Vying for Dominance: An Experiment in Dynamic Network Formation

Nathaniel Neligh*

Sep 4, 2018

Abstract

Centrality can provide many advantages. 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 predicts that agents entering the network at specific times will make more connections than is myopically optimal in hopes of later receiving additional connections. In a laboratory experiment, we find that players do exhibit this "vying for dominance" behavior, but do not always do so at the predicted critical times. A model of heterogeneous risk aversion best fits the observed deviations from predictions.

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 how their structural features develop. One common structural feature among networks is that they often have a few very central nodes (i.e. nodes that are "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]—as such, it is natural to ask: why are some nodes are more central than others?

In this paper, we explore the previously unexamined hypothesis that *timing of entry into a network* plays a critical role in determining which nodes become the most central. We interrogate this hypothesis by proposing a new dynamic model of network formation and testing it in a laboratory experiment. Previous models have suggested that fundamental differences in node traits and equilibrium selection determine a node's centrality, but our model suggests that when nodes enter a network can also have a large impact.[3] Network games, like many economic systems, may generate advantages for those moving at certain times. For example, it is common wisdom in the technology industry that startup timing, when a new firm joins the market, is critical to eventual success. In general, it is neither the first firms to enter an industry nor the last that become the most successful but those joining somewhere between the two extremes.[4]

*Economic Science Institute, Chapman University: neligh@chapman.edu. We are grateful to Mark Dean and Alessandra Casella for their guidance and and advice on this project. Thanks also to Navin Kartik, Pietro Ortoleva, Judd Kessler, Corinne Low, Yeon-Koo Che, Andrea Prat, Miachael Woodford, Jacob Leshno, and all the participants in colloquia at Columbia University who have provided feedback and advice. We acknowledge the financial support of grants from the Columbia Experimental Laboratory for the Social Sciences. The author of this paper has no financial interests related to the outcome of this research. Research was conducted with approval from the Columbia University IRB.

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

[2]For theoretical evidence see: Kranton and Minehart (2001); Blume et al. (2009); Apt et al. (2016); Chen and Teng (2016)

For empirical evidence see: Powell et al. (1996); Sarigöl et al. (2014);Rossi et al. (2015); Pollack et al. (2015)

[3]For a discussion of the existing models, see Section 1.1.

[4]See Lilien and Yoon (1990).

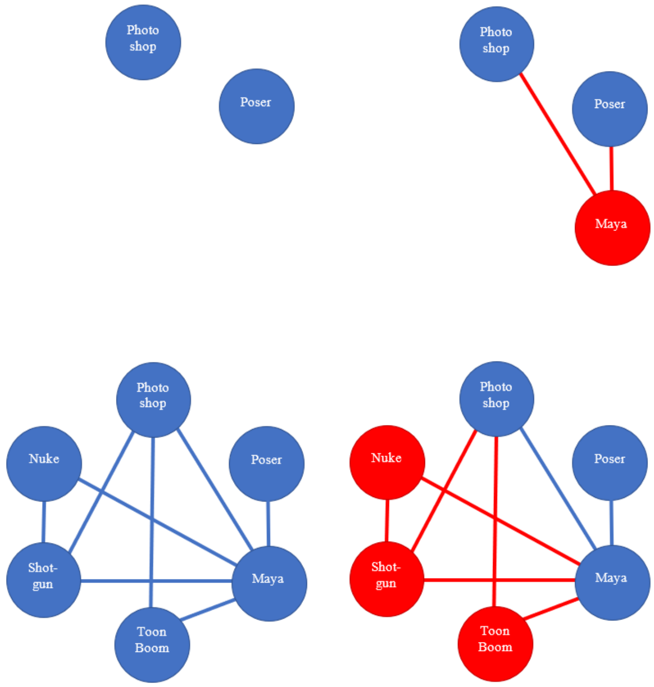

Figure 1: The formation of a network of compatibilities in animation software starting at the top left and going clock-wise.

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. Compatibilities are absolutely critical to the success of animation softwares, because in any project of reasonable size, multiple pieces of software will inevitably come into play. As such, the average consumer will generally consider compatibilities as a major factor when making purchasing decisions. Note that the motivations of the software creators in the example have been abstracted and substantially simplified for the sake of clarity.

At the start, the network contained Photoshop and Poser, unconnected. Later, Maya entered the network and connected to (i.e. made compatibilities with) both of the existing pieces of software. After that, three additional animation softwares (Shotgun, Toon Boom, and Nuke) joined the network, and they all connected to Maya as they did. By 2017, Maya had become the dominant (most central) firm in the animation software market, with more compatibilities than any other software in the network, beating out both Photoshop and Poser.[5]

Why do we see this pattern of play? One possibility 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 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.

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 decisions taking into account their expectation of future moves as well as the results of previous actions. According to our

[5]Note that other software (in particular Photoshop) may be dominant in other networks. We are restricting our attention to the animation software network in this example.

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 solutions depend only on static features of the network.

The model we introduce and test in this paper reflects a highly simplified and generic version of the true mechanisms behind real world network growth. As such, the experiment can be thought of as a way of establishing an upper bound on the rationality of average individuals in network formation games as well as helping us to identify the behavioral factors that are most relevant to decision making in a network growth context.

In Section 2 we introduce the model, and then in Section 3 we discuss the basic features and results of that model. We find that, 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 types of move which contribute to the strategic richness arise in many equilibria: *myopic* moves and *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.

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

Due to strategic complexity, solving the game can be difficult for large networks, so we focus primarily on small networks. Studying smaller networks is also more practical, as such environments are easier to implement in the laboratory. Surprisingly, if the network is small enough, then the only equilibrium moves are playing myopically or vying for dominance, making the solution easier to find and interpret. Note that these behaviors are not only relevant for the study of smaller networks. In Neligh (2017) we also consider a restricted version of the game where players will either play myopically or vie for dominance even in arbitrarily large networks.

In Section 3.5, we solve the game with five players, the setting used in the lab. Player 5's move is simple; he should choose a myopic move. The interesting behavior is that of the intermediate players, 3 and 4. These players may choose to vie for dominance in order to potentially gain connections from later players. These players' incentives depend on their expectation of later nodes' behavior and the number of competing central nodes they face when they enters the network. The trade-off between vying and myopic play also naturally depends on the cost of connections, leading to four parameter regions of interest.

Player 4 can only ever profitably vie for dominance in the two lower cost regions. Network state must also be considered, however: in the costlier of the two previously mentioned low cost regions, 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 Player 3's choice to vie for dominance.

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 two different predicted patterns of play. The solution predicts 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 them to the predictions of the model. In general, most players either vied for dominance or chose myopic moves, as predicted. However, players did not always vie for dominance at the predicted critical times. Broadly speaking, the model does better in predicting the play for later movers, as one would expect. The prediction that Player 5 always plays myopically is well supported. For Player 4, the comparative statics are correct, but the levels were sometimes wrong. Player 4 is predicted to vie when costs are low and when Player 3 did not vie, and he did vie most often under those conditions. However, Player 4s only vied 22% of the time in that scenario, a number far different from the 100% predicted by the theory. For Player 3 even the comparative static was wrong: contrary to predictions, Player 3s vied more when costs were low.

Overall, players vied for dominance less often than expected, and 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 risk averse players may avoid vying for dominance. Vying for dominance is a risky, investment-like, behavior, so risk aversion is a natural explanation for the observed deviations. In Section 6 we explore the possibility that risk aversion is influencing subjects' choices. We elicit risk preferences using multiple price lists 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, and we find that it fits well with the aggregate moments in the data.

We conclude then that timing does play a strong role in whether nodes have the opportunity to become central. However, risk aversion can prevent some players from exploiting their opportunities. It takes the right type of player joining the network at the right time to achieve high centrality.

1.1 Literature: Theory and Experiments on Network Formation

Before we present the model, we provide an overview of the literature concerning theories and experiments on networks and network formation. In many previous models of network formation the set of solution networks can be defined based on static features of the network such as efficiency, individual rationality, or stability. These features are static in the sense that they can be assessed with from a single instance of a network with no need to consider the dynamics of the network formation process. Formulations in which solutions depend on static features of the network are appealing because they allow for clean results and good tractability, but they do eliminate a number of dynamic questions from consideration. In models of this type, structural features such as centrality are determined by equilibrium selection and fundamental properties of the nodes. As such, investigating the relationship between entry timing and node centrality will require a different type of model.

Network formation models were introduced to modern economics by Jackson and Wolinsky (1996) and their model of cooperative network formation. In this model, a network is "pairwise 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. Pairwise stability is a static concept, so dynamics do not play a role in determining network structure.

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 with solutions being referred to as "Nash Networks". Bala and Goyal (2000) also introduce dynamics into their models, although players are assumed to behave myopically, so the process will always converge to a "Nash Network". Therfroe, the set of possible final networks again depends only on static features. In a similar vein, the model of Watts (2001) assumes that players myopically update their connections, which causes them to converge to a pairwise stable network.

There are several papers that do include dynamics and forward looking strategic agents but payoffs or game structures are such that the set of possible outcomes again depends only on some static network feature. For example Currarini and Morelli (2000) and Mutuswami and Winter (2002) both employ mechanisms which guarantee that only efficient networks can be supported in equilibria of their game. Song and van der Schaar (2015) find that their network formation process can converge to any network which satisfies an individual rationality constraint requiring that each player make a payoff of at least zero.

The natural sciences related literature explores models of network formation which do not have solution sets that can easily be defined using static features of the model, but these models often lack some component of the strategic complexity required to generate some of the interesting behaviors we are interested in, such as vying for dominance. 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. The model of Kim and Jo (2009) does include optimizing agents, but connections only provide immediate benefit, so agents have no need to be forward looking.

While agents are strategic and forward looking in the 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. Like our model, theirs 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 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, a type of action which is central to the present paper.

Experiments have found that network structure can have large impacts on behavior in trading games,[6] public goods games,[7] and group decision making games.[8] A good review of older experiments examining the role of network structure in determining economic outcomes can be found in Kosfield (2003).

Experiments have also 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 network formation games. As such, there are often several players all attempting to become central in these experiments. Second, is the role of heterogeneity. Players' differences, whether inherent or exogenously given, have a large impact on behavior.

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, particularly in players' risk aversion, can influence the competition for centrality in our experiment.

Competition for centrality is a very important feature of our model as well. However, while this competition for centrality has been a confounding factor in previous studies, it is a direct prediction in our model. As such, our model will allow us to discuss the phenomenon with more rigor and detail.

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 experiment 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 are incetntivized to 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 our model and to examine the importance of entry timing in determining node centrality.

[6] Charness et al. (2007)
[7] Carpenter et al. (2012); Charness et al. (2014)
[8] McCubbins and Weller (2012); Kittel and Luhan (2013)
[9] Bernasconi and Galizzi (2005); Kearns et al. (2012); Carrillo and Gaduh (2012); van Leeuwen et al. (2013)

2 The Game

We now present our 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.

There is a set of players represented by nodes indexed $j \in \{1, ..., J\}$. Networks are represented as $G = \{\mathbf{n}(G); \mathbf{h}(G)\}$ where $\mathbf{n}(G)$ is a set of nodes, and $\mathbf{h}(G)$ is a set of edges represented by pairs of nodes. Connections are undirected. 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 he 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. Note that this definition of strategy is not restrictive, because there is a one-to-one mapping between network states and action histories.

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.

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

$Y \in \mathrm{I\!R}$ is a constant base payoff.[10] $C|\mathbf{h}_i|$ is the total cost of connections by individual i who purchased the set of connections $\mathbf{h}_i$. $C \in \mathrm{I\!R}^+$ is the constant cost of connections. $B\zeta_i(G_J, \delta)$ is the benefit from centrality where $B \in \mathrm{I\!R}^+$ 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.[11] 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, allowing us to better isolate the impact of the novel dynamic game structure. The payoff function is similar to that used by Watts (2001) and Jackson and Wolinsky (1996).[12] This type of payoff is most directly applicable to systems in which some beneficial opportunity 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 there are benefits to centrality, as centrality measures are often highly correlated, especially in networks with low diameter.[13]

This model makes some fairly restrictive assumptions, but many of the interesting results carry through in some form even when the assumptions are loosened. We discuss several extensions Section 7, and additional extensions, particularly to larger networks, in Neligh (2017).

[10] This component is not important in discussing the major results of the model, but it will become relevant later on in the discussion of the experimental game and risk aversion.

[11] We sometimes refer to $d_{ij}(G_J)$ as d_{ij} when it will not cause confusion, because no other distances are payoff relevant.

[12] Their payoff function is has $Y = 0$ and $B = \delta$, but otherwise is identical. The difference only becomes important when we begin to discuss the experiment and risk aversion.

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

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.[14] Because this is a finite game of perfect information, multiplicity of equilibria derives from the manner in which players resolve indifferences. As such, in order to pin down an equilibrium we need a set of parameters and a tie-breaking rule.

3 Results

3.1 Efficiency

In economic models of network formation, theorists are generally most concerned with which networks form and which networks are efficient. We first explore the structure of the efficient network, which depends on the parameters of the problem and in particular on 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 the game.

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 proof, see Appendix A.1.

This result effectively identical to the efficiency result of Jackson and Wolinsky (1996) up to a factor of two that reflects the difference in the social cost of connections betwee the two models. In our model the social cost of a connection is C, paid by one party, while in Jackson and Wolinsky (1996), the cost is $2C$, because each involved party pays C.[15]

Note that, while the sequential nature of the game does impose limits on the set of feasible networks, it does not impose limits on the possible network "shapes" other than connectedness.

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 can not be formed by any SPE.

[14] Note that the indifferences in this game are due to structural symmetries 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] Also, the empty network is never efficient in our model, because it is never feasible.

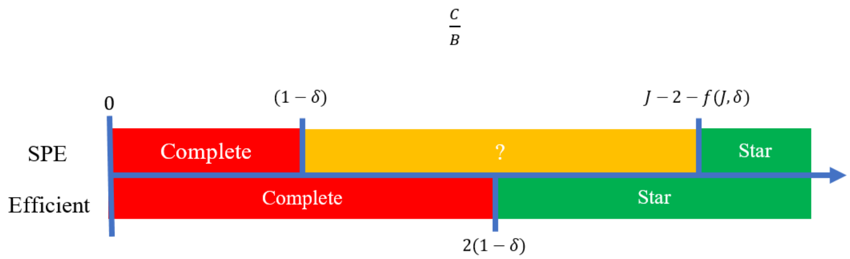

Figure 2: Visualization of parameter regions of interest. These visualized threshold locations imply low δ or high J. If the reverse is true, the second SPE threshold can be to the left of the efficiency threshold. $f(J,\delta) = \frac{\delta-\delta^{J-3}}{1-\delta}$.

For proof, see Appendix A.2.

As in Jackson and Wolinsky (1996), when $\frac{C}{B} < 1-\delta$ the complete network forms as the only outcome of subgame perfect equilibria, and when $\frac{C}{B} > 1-\delta$ it cannot form.

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. As such, we cannot guarantee a minimally connected network unless costs are relatively very high.

Proposition 3: 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 form any in SPE's of the game.

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 3 is tight as long as δ is small in a weak sense. If $\frac{C}{B} < J-2$ and δ is sufficiently small then there is a tie-breaking rule such that the subgame perfect equilibrium defined by that tie-breaking rule, C,B, and δ produces some network that is not a star network with non-zero probability. For proof of Proposition 3 and it's tightness when δ is small see Appendix A.2.

3.3 Summary of Results

These results are broadly intuitive and are generally quite robust to small changes in the assumptions of the model. They 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 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 results differ from those found in Jackson and Wolinsky (1996) and Watts (2001) in a number of key ways. First, the fact that we are using non-cooperative network formation shifts the efficiency threshold which allows for the possibility of inefficient under-connection. Inefficient under-connection arises when $(1-\delta) < \frac{C}{B} < 2(1-\delta)$, and inefficient over-connection can arise when $(J-1) - \frac{1-\delta^{J-3}}{1-\delta} > \frac{C}{B} > 2(1-\delta)$. Also, the empty network is not feasible, so it neither forms nor is efficient.

Second, the dynamics of the network formation process substantially change the range in which non-degenerate networks can form and the nature of those network. Non-degenerate networks can be supported for potentially much higher cost levels in our model, because additional connections can provide potentially large future benefits. The non-degenerate networks also tend to be less symmetrical, since order of entry generates inherent asymmetries. We will discuss the structure of non-degenerate networks in more detail in Section 3.5.

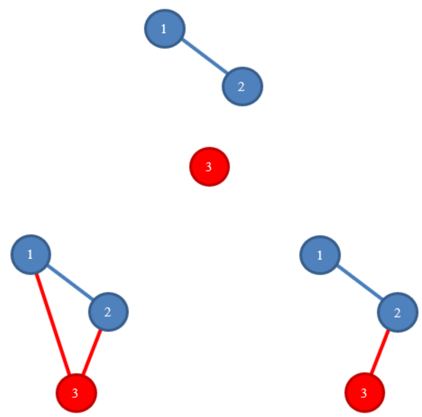

Figure 3: An unattached Player 3 (top) may choose to make a myopic move (lower right) or vie for dominance (lower left).

3.4 Myopic Moves and Vying for Dominance.

The question naturally arises of what happens in the non-degenerate region (yellow) region. Equilibrium behavior in this region is rich and can include a lot of strategically interesting actions including myopic moves and vying for dominance. As discussed in the introduction, a *myopic* move is a move which would be optimal if the game ended immediately after that move. *Vying for Dominance* means making a move which results in the player being of the most central (dominant) nodes immediately after.

In all of the cases we discuss in this paper, a myopic move involves making a single connection to one of the 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, they are simply behaving in a way that is consistent with myopia. As we will show, the myopic action is often optimal in the subgame perfect sense.

In general, even more complex types of behavior exist in the non-degenerate 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 myopically 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 backwards induction would require looking at the payoffs associated with $\prod_i = 2^J 2^{j-1} - 1$ (more than 2^{J^2-3J}) possible networks.

In order to make things solvable we focus on an environment in which vying and myopia are the only equilibrium moves: small networks.

Proposition 4: Vying for dominance and playing myopically are the only types of behavior which occur in equilibrium for networks with $J = 5$ and $C/B > 1$.

For proof see Appendix A.3. Note that Proposition 4 is actually quite robust to changes in the assumptions of the model. A result in Appendix A.3 shows how the proposition can be robust to the inclusion of risk preferences and behavioral noise. Section 7 also discusses ways that Proposition 4 can be robust to changes in the choice structure of the game.

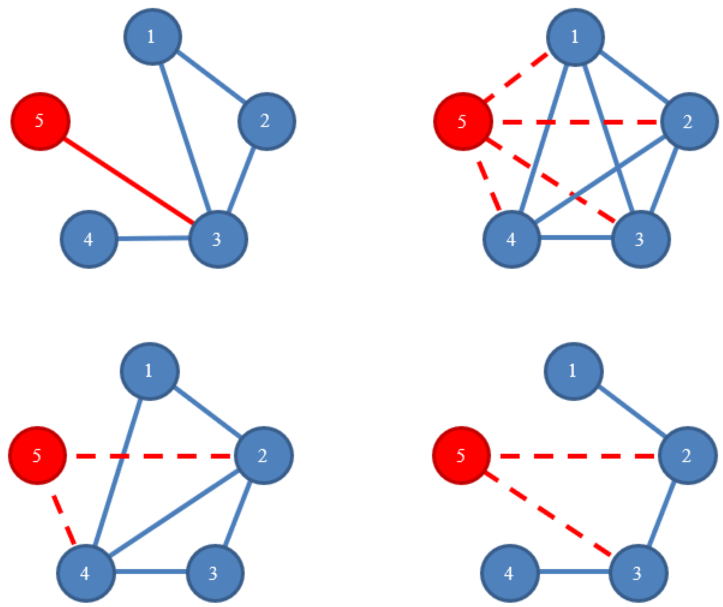

Figure 4: Several examples of possible move by Player 5. In cases with dotted lines, Player 5 picks one connection at random.

3.5 The Five-Node Game

In this section we will solve the game with five nodes and $C/B > 1$.[16] Recall that solutions to the game are characterized by the parameters and the tie-breaking rule. We use the following:

Definition: The *uniform random tie-breaking rule* states that when a player is indifferent between any number of moves during the backwards induction process, he will choose choose between them randomly with each optimal move having equal probability.

This tie-breaking rule is chosen for two reasons. First, it doesn't ex-ante favor any node. This is important, because we are trying to investigate the relationship between node centrality and entry timing. A tie-breaking approach that favors one entry time over another may artificially influence this relationship. Since nodes are identified by their entry timing in this game, that means we want a tie breaking approach which favors all nodes equally. Second, uniform random tie-breaking matches well with experimental data as we show in Appendix B.2. We discuss other tie breaking approaches in Appendix A.7.

We now solve the game using backwards induction. Note that we will be taking advantage of Proposition 4 to simplify the discussion of the game.

3.5.1 Player 5

Regardless of network configuration, Player 5 will always connect to a single dominant node, i.e. the myopic move. This is the best one connection move, and the benefits of multiple connection moves are 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.

[16] The game is solvable when $C/B < 1$, but the behaviors are very different from those in other regions, so that case is covered in the Appendix A.4.

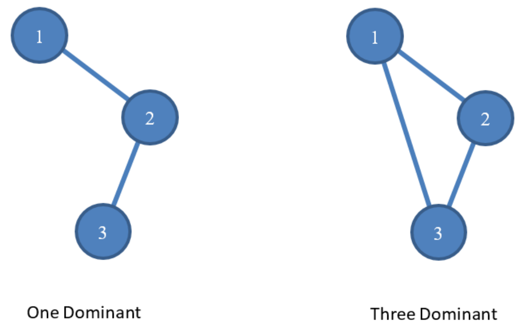

Figure 5: Possible networks faced by Player 4

Move	Cost	Immediate Benefit	Prob P5 Connection	Expected Future 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	$0.25B + 0.5\delta B + 0.25\delta^2 B$	$2.25B + 1.5\delta B + 0.25\delta^2 B - 2C$

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.

3.5.2 Player 4

Player 4's move can depend on the type of network he is facing. Ignoring the redundant case,[17] he can face two possible networks. 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 such networks. We consider the two cases created by the two networks separately.

Player 4 Facing One Dominant Node: From Proposition 4, we know that Player 4 will either play myopically or vie for dominance. There are therefore three possible moves he can make.

When Player 4 chooses a myopic action, he connects to a single dominant node. His single connection has a cost of C, and he receives 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 dominant node after the myopic action, there is no chance that player 5 will connect to them. Player 4 will then only expect to gain δB from a second degree connection with Player 5.

Player 4 can also vie for dominance by connecting to all existing nodes. Making three 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 (otherwise Player 5 will connect a distance of two away). Therefore, he will have an expected benefit of $0.5B + 0.5\delta B$ from his interaction with Player 5.

Finally, Player 4 can vie for dominance by making two connections, one to each non-dominant node. This move is never optimal.[18]

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

As such, Player 4will always play myopically or make the three connection vying move. We can see which is preferred by comparing the expected utilities and arriving at the following lemma.

[17]The two networks in which Player 3 connects to a single node are identical up to relabeling Nodes 1 and 2

[18]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.

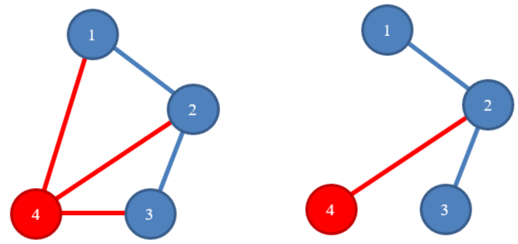

Figure 6: Visual representations of the potentially optimal moves of Player 4 facing one dominant node: vying for dominance (left) and myopic (right).

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

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

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 for Dominance)
- If $\frac{C}{B(1-\delta)} > 1.25$, Player 4 facing one dominant node will connect to that node (Myopic Action)

See Figure 6 for visualizations of the optimal moves of Player 4 facing one dominant node.

Player 4 Facing Three Dominant Nodes: In this situation, Player 4 only has two real moves. By proposition 4, he can make one connection (myopic) 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 if Player 4 vies for dominance, he only gets a one quarter chance of receiving a connection from Player 5, because there will be four dominant nodes. Vying and myopic moves are again the only optimal moves with vying being optimal up to some cost threshold, but the threshold has decreased, because the payoffs from vying for dominance have decreased.

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

- 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 7 for visualizations of the optimal moves of Player 4 facing three dominant nodes.

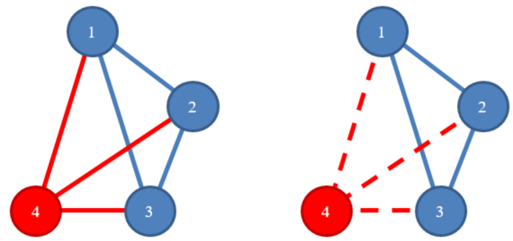

Figure 7: Visual representations of the possible moves of Player 4 facing three dominant nodes

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 redundant case). He can make one connection (myopic) or two connections (vying for dominance). See Figure 8 for a visualization of the moves Player 3 can make.

Player 3 does have to consider, however, how his own choice effects Player 4's choice to vie for dominance. Table 3 presents the choices and trade-offs faced by Player 3.

Generally speaking, Player 3 wants to vie for dominance when the cost of connections is relatively low for the same reason as Player 4 does, potentially receiving connections from Players 4 and 5. Player 3 is generally more willing to vie, because he has a lower cost for doing so. 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 for Player 3. In this range, vying is not profitable for Player 4, so he plays myopically.

However, 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 whether Player 3 has vied for dominance. Player 3 wants Player 4 to vie, because Player 4 vying provides a connection to Player 3 with certainty. As such, Player 3 will play myopically in this range in order to get Player 4 to vie.

When $\frac{C}{B(1-\delta)} \in (\frac{1}{1-\delta}, 1.125)$, 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 $\frac{C}{B} > 1$

- $\frac{C}{B(1-\delta)} \in (\frac{1}{1-\delta}, 1.125)$, Player 3 vies for dominance
- $\frac{C}{B(1-\delta)} > (1.125, 1.25)$, Player 3 makes a myopic move
- $\frac{C}{B(1-\delta)} \in (1.25, \frac{5}{3})$, Player 3 vies for dominance
- $\frac{C}{B(1-\delta)} > \frac{5}{3}$, Player 3 makes a myopic move

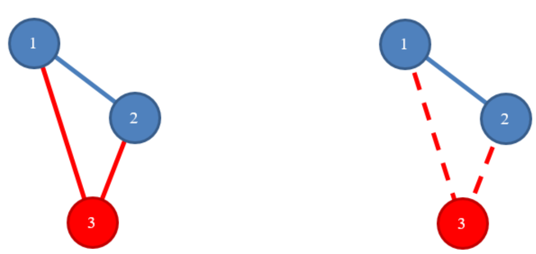

Figure 8: Possible moves for Player 3. Vying (left) and myopic(right). In the case with the dotted line, one connection is picked at random.

$\frac{C}{B(1-\delta)}$	P4 Vie Given P3 Vie	P4 Vie Given P3 Myopic	Vie Expected Connections	Myopic Expected Connections	Pref
$(\frac{1}{1-\delta}, 1.125)$	Yes	Yes	1.25	1	Vie
$(1.125, 1.25)$	No	Yes	2/3	1	Myopic
$(1.25, 1.\bar{6})$	No	No	2/3	0	Vie
$(1.\bar{6}, \infty)$	No	No	2/3	0	Myopic

Table 3: Player 3 trade-offs

Lemmas 1, 2, and 3 give us four parameter regions of interest. Players 1 and 2 have no decisions to make.

3.5.4 Summary

We can compile the behaviors into solutions for eachof the parameter regions discussed.

Proposition 5: The solution to the five-node game with $C/B > 1$ can be characterized by the following table:

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

For visualizations of what typical networks in each region can look like see Figure 9.

4 The Experiment

4.1 The Experimental Game

We conducted an experiment in order to test whether players are able find the previously described subgame perfect equilibrium of the game and to characterize any systematic deviations from the equilibrium behaviors. In the experiment, players play the game with a slightly modified payoff function which simplifies 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 distance two node. Nodes which are farther than two away provide no benefit. We used 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) and (3) discussed above with typical networks like those represented in Figure 9.

4.2 Setup

We conducted all sessions at the Columbia Experimental Laboratory for the Social Sciences (CELSS). The experiment was implemented using the zTree experimental platform,[19] and participants were recruited from the CELSS subject pool which is managed using the Online Recruitment System for Economic Experiments (ORSEE).[20]

Data was gathered during eight sessions from a total of 114 subject. The table in Appendix B.8 summarizes treatment and demographic information for each session. Note that the question batteries were only included with $C = 110$ treatments, as we are primarily using data form the batteries to explain vying for dominance, and this behavior occurs more often and in more situations in the low cost treatments.

In thie session, each round corresponded to the creation of one network for each group. Players were grouped in sets of three representing Nodes 3 through 5[21] and were randomly rematched and reordered for each round. Players were assigned to groups in a completely random fashion with no regard to previous groupings, and within each groups players were randomly assigned a role, again completely randomly. This was done to minimize potential repeated game effects. 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.

[19] Fischbacher (2007)
[20] Greiner (2015)
[21] Nodes 1 and 2 have no choices and so no players were assigned to those roles.

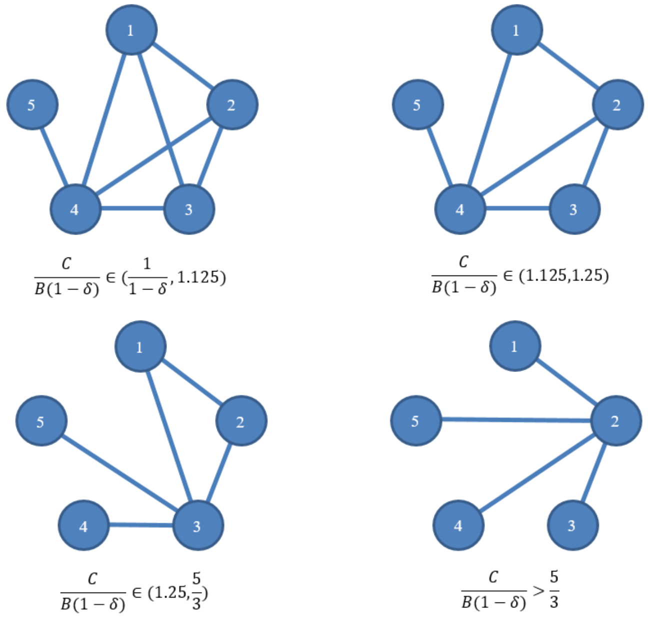

Figure 9: Typical outcome networks by region.

Figure 10 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 a player is satisfied with a set of connections, he or she 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 $C = 110$ treatments included question batteries to the end of the experiment. We included sets of questions designed to elicit risk preferences, beliefs, and personality characteristics.

The first battery of questions elicited risk preferences and 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 the Big Five Inventory of personality questions, aimed at detecting subject's entrepreneurial inclinations.[22] Participants were paid two dollars for completing the personality battery. We describe the three series of questions in more detail in Section 6 (for risk batteries) and in Appendix C (for the other batteries). [23]

4.3 Predictions

4.3.1 Move Predictions

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

- **Prediction A:** All players in both treatments will either play myopically or vie for dominance
- **Prediction B:** Player 5 will play myopically
- **Prediction C:** Player 4

[22] Zhao and Seibert (2006)

[23] Screenshots of the BFI questions were not included as the questions were identical to those in John and Srivastava (1999).

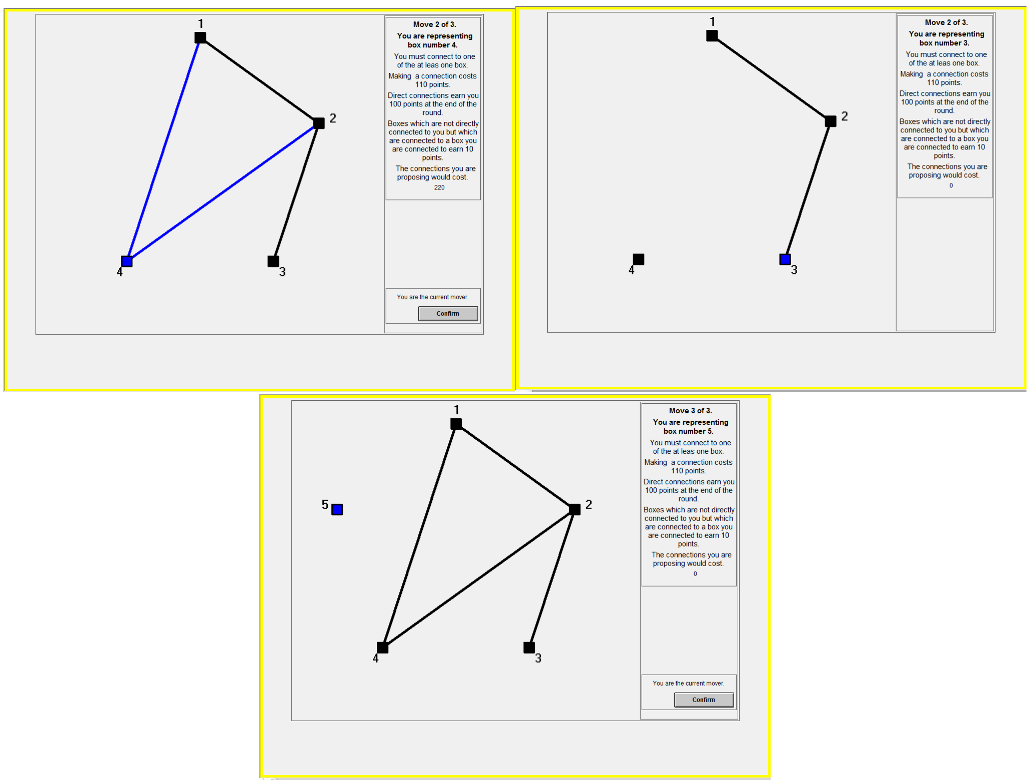

Figure 10: Examples of what players see during the network formation process. Beginning from the top left we have: (1) what Player 4 sees while making a decision; (2) what Player 3 sees while Player 4 makes a decision; and (3) what Player 5 sees after Player 4 made a decision.

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

Table 4: Player Payoffs: Predictions

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 5: Probability of Final Dominant Node: Predictions

- Player 4 will play myopically when $C = 140$ or when $C = 110$ and facing three dominant nodes
- Player 4 will vie when there is one dominant node and $C = 110$

- **Prediction D:** Player 3
  - Player 3 will vie in the $C = 140$ treatment
  - Player 3 will play myopically in the $C = 110$ treatment

4.3.2 Payoffs and Centrality

Proposition 5 can also be used to predict player payoffs. Table 4 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 payoffs. Predictably, 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 tend to vie for dominance and therefore make more connections on average.

When the connection cost increases from $C = 110$ to $C = 140$ welfare drops for two reasons. First, the cost of each connection increases which impacts welare directly. Second, the total number of connections goes down, decreasing the efficiency of the outcome network. The high cost treatment generates five connections while the low cost treatmentgenerates six.

In order to examine the relationship between entry timing and centrality we predict how often each node will end up as the most central at the end of the network formation process. Table 5 provides the probability that each player will end the game as the dominant node under Proposition 5.

Moving from the high cost to the lower cost treatment, probabilities shift such that the later nodes end up as the final dominant node more often. Lower costs mean more opportunities for later players to profitably vie for dominance. It is notable that, while there is a distinct early mover advantage in terms of average payoffs, such an advantage is not present in terms of centrality, with later nodes sometimes having strictly higher chances of achieving centrality than their early moving counterparts. As such, it is important to distinguish between these types of advantages. This dichotomy is analogous to the fact that in real world environments, the most important and connected firms/individuals may not be the most profitable or the most well off. A node that vies for dominance in equilibrium will do better than a node in the same position that fails to vie, but they may not do as well in terms of payoffs as a non-vying node who joined earlier in the process.

5 Results

In this section we summarize the results and compare them to the predictions about player behavior. We also discuss efficiency and node centrality in observed networks. Analysis is performed at both the network and the decision level. In sessions with $C = 110$, we observed a total of 1875 moves by 75

Treatment	Vie	Myopic	Total	Random Benchmark Total	One Connections Other	Multiple Connections Other
$C = 110$	0.075	0.707	0.782	0.309	0.076	0.141
$C = 140$	0.018	0.836	0.854	0.265	0.101	0.043

Table 6: Aggregate Move Proportions Excluding Player 3

subjects, or 635 networks; in sessions with $C = 140$ we observed a total of 975 moves by 39 subjects or 325 networks. [24]

Appendix B.1 examines the role of time trends in the data to look for evidence of learning and fatigue and find that behavior is largely consistent over time.. Because player behavior does not change over the course of a session in a way that substantially impacts results, we do not control for time trends or observed history of play in the analysis in the body of the paper. We do, however, find evidence of substantial cross subject heterogenetiy, so standard errors in decision level analysis are clustered at the subject level. All regressions are done using OLS with heteroskedastically robust standard errors.

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. We find deviations from the predictions of uniform random tie-breaking at the 5% level, but these deviations are not large enough in in magnitude to change predicted equilibrium behavior for the treatments we employed. Details are available in Appendix B.2

5.1 Prediction A

We now consider player behavior as it relates to the predictions. First, we examine the aggregate data and consider whether the data is consistent with Prediction A which states that all moves should be either myopic or vying for dominance. Table 6 reports the aggregate proportion of each move type. Player 3 is excluded, because he can only choose to vie for dominance or play myopically. This makes his data uninteresting for investigating Prediction A.

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.[25] 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 total proportion of vying and myopic moves is quite large both compared to the proportion of non-myopic, non-vying moves and compared to the random benchmark. From these results, we can say that Prediction A seems to generally hold in the data.

5.2 Prediction B: Player 5

In this section we look at the behavior of Player 5. Prediction B states that Player 5 should always connect to a single dominant node, which is the myopic move. Figure 11 shows aggregate the move proportions of Player 5.

Moves are categorized as myopic (as defined above), multiple connections, or one connection not myopic. One connection non-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 can be thought of as larger errors, because making multiple connections can be fairly costly relative to the optimal move.

[24] Data is pooled across accross groups, rounds, and sessions as is common for the type of analysis being conducted here, similar to network games experiments such as Charness et al. (2007),Charness et al. (2014), Carpenter et al. (2012), and Carrillo and Gaduh (2012).

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

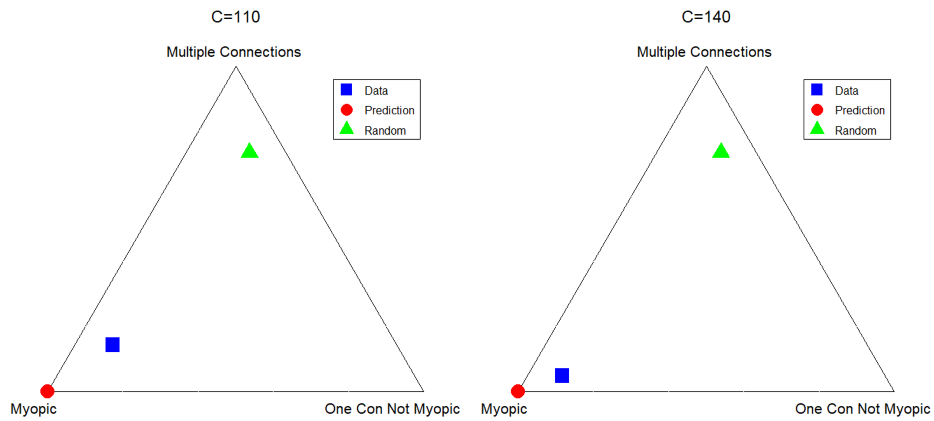

Figure 11: Player 5 move proportions. The closer a dot is to a corner, the more of the corresponding move type occurs. "Myopic" means one connection to a dominant node. "One Con Not Myopic" means one connection to a non-dominant node. "Multiple Connections" includes all moves with more than one connection.

	Estimate	Std. Error	Pr(>|t|)
(Intercept)	0.217	0.034	0.0000
C=110 3 Dominant	-0.131	0.034	0.0001
C=140 1 Dominant	-0.171	0.039	0.0000
C=140 3 Dominant	-0.208	0.035	0.0000

Table 7: 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.

The three colored dots represent the predicted move proportions (red circle), the observed move proportions (blue square), and the move proportions of a hypothetical player who chose their move in a uniform random way green triangle).[26] 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 B.

5.3 Prediction C: Player 4

We now look at the behavior of Player 4. Recall that Player 4's optimal move can depend on the network he faces. Prediction C states that Player 4 should vie for dominance when facing one dominant node in the low cost treatment and choose a myopic move in any other cost/state combination.

Figure 12 shows the aggregate 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. As before, dots of different shapes and 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.

Figure 12 shows that Player 4 data matches the theoretical predictions fairly well in the cases where myopic actions are predicted, so we can say that component of Prediction C is generally supported. However, when one dominant node with $C = 110$, we see more myopic actions and more "Other" Sthan

[26] This distribution was calculated by looking at the empirical distribution of network states faced by Player 5's and putting a an equal weight on each valid move available in that network state.

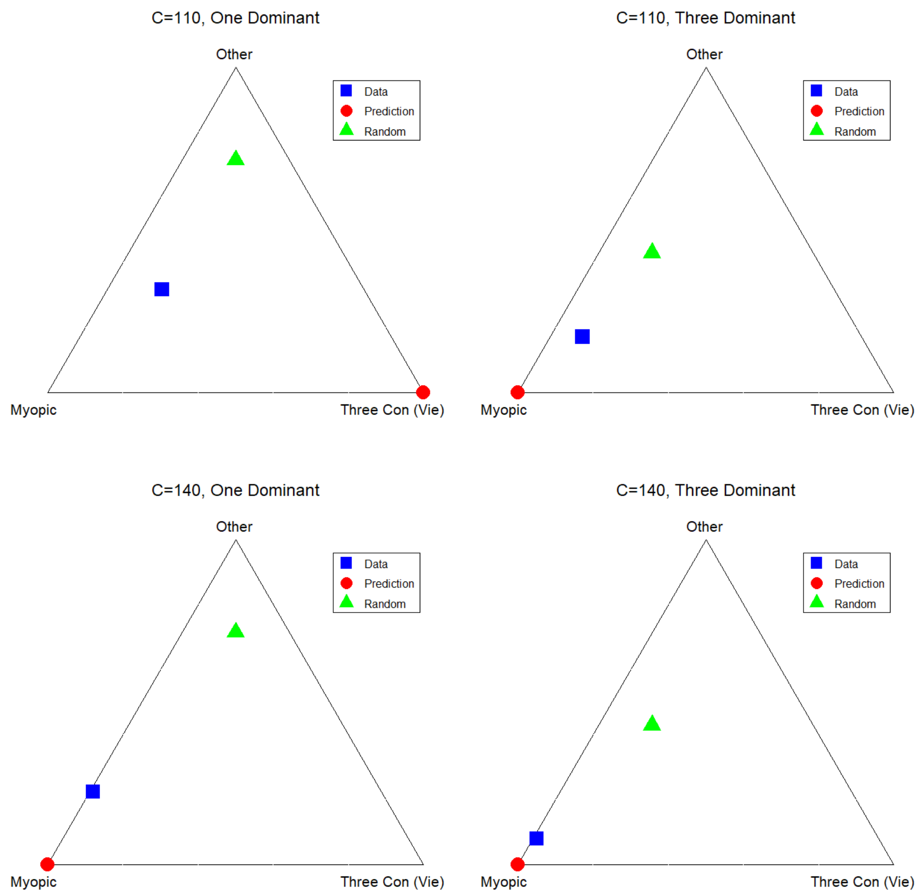

Figure 12: Player 4 move proportions. The closer a dot is to a corner, the more of the corresponding move type occurs. “Myopic” means one connection to a dominant node. “Three Con (Vie)” means three connections. “Other” includes all moves not covered by the other two categories.

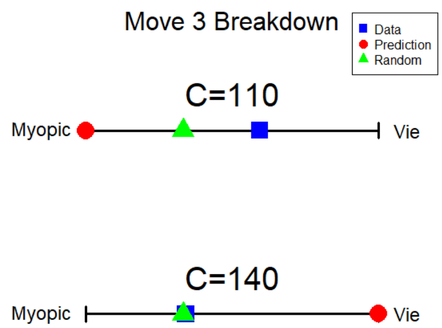

Figure 13: Move proportions for Player 3

vying, but the theory predicts vying in equilibrium.

That component of Prediction C does hold in the comparative sense, however. The proportion of three connection vies is higher for Player 4 facing one dominant node in the $C = 110$ treatment than for any other cost/state combination. Table 7 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 situations are significant and negative, meaning that Player 4 vied for dominance significantly more when $C = 110$, facing one dominant node than in any other cost/state combination.

5.4 Prediction D: Player 3

Consider now the behavior of Player 3. Player 3 only has two possible actions (treating the two one-connection moves as identical): he can either choose the myopic move or he can vie for dominance. Prediction D states that Player 3 should vie in the high cost treatment and play myopically in the low cost treatment.

Figure 13 reports the move proportions for Player 3. As before, shapes are used to 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 $C = 140$. Prediction D does not hold well in the data in either the absolute or the comparative sense.

5.5 Efficiency and Dominance

Before we move on to the interpretation of the data, it is important to note how the observed behaviors translate into node centrality and efficiency of generated networks, because these are ultimately the network features of interest.

Figure 14 shows the predicted and actual payoffs for players in each position in both treatments. In general, payoffs were quite close to the prediction. Players earned fewer points than predicted with the exception of Players 2 and 3 in the high cost treatment who likely benefited from the fact that players displayed a small bias against connecting to Node 1. Overall each network generated an average of 1357 points in the $C = 110$ treatment and 1172 points in the $C = 140$ treatment, a loss relative to predictions of 4.4% and 2.4% percentrespectively.

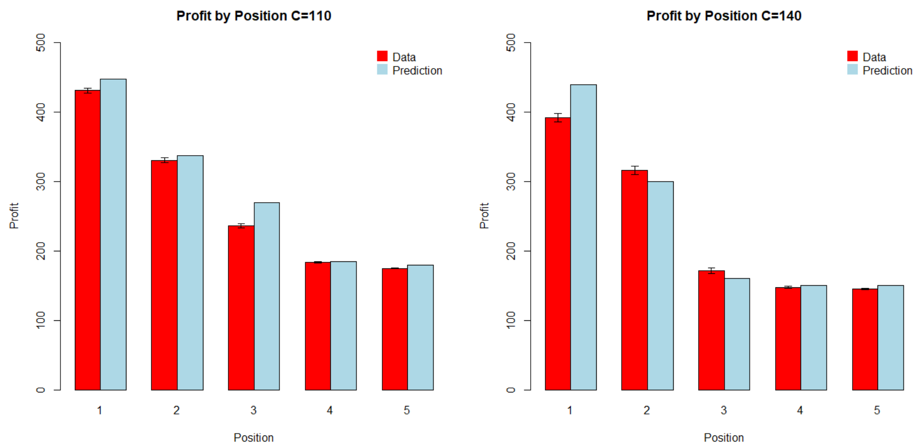

Figure 14: Average payoff by position, data and prediction. Note that payoffs for position 1 and 2 are hypothetical as no real players were associated.

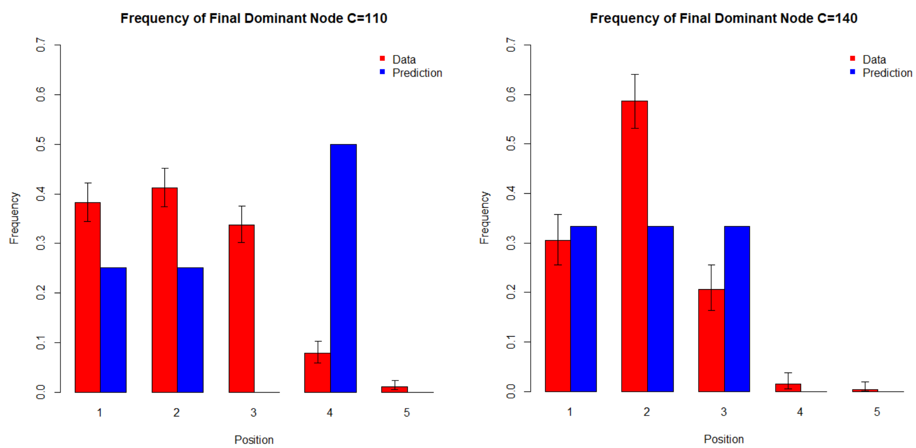

Figure 15: Histogram of the frequency for most connected node at the end of the network formation process $C = 110$ (Left) and $C = 140$ (right). Red bar show the observed proportion while blue bars show the predictions. Error bars show a 95% CI for the observed proportion.

The connectivity and hence efficiency of constructed networks did respond to the difference in connection costs. 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, which is a substantial difference, though less than predicted.

Figure 15 shows the frequencies of each node being the dominant node in the final network.[27] In the $C = 110$ case there is a definite shift earlier relative to predictions with earlier nodes being dominant more often than the theory would suggest. Player 3 is the final dominant node more often than Player 4.

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 also see a mild bias towards Node 2 over Node 1 in the $C = 110$ game. It is unclear why the bias in favor of Node 2 exists, but it does not have a strong influence on incentives of later players.

It is interesting to note that, while there are large deviations between the theory and data with

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

regards to which nodes end the game with a dominant position, there are only small differences between theory and data with regards to payoffs. This discrepency is not entirely surprising since, the subjects in positions 4 and 5 tend to take moves which are only slightly suboptimal, and since players often have the capacity to compensate for deviations from equilibrium in ways that preserve their payoffs.

5.6 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. The symbol **(T/F)** indicates a prediction which holds in a partial or comparative sense.

- **(T)Prediction A:** All players in both treatments will either play myopically or vie for dominance.
- **(T)Prediction B:** Player 5 will play myopically
- **Prediction C:** Player 4
  - **(T)**Player 4 will play myopically when $C = 140$ or when $C = 110$ and facing three dominant nodes
  - **(T/F)**Player 4 will vie when there is one dominant node and $C = 110$
- **Prediction D:** Player 3
  - **(F)**Player 3 will vie in the $C = 140$ treatment
  - **(F)**Player 3 will play myopically in the $C = 110$ treatment

The question then arises of why we observe these particular deviations from the theory.

5.7 Discussion

In order to better understand the deviations from the theory, it is useful to examine whether players in different situations are best responding. Figure 16 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 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.

Player 4 in the $C = 110$ facing one dominant node treatment does not seem to lose much by deviating from equilibrium. This comes from the fact that mistakes by player 5 tend to decrease the value of vying for dominance in this position while improving or not impacting the value of choosing myopic moves. This decrease in the benefit of vying may contribute to the lack of vying in this position.

Overall, it seems as though the only players who are not best responding the majority of the time are Player 3 in the $C = 140$ treatment and Player 4 facing one dominant node in the $C = 110$ treatment. In both of these cases, the payoffs from equilibrium and non-equilibrium moves are quite close, which means other features of player preferences could have a substantial impact. One very natural explanation is that risk aversion may be preventing some players from vying for dominance. Vying for dominance is inherently risky relative to choosing a myopic move, requiring an investment in connections which may not pay off. See Table8 for a comparison of the variability of payoffs from myopic and vying moves.

A number of factors point to risk aversion being a reasonable explanation for observed behavior. First as summarized in Table 9, players vie for dominance less often than predicted.

Second, players vie for dominance more in situations where the reward is higher. This is consistent with players having some type of aversion to vying activities which can be overcome with sufficient increases in expected payoffs. For each of eight the possible equilibrium relevant cost/state

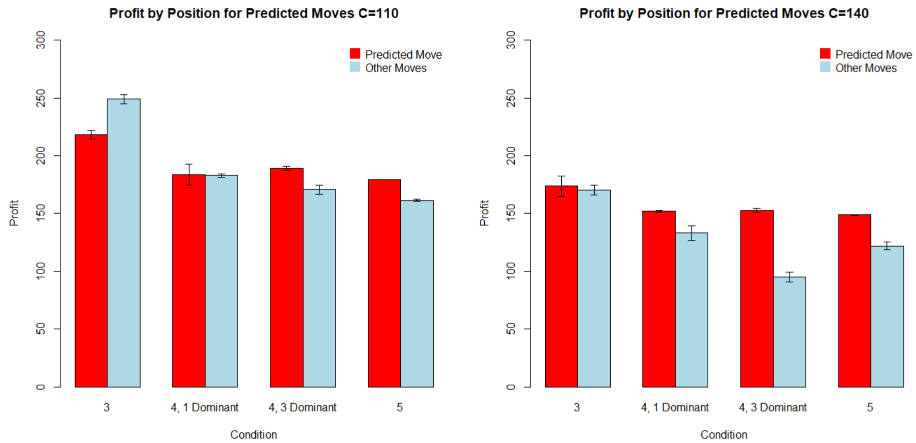

Figure 16: Observed payoffs by situation for players making the predicted move and players not making the predicted move

		Player 3	Player 4, 1 Dominant	Player 4, 3 Dominant
C=110	Vying	74.5	45.6	42.3
	Myopic	56.0	15.3	29.6
C=140	Vying	78.9	-	-
	Myopic	50.2	14.0	20.3

Table 8: Standard deviation of observed payoff after vying and myopic moves. Blanks indicate situations with fewer that three observations.

Treatment	Prediction	Observed
$C = 110$	0.333	0.27
$C = 140$	0.333	0.117

Table 9: Fraction of Players Vying for Dominance

	Estimate	Pr(>|z|)
(Intercept)	0.22283	0.000
Expected Gain from Vying (Vs Myopic)	0.00002	0.000
AIC	1808.96	

Table 10: Regression of vying by connecting to all nodes against expected gain vs myopic. Errors clustered at the subject level.

combinations,[28] 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. These values were used to calculate the empirical average gain from vying, relative to the myopic moves, in each situation. In Table 10 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.

Based on these facts and the intuition, it seems plausible that players might have some aversion to vying for dominance which can be overcome by higher expected gains.

6 Risk Aversion

In this sections, we explore the possibility that risk aversion may be preventing players from vying for dominance by eliciting risk preferences and using them to predict vying behavior. We also examine whether the aggregate patterns of play are consistent with an equilibrium in which players have heterogeneous risk averse utility functions.

6.1 Risk Aversion Data

As described earlier, we concluded two of the low cost sessions 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. One row from one list was chosen randomly, and players were paid based on their chosen gamble for that row. See Appendix C for screenshots and additional details of list construction.

In this paper, rather than estimate a single summary measure of risk aversion for each player, we estimate several situation-specific measures. Each of the MPLs was designed to mimic the trade-off between vying for dominance and choosing a myopic move in one situation, and a separate risk aversion measure will be estimated for each MPL.[29] We do this because there is evidence that risk aversion is a complex and multi-dimensional object[30] which may not be well captured by our specification. By using this type of situation specific measure we ensure that we are at least capturing some monotone transformation of the component of risk aversion which is relevant to each decision of interest.[31]

The left hand side of each list was a gamble mimicking the equilibrium payoffs from vying for dominance 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 as well as gambles which closely match the observed payoffs from myopic actions near the middle of the list. The point at which the player switches from the left gamble to the right gives us information about their situation specific risk preferences, which should help us predict whether that player is likely to choose a myopic move or vie for dominance in the situation corresponding to the list.

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

[29] We chose only situations in which the average payoff from vying was higher than the average payoff from choosing a myopic move, because these are the environments in which a risk averse individual may wish to vie.

[30] Kahneman and Tversky (1979), Hershey and Schoemaker (1985), and Sprenger (2015).

[31] Because MPL 1 and MPL 2 are potentially measuring unrelated characteristics, using a summary or instrumental variables based value for risk aversion may not be a valid approach.

MPL 1 mimics the trade-off of 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.[32]

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. 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 two risky options.[33] This could lead to differences in observed risk preferences between MPL 2 and MPL 1.

6.2 Risk Aversion Data Usage

The data from the MPLs was used in two ways: first, to categorize players based on their risk preference type and second, to estimate a risk aversion parameter. Based on whether each player's switch point was above or below the risk neutral switch point for a given list, the players were categorized as risk loving, risk neutral, or risk averse.[34] 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 sense of the distribution of risk aversion and the regularity of player choices, but they are not used in further analysis.

The data was used to estimate a situation specific risk aversion parameter from a CRRA model using the same stochastic model as Holt and Laury (2002).[35] A higher risk aversion parameter means a more risk averse utility function. Note that we are using the CRRA, so we assume

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

This model can be estimated for all players, even those who have behavior that is not consistent with expected utility maximization. If we focus on subjects that 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.[36] As such, our results should not be sensitive to our choice of utility function. The critical feature of the setup is that we construct lists which allow us to capture the correct component of risk aversion.

There is no theoretical reason in our framework to choose one utility function over another 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 robust to changes in the functional form, we will use the CRRA form in the main text and repeat the analysis with the CARA form in Appendix B.3. Note, there are some empirical reasons we might prefer CRRA to CARA, which are discussed in that section.

Tables 11 and 12 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 is likely fairly noisy, and the measure may also contain some information about how much effort and attention players are devoting to the game.

[32] 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 Appendix C for details.

[33] 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).

[34] 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 MPL 1, two switch points were classified as risk neutral.

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

[36] Note that in their paper, Holt and Laury (2002) use a power utility function with two parameters which nests both CARA and CRRA, but we only want a one-dimensional measure of risk aversion.

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

Table 11: Summary Statistics Estimated η_is

	Risk Averse	Risk Neutral	Risk Loving	Undef
Panel 1	8	7	3	12
Panel 2	1	12	3	14

Table 12: Estimated Risk Preference Type

6.3 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 to some type of non-standard preferences. The personality characteristics were included primarily as a control for individual heterogeneity in preferences for preference vying behavior 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.[37]

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 he or she would receive a connection from a later node. See Appendix C 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.[38] The elicited beliefs were used to construct an expected number of connections gained from vying for each player in each situation. Table 13 reports summary statistics about the additional number of connections players expected after vying relative to the myopic move.

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, Openness, and Extroversion. Table 14 summarizes the mean and the standard deviation for each category among our subjects.

6.4 Risk Preference and Vying

We now examine whether risk aversion can predict vying behavior in situations where vying for dominance is on average payoff improving relative to myopic moves in the $C = 110$ treatment (Player 3 and Player 4 facing one dominant node). We look at these scenarios in particular, because these are

[37] See Zhao and Seibert (2006)

[38] Using a quadratic scoring rule paying in probability points following Schotter and Trevino (2014)

	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 13: 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 14: Summary of Big Five Personality Metrics. Each measure is the sum of scores from related responses (with scores reversed where appropriate). “Out of” indicates the maximum score for each characteristic.

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 15: 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. ($< 0.1^{*}, < 0.05^{**}, < 0.01^{***}$)

the only two scenarios in which risk averse players may wish to vie for dominance in the low cost treatment. In all other situations, players should play myopic moves regardless of risk aversion. We only elicited risk aversion in $C = 110$ treatments for efficiency reasons; these treatments generated more vying behavior.

On the left side of Table 15 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 the appropriate η was negative and significant. In addition, no other coefficients other than the intercept are significant in any specification.

The right side of Table 15 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 one of the controls (neuroticism) is also significant and negative.[39]

In general, risk aversion appears to be influencing vying for dominance behavior. For contrast in Appendix B.5, 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 $C = 110$, a condition in which risk averse individuals should never vie for dominance. As predicted, there is not

[39]It is possible that the neuroticism measure is capturing some features related to the ability to easily deal with new situations. To some extent, Neuroticism may be thought of as an inverse of willingness to explore. Zhao and Seibert (2006) found that neuroticism is negatively correlated with entrepreneurial tendency, so the relationship here is not entirely surprising.

a significant impact. In addition, the coefficients on the risk aversion are substantially smaller.

Some readers may be concerned that our risk aversion measures may also be picking up information on participant confusion or attention due to the large number of undefined subject. As we show in Appendix B.4, the players with undefined risk types do not have a substantial impact on magnitude the relationship between risk aversion and vying, although the reduction in power does remove the significance of risk aversion in several specifications.

The fact that beliefs do not have a significant impact on vying in Table 15 may raise concerns that the measurement of beliefs was very noisy. This may be true to some extent, but colinearity between beliefs and risk aversion also plays a large role. If we remove risk aversion from the estimation, coefficients on beliefs become consistently positive and grow substantially in magnitude. In one specification they become significant at the five percent level. Regressions without risk aversion measures are included in Appendix B.6. These results indicate that overall risk aversion is better at explaining vying when controlling for beliefs than beliefs are when controlling for risk aversion. This seems to indicate that the risk aversion measure has more unique information about vying behavior than the beliefs measure.

6.5 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 consider a broad class of models with heterogeneous risk aversion and compare the equilibrium predictions of those model to the data.

The model class we use has a very general structure. We assume players have utility functions:

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

where $g(x)$ is a concave function of x which is drawn from some population G with replacement immediately before Player i makes their choice.[40] We assume that players know G but not the precise g_i of other players. We are ruling out are risk seeking behavior and behavior inconsistent with expected utility maximization. We still assume random tie-breaking

While the model class is very general, it has fairly specific predictions in this context. The first fact of note is that Proposition 4 will still hold in any equilibrium of a model in this class.

Risk Aversion Prediction A: In game with heterogeneous risk aversion Players will only ever choose myopic moves or vie for dominance.

This prediction comes as a special case of the robustness result in Appendix A.3. In addition, only a few situations can generate a non-zero amount of vying play in an equilibrium of a model in this class.

Risk Aversion Prediction B: In game with heterogeneous risk aversion, when $C = 110$, only Player 3 and Player 4 facing one dominant node can vie.

Risk Aversion Prediction C: In game with heterogeneous risk aversion, when $C = 140$ only Player 3 can vie.

Proofs for Risk Aversion Predictions B and C are in Appendix A.5. Table 16 summarizes in which situations vying may be possible in an equilibrium of the type described.

6.5.1 Aggregate Evidence for the Risk Aversion Model

We now examine how well the move restrictions generated by the model class with heterogeneous risk aversion match the data. In Table 17 we show the vying probability in each situation, ordered from most observed vying to least. The three highest proportions of vying for dominance behavior occur in the three situations where it is possible for vying to occur in a heterogeneous risk aversion model.

[40] In the experiment, players are not actually being drawn with replacement, so players could theoretically update their beliefs based on behavior earlier in the round, but we are assuming that the population is large enough and signals are weak enough to make this irrelevant.

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

Table 16: Potential for Vying in the General Model with Risk Aversion

	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 17: Vying Proportions: Data and predictions from the Risk Averse Model with risk neutral predictions for reference.

The question then arises of whether it is possible to construct a specific population of utility functions which will generate the observed proportions of vying in those three situations with reasonable utility functions. 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 $\kappa = 1$.

From the data, we estimate very different values: $\tilde{\alpha} = 0.151$, $\tilde{\beta} = 0.576$, and $\tilde{\kappa} = 0.342$. These proportions can be supported by an infinite number of possible utility function populations. One such population is $u_i = g(x, 2000, 5)$ with probability 0.151, $u_i = g(x, 2000, 5.5)$ with probability 0.191, $u_i = g(x, 2000, 5.7)$ with probability 0.234, and $u_i = g(x, 2000, 10)$ with probability 0.424 where $g(x, b, \eta) = \frac{(b+x)^{1-\eta}}{1-\eta}$. This is effectively a population of CRRA utility functions modified to give each player a baseline wealth of 2000 points.[41] To see how this population was constructed to match the data, see Appendix A.6.

While it is difficult to characterize the set of all possible combinations of α, β, and κ that can be supported by a subgame perfect equilibrium with some population of utility functions, we do know that not all combinations can be supported. 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.

The predictions of the risk averse model are able to fit all the major moments of the data to a first approximation. Risk aversion seems like the best candidate to explain the data both on an individual choice level and in aggregate. It seems reasonable then to consider risk aversion as the best explanation for deviations from the baseline theory. In Appendix B.7, we consider alternative behavioral models, but we do not find any of them to be promising as alternative explanations of the data.

7 Extensions

While the model discussed in this paper does make some strong assumptions about the structure of the game, many of the results have analogous forms which hold when the assumptions are loosened. One strong modeling assumption that we can loosen is the requirement that players make at least one connection as they enter the network. If we eliminate that requirement, vying and myopic behaviors

[41] 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.

still play a critical role in how the network develops. Consider an alternate game which is identical to the base game except players may choose empty connection sets. The following variant of Proposition 4 holds in this modified game.

Proposition 4 Alternate 1: If $C/B > 1$ and $J = 5$, all players will either make one connection to a single dominant node, make no connections, or vie for dominance.

Proof in Appendix A.8. This alternate proposition shows that while the less restricted version of the game does allow for a more complex set of behaviors, the major lessons from the baseline game are still somewhat applicable.

Another modeling choice that we can generalize is the assumption that players can only make connections as they join the network. Consider a second alternate version of the model in which after all players have joined the network, players are given the opportunity to move again in some arbitrary sequence which is commonly known at the start of the game.[42] When a player moves after their initial entry into the network, they can choose to add connections to any node they are not already connected to. In this modified game, the following alternate Proposition 4 holds.

Proposition 4 Alternate 2: If $C/B > 1 + 2\delta - \delta^2 - \delta^3$ and $J = 5$, then all players will either make one connection to a single dominant node, make no connections, or vie for dominance.

Proof in Appendix A.8. The single movement opportunity of the base model is not a critical assumption for generating vying for dominance behavior. Even when we allow for different move timings, order of entry generates an incentive for earlier players to compete for connections from late joiners.

These alternate propositions are by no means an exhaustive list of possible ways that extensions to the model can generate similar results. They merely serve as an illustration of the fact that vying for dominance behavior is not exclusive to the restrictive model we used in this paper. As long as centrality is beneficial, players join in order, and connections are costly both costly and persistent, vying for dominance can play a role.

We discuss the above extensions and others, such as node heterogeneity and in more detail in Neligh (2017).

In Neligh (2017) we also show that these results are not just relevant to the study of small networks. In that paper we discuss a variant of the game which remains tractable even for very large networks. In the Dominant Node Restricted Game, players are required to connect to at least one dominant node as they enter the network. Under this restricted model, we have a result analogous to Proposition 4 whereby in any Markov perfect equilibrium players will always connect to all existing nodes or to some subset of the current dominant nodes. 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).

Furthermore, when players break ties in favor of connecting to newer nodes, the Dominant Node Restricted Model predicts periodic vying for dominance with the time intervals between vying players increasing exponentially over time. Players who vie for dominance do so because they want connections from later 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. Players who vie for dominance, therefore, need a period of non-vying nodes after their move to make up for the cost of vying. 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 moves needed to make a vying move profitable is increasing over time.

[42] The sequence can be completely arbitrary. Some players may move many times, while other players may only move once as they join the network.

8 Conclusion

In this paper, we presented a new network formation model that uses dynamics and forward looking strategic agents to explore the relationship between entry timing and node centrality. The model predicted a number of novel behaviors including "vying for dominance" 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. Only by vying can a player potentially become the dominant node in the network.

We provided 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 simplify the game for greater tractability by focusing on small networks.

In the latter part of this paper, we solve the game with 5 nodes 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 better 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 explored the possibility that risk aversion may have been preventing players from vying for dominance. 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, in general, did not have such a relationship in general. We also examined the equilibrium predictions of a general class of models with heterogeneous risk aversion. Models in this class always predict that vying for dominance should occur in only three situations. In the data, players vied for dominance in those three situations more than in any others. Overall, risk aversion was found to be the best explanation for observed behavior.

We conclude then that entry timing can provide players with the opportunity to attempt to become central, but only the more risk neutral and entrepreneurial players will choose to exploit that opportunity. Only by having the right player enter the network at the right time can high centrality be achieved.

There are also several important implications of the model which do not fit into the framework of the discussion on the experiment. For example, the fact that different cost levels lead to network configurations of different efficiency suggests that a tax subsidy scheme on connections could be welfare improving in many cases. A planner could impose a flat tax mirrored by a subsidy on connections, thereby enforcing the network structure associated with any cost level. The number of connections made at a given cost level is generally predictable, making budget balance easy to achieve at least in theory.

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.

Subsidizing connections in this manner in order to generate a complete network is welfare increasing as long as $C < 2(B-b)$. By Proposition 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. Testing the role of taxes and subsidies in network formation would be a natural direction for further research.

Another possible avenue is to explore the role of entry timing and risk aversion in a context of immediate economic interest. The predictions of this model are most applicable in systems where centrality is beneficial and connections are both costly and stable. These characteristics hold true in a number of interesting networks, including citation networks and networks of professional mentorship.

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 from Neligh (2017) could provide an appealing way to apply the lessons from this model to larger groups.

Neligh (2017) covers many theoretical extensions and modifications to the baseline game at a broad level but many extensions are not explored in great depth. Additional work may uncover surprising and useful findings in these different contexts. There is also further work to be done in improving the tractability of the baseline game with larger networks.

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 features of interest. This paper serves to provide an example for future work, showing how by taking the dynamic and strategic elements of network formation seriously, we can generate intuitive and novel predictions about the evolution of important systems.

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).

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 networks. *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).

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

Neligh, N. (2017). The theory of 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).

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.