generate the observed vying proportions ($\tilde{\alpha}$, $\tilde{\beta}$, and $\tilde{\kappa}$) as outcomes of a subgame perfect equilibrium of the game with heterogeneous expected utilities.

Proof of Empirical Risk Aversion Proposition: For convenience we will refer to the situations in which players may vie by their corresponding Greek letters (For example Player 3 in the $C = 110$ treatment is situation β). Given that we have observed $\beta > \kappa > \alpha$ and estimate the population α, we are going to want a set of four utility $\{g^1, g^2, g^3, g^4\}$ functions with the following properties:

- Players will vie in situations α, β, and κ if they have g^1 as their utility function.

[42] connecting to one non-dominant node is strictly worse

- Players will vie in situations β and κ if they have g^2 as their utility function but not in situation α.
- Players will vie in situations β if they have g^3 as their utility function but not in situations α and κ.
- Players will never vie if they have g^4 as their utility function.

Note that what constitutes a valid set of utility functions will change based on α, because Player 3's incentives change based on the behavior of Player 4 in the $C = 110$ treatment. Similarly, if one observed a different ordering over vying proportions, one would need to switch the situations for which each utility function predicts vying in the corresponding manner.

If we can find this set, we can easily construct a population which generates the observed values for α, β, and κ. Simply say $\tilde{\alpha}$ of the population has utility function g^1, $\tilde{\kappa} - \tilde{\alpha}$ of the population has utility function g^2, $\tilde{beta} - \tilde{\kappa}$ of the population has utility function g^3, and everyone else has g^4. Using the equilibrium characterization from Appendix Section 8.3, this gives us our result.

Now all that remains to be seen is whether such a set of function exists that can fit our observed. Consider utility functions of the following type:

$$g(x, b, \eta) = \frac{(b + x)^{1-\eta}}{1 - \eta}$$

We can satisfy all the requirement by picking $g^1 = g(x, 2000, 5)$, $g^2 = g(x, 2000, 5.5)$, $g^3 = g(x, 2000, 5.7)$, $g^4 = g(x, 2000, 10)$.

8.5 Comparison of Tie Breaking Approaches

We use random tie-breaking as our primary benchmark, because it seems to be most plausible in the data, but other tie-breaking approaches can lead to other subgame perfect equilibria.

It can be illustrative to compare the random tie-breaking results to those derived from other tie-breaking approaches. This comparison can provide the reader with a better understanding of the range of behaviors that are possible under SPE. We will be comparing the uniform random tie-breaking approach to two other tie-breaking approaches which are opposite extremes. Under the stability seeking tie-breaking approach, players resolve indifferences in favor of connecting to the oldest node. [43] Under the novelty seeking tie-breaking approach players resolve indifferences by connecting to the newest node.[44] For the sake of simplicity we will assume risk neutrality throughout this section.

Figure 21 shows how the total connectivity of the resulting network under each tie-breaking approach changes with C. It is important to note that the tie-breaking approach which yields the most connections changes as C increases. Furthermore, we can immediately determine which tie-breaking outcomes are more efficient. To the left of the purple line, more connected outcomes are more efficient, while to the right of the line they are less so. Stable, random, and novel tie-breaking all have parameter regions where they yield the most efficient outcome.

Figure 22 shows the payoffs made by each player in expectation in each of the discussed equilibria. Payoffs for all players other that Player 1 tend to decrease as C increases, since these players must make at least one connection. The payoffs for Players 3-5 converge as C increases, because for high C they each make one connection resulting in a star network.

Under the random tie-breaking approach, there is a very thin spike around $C = 100$ indicating the changeover from the mode wherein Player 5 sometimes makes two connections to the mode wherein Player 5 makes exactly one connection. Player 2 does better than the other connection making players under both the random and stability seeking tie-breaking approaches, because he can be the center

[43]When choosing between nodes with multiple connections, look first at the oldest node in each move then going on to the second oldest in each and so on until one is older. An absence of nodes is considered to have an age of 0. Most of these caveats will not come into play in the experimental game, but they are important in order to make sure that the tie-breaking approach always produces a unique well defined response.

[44]When choosing between nodes with multiple connections, look first at the newest node in each move then going on to the second newest in each and so on until one is newer. An absence of nodes is considered to have an age of 0. Again, most of these caveats will not come into play in the experimental game.

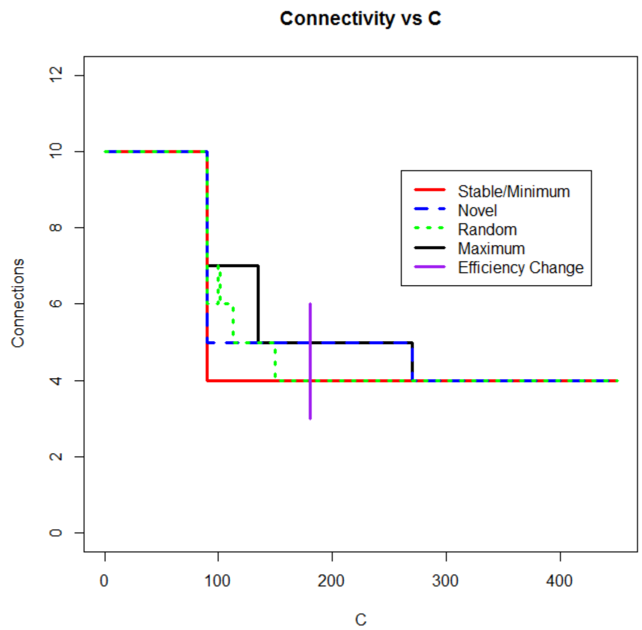


Figure 21: Connectivity versus cost. Purple line indicates the changeover between connections being more efficient and connections being less efficient.

of the resulting star. In is notable the, under the novelty seeking tie-breaking approach, when $135 < C < 180$ for Player 3 is greater than the payoff for Player 2. This is a rarity. In general earlier players have more opportunities to receive future connections and hence they have higher payoff. The early mover almost always has the advantage as far as payoffs are concerned.

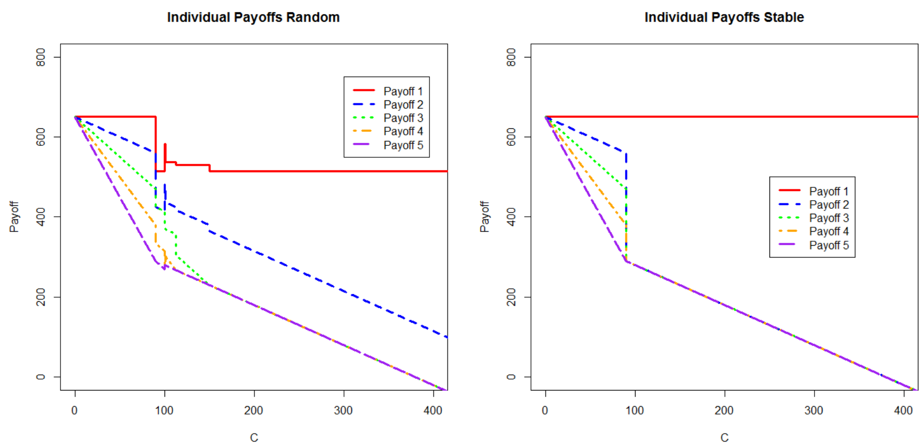
Figure 23, provides a graph of gini coefficient of expected payoffs for each tie-breaking approach discussed. Looking at all nodes (Upper Left), it seems like the equality of the outcomes is well ranked with tie-breaking approaches that favor to later nodes leading to a more equal outcomes. The inequality is generally increasing with C except potentially at the mode change point of $C = 90$. This increase is largely driven by two factors. First, Player 1 does not have to make a connection, and therefore is unaffected by connection cost. Second, only Player 1 and Player 2 can be the center of the star network for high C.

If we eliminate Player 1, the smoothness of the increase goes away for most tie-breaking approaches (see Figure 23 Upper Right). If we eliminate both Player 1 and Player 2 then we see a non-monotonic relationship between inequality and C as in Figure 23 (Lower), usually peaking at $C = 90$. In general, decreasing C from 140 to 110 should decrease the inequality of the expected outcome regardless of which of these tie-breaking approaches are employed. This result does not change if one eliminates Player 1 and Player 2 from consideration.

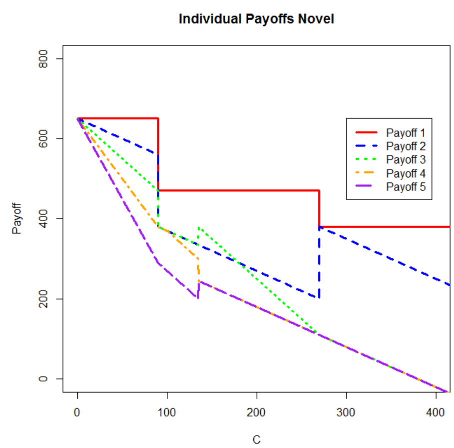
9 Appendix B: Additional Data

9.1 Node Dominance in the Final Network

The theory gives us predictions about which nodes should be the most connected nodes at the end of the network formation process. In the $C = 110$ treatment, Node 4 should be the most connected node with frequency 0.5, and Nodes 1 and 2 will be the most connected final node the remainder of the time, split evenly between them. In the $C = 140$ treatment, Node 3 should all have a 1/3 chance of being the final dominant node, and Nodes 1 and 2 should each have a 1/3 chance. The histograms of the frequency with which each node was one of the final most connected are presented in Figure 24.



Vying for Dominance: An Experiment in Dynamic Network



Formation

Figure 22: Payoffs in the experimental for each player by cost under different tie-breaking approached.

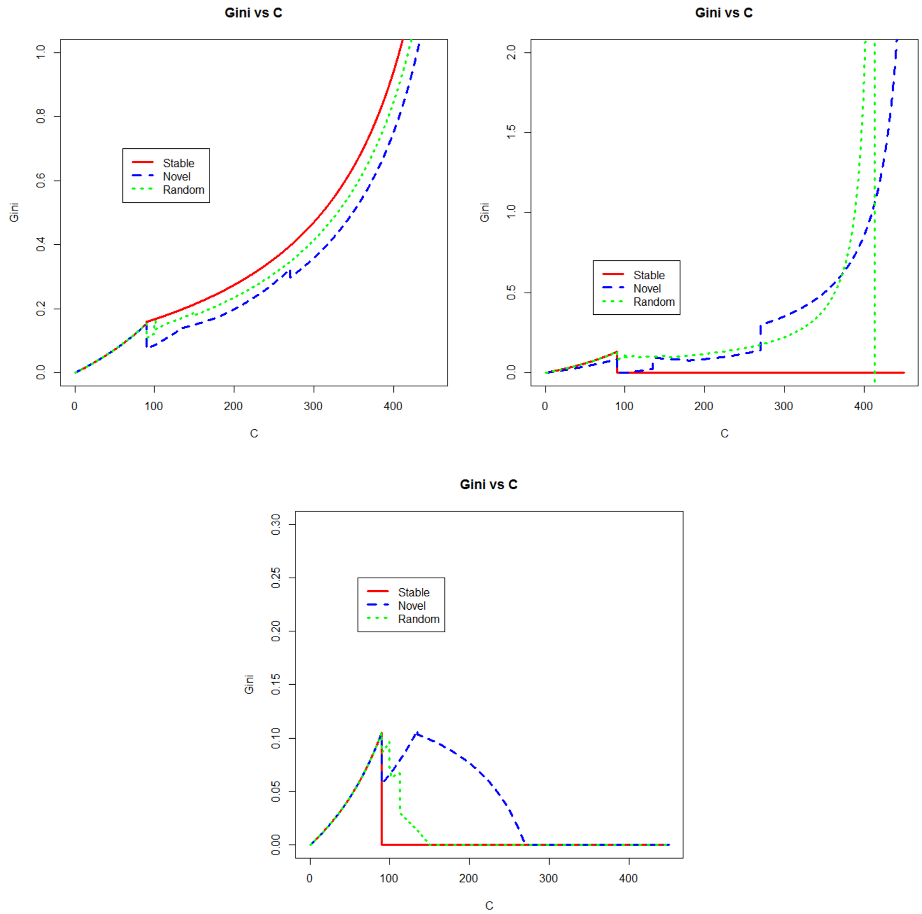


Figure 23: Graph of gini coefficient of expected outcome against C (Upper Left), Graph of gini coefficient of expected outcome against C, no Player 1 (Upper Right), Graph of gini coefficient of expected outcome against C, no Player 1 or Player 2 (Lower)

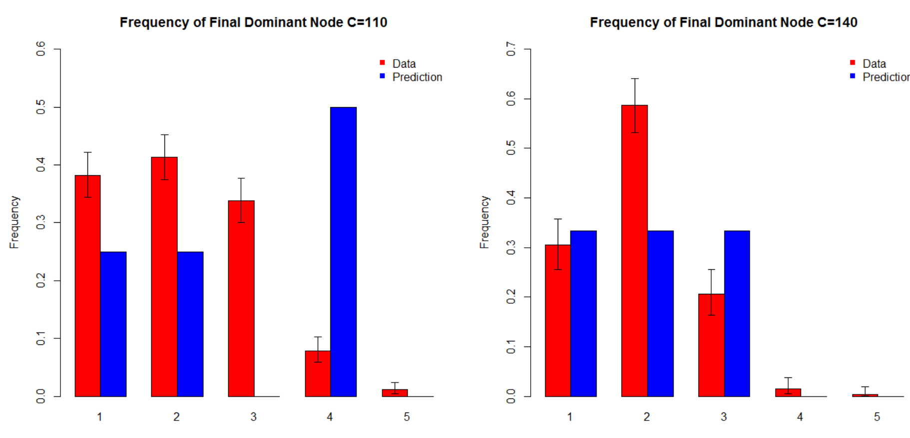


Figure 24: Histogram of the frequency for most connected node at the end of the network formation process $C = 110$ (Left) and $C = 140$ (right). Red bar show the observed proportion while blue bars show the predictions. Error bars show a 95% CI for the observed proportion.

	MPL 1	MPL 2	MPL 3
CRRA -log like	101.049	114.771	84.121
CARA -log like	115.120	207.944	157.588

Table 20: Negative log likelihoods for observed choice behavior in each MPL using individual level risk parameters estimated with CRRA and CARA utility functions.

	Mean Eta	SDev Eta	Mean Mu	SDev Mu
Panel 1	0.200	0	2	0
Panel 2	0.100	0.087	0.119	0.103
Panel 3	0.200	0	0.200	0
Overall	0.507	1.291	0.748	0.682

Table 21: Summary statistics on CARA model parameters

Note that, in addition to the deviations mentioned in the main body, we also see a strong bias towards Node 2 in the $C = 140$ game, and a much milder bias in the $C = 110$ game. It is unclear why this is the case, but it does not have a strong influence on incentives. Aside from the bias in favor of Node 2, the distribution of final dominant nodes seems to be shifted earlier.

9.2 CARA vs CRRA Risk Aversion Estimates

While there are no theoretical reasons to pick one functional form over another in this paper when dealing with consistent subjects, we can compare theories on how well they predict the actual data. We estimated risk preferences using both CRRA and CARA utility functions. The negative log likelihoods for fitting the MPL data are reported in Table20. The CRRA utility function fits the choice data better in all cases. It also produces more significant results in later analysis. Results using estimates from the CARA model are similar but slightly weaker. They are reported in Appendix Section 9.2.

Some readers may find it surprising that the CRRA utility function fits better than the CARA function given the role that wealth effects play in CRRA and the small size of the gambles relative to the presumed true wealth of the subjects. If one looks at the original paper of Holt and Laury (2002), however, one can see that the estimate coefficient corresponding to CRRA risk aversion was more significant than the estimate corresponding to CARA utility (15.8 standard errors from zero versus 1.1, [-values not reported). It may be that players are treating each gamble as a separate prospect, independent of wealth.

Still as a robustness check, we report the results of analysis as in Section 6 but instead of estimating the CRRA risk aversion function, we estimate the CARA risk aversion function

$$u_i(\pi) = \frac{1 - \exp -\eta\pi}{\eta}$$

using Holt and Laury (2002) action probabilities as before

Note that we are again using η as the risk aversion coefficient in order to economize on notation, although the quantity being measured is not the same. Table 21 summarizes the estimated values for η.

The values are very tightly clustered around 0.1, although there is variation. CARA does not work well, largely because it is very sensitive to people displaying undefined risk types, which is quite common in our data. As such μ is quite high, which may lead to a flat objective function and inaccurate estimates of η.

The CARA estimates do still have some predictive power. Tables 22 present regressions of vying behavior on player characteristics with the η_is estimated using the CARA utility function

Results for Player 4 facing one dominant node are similar to the CRRA case, but here the risk aversion is not significant. The coefficients on risk aversion are still negative, and sometimes close to

	Player 4 Facing One Dominant Node			Player 3		
MPL 1 η	-0.245^{*}	-0.168	-0.172	-0.053^{*}	-0.066	-0.073^{**}
	(0.059)	(0.301)	(0.243)	(0.077)	(0.107)	(0.045)
Intercept	1.057	0.246	0.258^{***}	1.237^{*}	0.672^{***}	0.700^{***}
	(0.455)	(0.113)	(0.001)	(0.074)	(0.000)	(0.000)
Expected Vie Gain	0.070	0.018		0.157	0.060	
	(0.740)	(0.925)		(0.267)	(0.691)	
Openness	-0.006			0.012		
	(0.498)			(0.261)		
Extroversion	0.004			0.0001		
	(0.817)			(0.994)		
Conscientiousness	0.0001			0.008		
	(0.994)			(0.385)		
Agreeableness	-0.022^{**}			-0.007		
	(0.034)			(0.427)		
Neuroticism	0.002			-0.023^{***}		
	(0.918)			(0.008)		
Adj R^2	0.027	0.003	0.013	0.147	0.055	0.057
Obs	94	94	94	250	250	250

Table 22: Predicting Vying for Dominance for Player 4 Facing One Dominant Node and Player $3C = 110$. Estimating η using a CARA utility function. The Player 4 regressions use ηs estimated from MPL 1, while the Player 3 regressions use η estimated from MPL 2. Errors clustered at the individual level. ($\leq 0.1^{*}$,$\leq 0.1^{**}$,$\leq 0.01^{***}$)

significant, but they never reach significance. The adjusted R^2s are substantially reduced.

Again results for Player 3 are similar to CRRA results. Risk aversion is significant in one specification only, but the coefficients are consistently negative.

9.3 Controlling for Undefined Risk Type

Table 23 repeats the analysis from 17 without the players who had undefined risk types on the MPL corresponding to each situation. While significance is impacted in some cases due to the large number of dropped subjects, the coefficients are very similar implying that the undefined subjects are not driving the results.

We can also address the question of whether undefined risk types are driving resuls more directly by regressing vying behavior against η, a dummy for undefined risk type in the corresponding MPL, and an interaction term. If the impact of η on vying for dominance is significantly different for players with undefined risk types, then the coefficient on the interaction term should be significant. Table 24 reports results of these regressions. In neither situation is the coefficient on the interaction term significant.

9.4 Impact of Risk Aversion When Vying is Not Optimal

9.5 Alternative Behavioral Models

There are a number of of common behavioral models that one might consider as alternatives to risk aversion for explaining the deviations from the base theory. In particular, many experiments find predictive success with random utility models like the QRE model of McKelvey and Palfrey (1998) and heterogeneous sophistication models like the Level-K model of Stahl and Wilson (1994). The risk aversion model does not provide a full stochastic model of player behavior like these models do, so it cannot be compared directly on AIC without imposing some error structure. There are, however, a few qualitative reasons to think that these models will be unlikely to provide a good fit for the data on their own.

Variable	Player 4 Facing One Dominant Node			Player 3		
MPL 1 η	-0.275	-0.289	-0.222	-0.089^{***}	-0.069	-0.048^{*}
	(0.268)	(0.224)	(0.298)	(0.000)	(0.120)	(0.059)
Intercept	1.205	0.402^{*}	0.276^{*}	3.142^{***}	0.842^{***}	0.7626^{***}
	(0.574)	(0.050)	(0.015)	(0.000)	(0.000)	(0.000)
Expected Vie Gain	-0.199	-0.191		-0.022	-0.150	
	(0.559)	(0.406)		(0.899)	(0.487)	
Openness	-0.008			-0.014		
	(0.722)			(0.146)		
Extroversion	-0.002			0.009		
	(0.898)			(0.161)		
Conscientiousness	0.011			-0.009		
	(0.460)			(0.425)		
Agreeableness	-0.017			-0.031^{***}		
	(0.347			(0.001)		
Neuroticism	-0.011			-0.031^{***}		
	(0.560)			(0.000)		.
Adj R^2	0.018	0.020	0.024	0.174	0.009	0.000
Obs	62	62	62	134	134	134

Table 23: **All players with undefined risk types on the corresponding MPLs removed.** Predicting Vying for Dominance for Player 4 Facing One Dominant Node and Player 3 $C = 110$. The Player 4 regressions use ηs estimated from MPL 1, while the Player 3 regressions use η estimated from MPL 2. Errors clustered at the individual level. ($\leq 0.1^{*}$,$\leq 0.1^{**}$,$\leq 0.01^{***}$)

Variable	Player 4 Facing One Dominant Node	Player 3
Corresponding η	-0.222	-0.048
	(0.014)	(0.056)
Intercept	0.276	0.762
	(0.296)	(0.000)
Undefined	0.113	-0.200
	(0.490)	(0.092)
η*Undefined	-0.115	0.006
	(0.648)	(0.921)
Adj R^2	0.050	0.084
Obs	94	250

Table 24: Predicting Vying for Dominance for Player 4 Facing One Dominant Node and Player 3 $C = 110$ as a test of whether having an undefined risk type influences the effect of risk aversion on vying for dominance. The Player 4 regressions use ηs estimated from MPL 1, while the Player 3 regressions use η estimated from MPL 2. Errors clustered at the individual level. ($\leq 0.1^{*}$,$\leq 0.1^{**}$,$\leq 0.01^{***}$)

Variable	Player 4 Facing Three Dominant Nodes		
MPL 1 η	-0.120	-0.100	-0.091
	(0.130)	(0.203)	(0.161)
Intercept	0.393	0.140	0.115**
	(0.183)	(0.260)	(0.025)
Expected Vie Gain	-0.043	-0.033	
	(0.756)	(0.814)	
Openness	-0.010**		
	(0.042)		
Extroversion	0.007		
	(0.131)		
Conscientiousness	-0.002		
	(0.677)		
Agreeableness	-0.001		
	(0.794)		
Neuroticism	0.003		
	(0.429)		
Adj R^2	0.042	0.008	0.014
Obs	156	156	156

Table 25: Predicting Vying for Dominance for Player 4 Facing Three Domiant Nodes and Player 3 $C = 110$ using ηs estimated from MPL 1. Errors clustered at the individual level. ($\leq 0.1^{*}, \leq 0.1^{**}, \leq 0.01^{***}$)

One of the moments of the data that we have devoted the most attention to is the proportion of Player 4s in the $C = 110$ treatment facing one dominant node who vie for dominance by making three connections. More players choose the myopic move in this case than vie for dominance. QRE and other similar random utility models[45] always predict that, in equilibrium, a player in a given state has actions a_1 and a_2 such that $E(u_i|a_1) \geq E(u_i|a_2)$ it must be the case that action a_1 is chosen weakly more often than action a_2.

Because Player 4s in the $C = 110$ treatment facing one dominant node make more points on average by vying than playing the myopic move, they should vie at least as often as they play myopic the myopic move.[46] Instead they vie significantly less. As such, the data does not seem to be consistent with models of this type.

Level-K also fails to predict several key moments of the data, and similar models are also unlikely to be more useful. Any player who is savvy enough to vie in position 3 should be savvy enough to vie when conditions are right in position 4, since the backwards induction reasoning is easier when there are fewer future players to consider.

For example, if level-0 is random behavior, then level-1 players would always play myopic moves.[47] We can use data on Player 5 in the $C = 110$ treatment to estimate the number of level-0 and level-1 or higher players. Looking at Player 5, we see that approximately 28% of players are level-0 when we account for the probability that a level-0 player might play a correct move by chance. Given that estimate, it is actually not optimal for Player 4 to vie when facing one dominant node, which fails to match payoff data, in which vying is better for Player 4 in that condition. If taking the myopic action is optimal for Player 4 facing one dominant node when $C = 110$, then the behavior of Player 4 in that case gives us an estimated 54% level-0 players, which is significantly different from the estimate we got looking at Player 5.[48]

If we instead assume vying for dominance is optimal for player 4, accounting for accidental correct moves by level-0 players, then from Player 4's data we estimate that the proportion of level-2 or higher

[45] those in the style of Block and Marschak (1960) RUM's in which the perceptual errors on utilities are drawn independent of the choices and utilities.

[46] Note that in this condition there is only one dominant node, so there is only one myopic move.

[47] See Appendix Section 8.2

[48] P-value of 0.000 given a standard test of proportions.

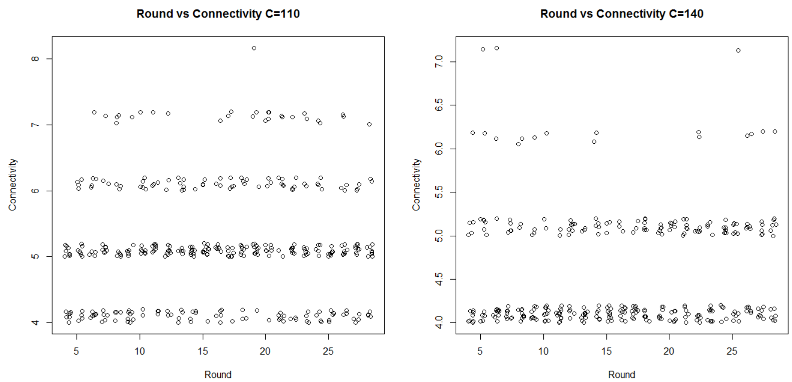


Figure 25: Connectivity plotted against round number. Data is perturbed to allow readers to see density.

players is approximately 11%. If we extrapolate from that, we should only see 22% of players vying for dominance as Player 3 in the $C = 110$ treatment. Instead we see 60% of Player 3s vying for dominance, which is significantly different from the prediction.[49] Models of heterogeneous sophistication fail to account for the data as well as models of heterogeneous risk aversion.

Learning and experimentation models also offer potential candidates for explaining behavioral deviations from theory, but there is little evidence of learning in the experiment. See Appendix Section 9.6 for data and learning.

9.6 Learning Effects

The last major potential determinant of outcomes that we look at is learning effects. Do players become more or less likely to take vie for dominance as the game progresses? Do they substantively change their behavior with experience? The answer to both questions seems to be no. Figure 25 show scatter plots of the round number against the connectivity of the network formed. Noise is added to the points to allow visual representation of density.

In both treatments there is no discernible effect of round experience on the connectivity of networks formed. There is a slight suggestion of a U-shaped effect in the $C = 140$ treatment, but this is primarily caused by a few outliers with little change in the average network. To check more formally, we regress the connectivity of networks on round number and find no significant effects. Results are reported in Tables 26 and 27.

Other summary values display a similar lack of pattern.

10 Appendix C: Batteries

In this appendix we provide screenshots of the question batteries used in some of the sessions. Figure 26 reproduces the screenshots from the belief elicitation question battery. We do not reproduce the Big Five Inventory here, because the questions were taken directly from John and Srivastava (1999)

The first panel is designed to elicit the risk preferences of players. In this panel, players make binary choices between gambles in a multiple price list in the manner of Holt and Laury (2002). Players are presented with three sets of ten choices each. The gamble on the left hand side of the screen (gamble A) is fixed while the gamble on the right hand side (gamble B) improves going down the page. At the top of the screen gamble A second order stochastically dominates gamble B while at the bottom of the screen the reverse is true. Players are paid the outcome of one gamble from among all risk elicitation questions chosen at random.

[49] P-value of 0.000 given a standard test of proportions.

	Dependent variable:
	connectivity
rep	0.009
	(0.006)
Constant	4.944***
	(0.110)
Observations	375
R^2	0.005
Adjusted R^2	0.003
Residual Std. Error	0.873 (df = 373)
F Statistic	2.059 (df = 1; 373)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 26: Regression of Network Connectivity on Round Number, $C = 110$

	Dependent variable:
	connectivity
rep	0.006
	(0.005)
Constant	4.374***
	(0.085)
Observations	325
R^2	0.005
Adjusted R^2	0.002
Residual Std. Error	0.630 (df = 323)
F Statistic	1.641 (df = 1; 323)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 27: Regression of Network Connectivity on Round Number, $C = 140$

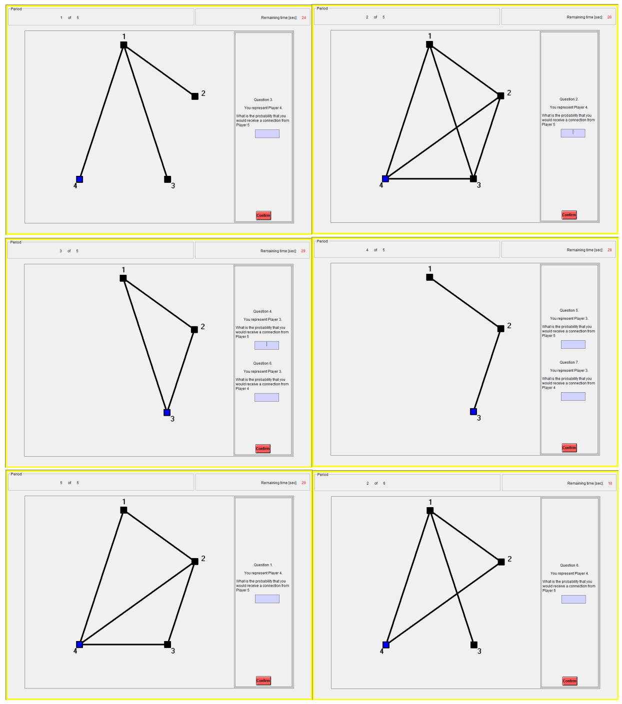
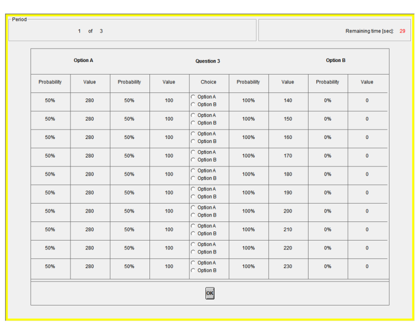


Figure 26: Screenshots from the belief elicitation question battery



Period

1 of 3

Remaining time [sec]: 29

Option A				Question 3	Option B			
Probability	Value	Probability	Value	Choice	Probability	Value	Probability	Value
50%	280	50%	100	Option A Option B	100%	140	0%	0
50%	280	50%	100	Option A Option B	100%	150	0%	0
50%	280	50%	100	Option A Option B	100%	160	0%	0
50%	280	50%	100	Option A Option B	100%	170	0%	0
50%	280	50%	100	Option A Option B	100%	180	0%	0
50%	280	50%	100	Option A Option B	100%	190	0%	0
50%	280	50%	100	Option A Option B	100%	200	0%	0
50%	280	50%	100	Option A Option B	100%	210	0%	0
50%	280	50%	100	Option A Option B	100%	220	0%	0
50%	280	50%	100	Option A Option B	100%	230	0%	0

OK

Figure 27: Screenshots of MPL 3

The gambles are adapted to be similar to those faced in the equilibrium with heterogeneous risk aversion and uniform random tie-breaking of Section6.4. MPLs 1 and 3 were designed to mimic the trade-off of Player 4 facing one dominant node in the low cost treatment and Player 3 in the high cost treatment. The left hand choice in both gambles has the same distribution of payoffs as vying for dominance in the corresponding positions. The right hand side includes varying constant payoffs with the expected payoff from a myopic move in the corresponding situation included on one row. See Figure 27 for a screenshot of MPL 3

MPL 2 mimics the trade-off faced by player 3 in the low cost treatment. The left hand gamble again has the same distribution of payoffs as vying for dominance in the corresponding situation. The right hand gambles all have the same distribution of payoffs as a myopic move for dominance in an equilibrium where Player 4 vies for dominance when facing one dominant node with a frequency of α. The frequency α is changes over the different questions in the list.

The second panel elicits the beliefs players hold about the actions of actions of others. Players are put into the position of Player 3 or Player 4 and given a hypothetical network. Players are then asked to estimate the probability that later players will connect their node given the network structure provided. These guesses were then compared to the actual observed sample probabilities from the first six sessions. Players are rewarded for one question from the belief elicitation panel chosen at random. Points were based on a quadratic probability scoring rule. A player who made a guess of $\tilde{p}$ for the awarded question with a real observed probability of $\bar{p}$ would receive an 800 point ($4) prize with a probability of $1 - (\tilde{p} - \bar{p})^2$. This payment rule should encourage subjects to truthfully reveal their beliefs about average probability of receiving a connection under the assumption of expected utility.[50]

The last panel elicits personality traits of players. We administer a battery of questions taken from the Big Five Inventory (BFI) of John and Srivastava (1999). The big five inventory is a common personality assessment method which rates people on five different personality characteristics: Extroversion, Agreeableness, Neuroticism, Conscientiousness, and Openness. Extroversion is a person's tendency to enjoy and be energized by social interaction. Agreeableness is the tendency to care about and wish to please other people. Neuroticism is a measure of emotional instability. Conscientiousness is a measure of how goal-oriented and organized one is. Openness is a tendency to be creative and try new things. There is no performance incentive for this section. Instead players are awarded a prize of 800 points ($4) for completing this section.

The BFI was chosen for two reasons. First, the BFI is fairly common in the academic literature on personality. Second, the BFI has been linked to entrepreneurial tendency, which is relevant to answering the question of why some players vie for dominance and others do not. Vying for dominance can be thought of as an entrepreneurial activity. Players make an investment in hopes of some uncertain future payoff. Zhao and Seibert (2006) report that entrepreneurs tend to be less neurotic, more extroverted, more open, less agreeable, and more conscientious than other managerial types. We will examine whether any of these characteristics are linked to the tendency to vie for dominance.

[50] Schotter and Trevino (2014)