6-2012

Behavioral Experiments on a Network Formation Game

Michael J. Kearns
University of Pennsylvania, mkearns@cis.upenn.edu

Stephen Judd
University of Pennsylvania, sjudd@seas.upenn.edu

Yevgeniy Vorobeychik
Sandia National Laboratories

Recommended Citation

Michael J. Kearns, Stephen Judd, and Yevgeniy Vorobeychik, "Behavioral Experiments on a Network Formation Game", . June 2012.

Kearns, M., Judd, S., & Vorobeychik, Y., Behavioral Experiments on a Network Formation Game, *ACM Conference on Electronic Commerce*, June 2012, doi: 10.1145/2229012.2229066

Behavioral Experiments on a Network Formation Game

Abstract

We report on an extensive series of behavioral experiments in which 36 human subjects collectively build a communication network over which they must solve a competitive coordination task for monetary compensation. There is a cost for creating network links, thus creating a tension between link expenditures and collective and individual incentives. Our most striking finding is the poor performance of the subjects, especially compared to our long series of prior experiments. We demonstrate that the subjects built difficult networks for the coordination task, and compare the structural properties of the built networks to standard generative models of social networks. We also provide extensive analysis of the individual and collective behavior of the subjects, including free riding and factors influencing edge purchasing decisions.

Disciplines

Computer Sciences

Comments

Kearns, M., Judd, S., & Vorobeychik, Y., Behavioral Experiments on a Network Formation Game, *ACM Conference on Electronic Commerce*, June 2012, doi: 10.1145/2229012.2229066

Behavioral Experiments on a Network Formation Game

Michael Kearns, University of Pennsylvania
Stephen Judd, University of Pennsylvania
Yevgeniy Vorobeychik, Sandia National Laboratories

Abstract: We report on an extensive series of behavioral experiments in which 36 human subjects collectively build a communication network over which they must solve a competitive coordination task for monetary compensation. There is a cost for creating network links, thus creating a tension between link expenditures and collective and individual incentives. Our most striking finding is the poor performance of the subjects, especially compared to our long series of prior experiments. We demonstrate that the subjects built difficult networks for the coordination task, and compare the structural properties of the built networks to standard generative models of social networks. We also provide extensive analysis of the individual and collective behavior of the subjects, including free riding and factors influencing edge purchasing decisions.

Categories and Subject Descriptors: Computer Applications [**Social and Behavioral Sciences**]: Economics, Psychology, Sociology

General Terms: Economics, Experimentation, Human Factors, Theory

Additional Key Words and Phrases: Social Networks, Game Theory, Network Formation

1. INTRODUCTION

In recent years, research from a variety of disciplines has established the universality of certain approximate structural properties of large-scale social, technological, organizational and economic networks. These properties include networks having small diameter, high clustering of connectivity, and heavy-tailed degree distributions. The apparent ubiquity of these properties, despite the diversity of the domains of the networks in which they appear, has led researchers to seek explanations in the form of models of network formation that can reliably generate the observed structures.

The most studied class of such models are *stochastic* network formation models, in which networks form through a decentralized process that generates local connectivity using randomization; examples include the classic Erdös-Renyi random graph model [Bollabas 2001], and the more recent small worlds [Watts and Strogatz 1998; Kleinberg 2000] and preferential attachment [Barabasi and Albert 1999] models. These models have been successful in providing simple and relatively general mechanisms generating common structural properties of large networks.

An important criticism of the stochastic formation models is that in real networks, connectivity rarely forms entirely randomly — rather, there is often some significant component of *purposefulness* when a new node or link is formed. Professionals join and form connections on a service like LinkedIn to enjoy the career benefits of being part of that network; new web sites are created to generate traffic, including by being

---

This work was conducted while Y.Vorobeychik was at the University of Pennsylvania.

EC'12, June 4–8, 2012, Valencia, Spain.

part of the network of links and the attendant search indexing benefits. While there may be elements of arbitrariness or stochasticity, networks generally arise from the self-interests of their constituents, and serve some collective purpose(s).

An alternative class of economic or game-theoretic models directly addresses the issue of self-interest and purpose in network growth. In *network formation games* [Tardos and Wexler 2007; Jackson 2005; 2010; Fabrikant et al. 2003; Borgs et al. 2011; Albers et al. 2006; Brautbar and Kearns 2011; Even-Dar et al. 2007; Even-Dar and Kearns 2006], individual players typically have utility functions with two competing components: there is a cost to join the network, usually in the form of purchasing links to other players (where the cost may be viewed as monetary, as in the connectivity and physical costs of adding a router to the Internet, or more cognitive, as in the time needed to create and maintain friendships on Facebook). But after joining the network, a player enjoys participation benefits (perhaps abstracted by some measure of their centrality in the network), and their overall payoff is the network benefit minus their cost of joining. It is common to equate the outcome of such games with their Nash equilibria, just as the stochastic models are analyzed for their statistically typical properties. While there has been growing interest in the theory of network formation games for several years now, to our knowledge there is not an accompanying behavioral literature.

In this paper, we describe among the first and largest human-subject experiments in a pure network formation game. These are the most recent in a long series of behavioral experiments on strategic and financial interactions in social networks [Kearns et al. 2006; Kearns et al. 2009; Judd et al. 2010; Kearns and Judd 2008; Chakraborty et al. 2010; Kearns et al. 2011], but represent a major departure from our prior experiments, where the networks examined were always exogenously imposed on the subjects. Here we *endogenized* the formation of the network structure itself as part of the experiment. While the theoretical literature on network formation games has focused on one-shot, simultaneous move games of full information (and even there, characterizations of equilibria are difficult and elusive), our experiments investigate a formation game of continuous, asynchronous moves and partial information.

In the experiments, subjects were given financial incentives to solve a collective but competitive coordination problem of *biased voting*, in which they must unanimously agree on one of two alternative choices, or receive no payoff at all. The competitive aspect arises from the fact that different players have different financial preferences for which of the two choices is agreed upon. We have previously studied this problem on fixed network structures [Kearns et al. 2009]; in the current experiments the subjects themselves had to *build* the network during the experiment, via individual players purchasing links whose cost is subtracted from their eventual task payoff. The nature of the biased voting task and the financial self-interests of the players sets up a clean strategic tension: in order to solve the biased voting problem, the players must collectively purchase enough links to establish some minimal global connectivity; but any individual player would prefer others to incur the costs of building this shared infrastructure.

A striking finding is that the players performed very poorly compared to our long series of prior experiments in which network structures were imposed exogenously. Despite clearly understanding the biased voting task, and being permitted to collectively build a network structure facilitating its solution, subjects instead appear to have built very *difficult* networks for the task. This finding is in contrast to intuition, case studies and theories suggesting that humans will often organically build communication networks optimized for the tasks they are charged with, even if it means overriding more hierarchical and institutional structures [Burns and Stalker 1994; Nishiguchi and Beaudet 2000]. We also report on a number of other aspects of subject behavior and performance, including structural properties of the built networks, comparisons to standard network formation models, and free-riding in edge purchasing.

2. EXPERIMENTAL DESIGN AND METHODOLOGY

The experiments reported here were held in three sessions with different pools of 36 subjects each. As in our previous experiments, subjects sat at networked workstations separated by physical partitions, and the only communication permitted was through the system. In each of many short experiments, subjects were given financial incentives to solve a global coordination problem via only local interactions in a network. In prior experiments, the network structure was a design variable that we chose and imposed exogenously in each experiment; here, the network structure was created by the subjects themselves, as described below.

We first describe the overarching collective task the subjects were charged with solving. In each experiment, subjects sat at individual workstations, and each controlled the state of a single vertex in a 36-vertex network whose connectivity structure evolved throughout the experiment. The state of a subject's vertex was simply one of two colors (red or blue), and could be asynchronously updated as often as desired during the one-minute experiment. Subjects were able to view the current color choices of *only their immediate neighbors in the current network* at all times. No communication between subjects outside the experimental platform was permitted.

In each experiment, each subject was given a financial incentive that varied across the population, and specified both individual preferences and the demand for collective unity. For instance, one player might be paid \$2 for blue consensus and \$1 for red consensus, while another might be paid \$1 for blue consensus and \$2 for red consensus, thus creating distinct and competing preferences across individuals. However, payments for an experiment were made only if (red or blue) *global* unanimity of color was reached; thus subjects had to balance their preference for their higher payoff color with their desire for any payoff at all.

At the beginning of each 1-minute experiment, the network over the players was typically *empty*: there were no edges, and thus every player controlled an isolated vertex and could see only their own color choice. Clearly reaching unanimity of color choice in the biased voting task is highly unlikely in such circumstances. Thus at any time throughout an experiment, subjects were free to *purchase* edges to other players at a fixed cost. Edge purchases were unilateral — they did not require approval by the player on the receiving end — but their benefits were bilateral, meaning that after the purchase both players could see each other's current color choices. The system GUI (see Figure 1) would dynamically evolve as new edges were purchasing, always showing a subject the color choices of their neighbors in the current network.

An important design decision is what information the players are provided to help them decide *which* vertices to purchase edges to. One possibility is *no* information: players could simply indicate their desire for a new connection, and the system could simply give them a new edge to a random player to whom they were not already connected. However, this would predetermine the network topologies built to be of a random, unstructured, Erdös-Renyi variety, and not particularly useful for the biased voting task. We thus decided to give subjects two pieces of information about their current non-neighbors: their current *degree* (number of connections), and their current *shortest-path distance* in the network from the subject. This allowed players to selectively purchase edges that seem relevant to the task. For instance, players could choose to buy edges to players with high degree who were distant from them in the current network (perhaps in the hopes that such players aggregate information on the other side of the network), or to players with zero or low degree (perhaps in the hopes of having strong influence over the color choice of such players). We emphasize that

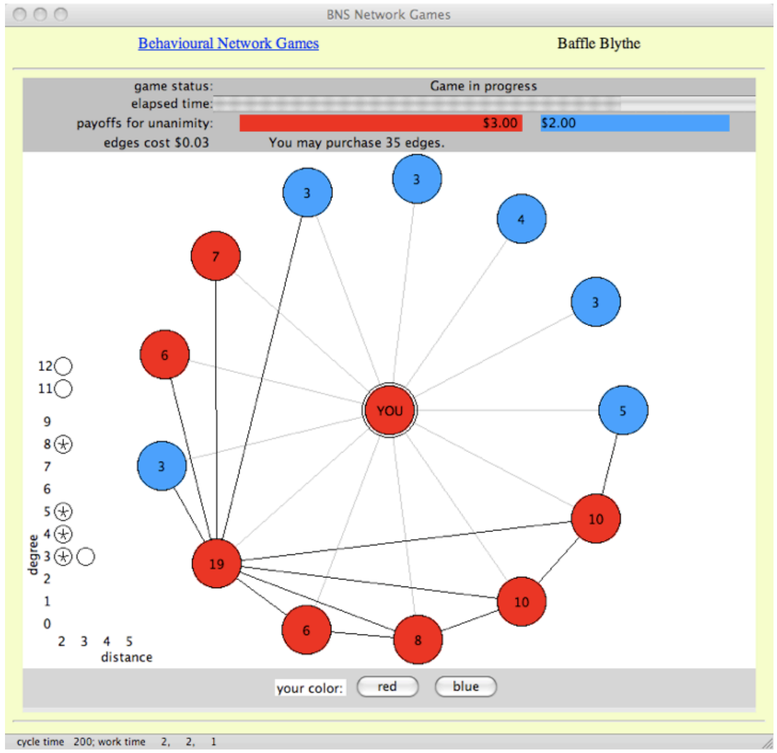

Fig. 1. Sample screenshot of player GUI in the network formation game. Around the vertex labeled "YOU", the central panel displays the player's current network neighborhood, indicating the current color of their neighbors as well as any edges between neighbors. In the action panel at the bottom, the player can change their own current color by clicking on the buttons labeled "red" and "blue". To the lower left of the central panel is the grid where players can select other vertices to purchase edges to. Each non-neighbor is represented by a circle whose grid position indicates their current degree and shortest-path distance from the player. (Vertices not currently in the same connected component as the player are shown as being at infinite distance.) If more than one vertex has the same degree and distance, the circle contains a "*" symbol. The player purchases edges by clicking on the desired circle, at which point their new neighbor will be incorporated into the neighborhood display. At the top, the time elapsed in the experiment is shown, along with the payoffs of the player, which are dynamically reduced as edges are purchased. The fixed cost per edge is also displayed.

this choice of informational design was not made with realism and generality in mind — obviously, in real social networks one does not have knowledge of the degree and distance of non-neighboring vertices — but rather potential relevance to the collective task, which seemed to us a more important experimental criterion.

We also note that this informational design also permitted the subjects, in principle, to collectively generate networks similar to those of well-studied stochastic models such as Erdös-Renyi (by simply ignoring the degree and distance values, and always choosing a random non-neighbor to connect to), or networks similar to those generated by preferential attachment (by ignoring distance information, and favoring purchasing edges to higher-degree vertices). We know from previous fixed-network experiments that such models generate networks generally favorable for human performance across a wide variety of tasks, including biased voting [Kearns et al. 2009]. Thus at least some behaviorally "easy" networks are collectively reachable within the given system design.

Each player's GUI had an edge-purchase panel in which each of their current non-neighbors was represented by a circle on a 2-dimensional grid, indicating that non-neighbor's current degree and distance from the player; see Figure 1. By simply clicking on the corresponding circle, a player would purchase an edge, and the new neighbor and their color would be dynamically incorporated into their network neighborhood display, and remain for the duration of the experiment. (Edge purchases were persistent and irrevocable.) If more than one non-neighbor had the same degree and distance, the grid circle would indicate so. As an experiment progressed, degrees increased and distances decreased in the growing network.

If the players failed to reach unanimity in the allotted time, all edge purchases were forgiven, and no payoffs were made; but if unanimity was reached at any point, the experiment was terminated, and a player's edge purchases were subtracted from their earnings on the biased voting problem to arrive at their net payoff. The system also enforced the condition that players must have strictly positive payoffs on successful experiments: thus, each player could only spend an amount on edges that was slightly and strictly *less* than their lower-payoff color in the biased voting problem. This prevents players from becoming "infinitely stubborn" in favor of their higher-payoff color if their lower payoff has been reduced to zero by edge purchases.

In a subset of experiments, conditions were as described above, but instead of starting with the empty network (which we shall refer to as *unseeded* experiments in the sequel), the experiment began with a "seed" network of edges that were provided free of charge to the players [Kleinberg 2000; Even-Dar and Kearns 2006]; players could then optionally purchase additional edges as above. Thus each experiment was characterized by the distribution of biased voting incentives of the players, the presence or absence of the seed network and its structure, and the fixed price of edges (which we varied from experiment to experiment). We shall comment on each of these design variables in the appropriate places as we describe our findings.

From a theoretical perspective, we thus presented our subjects with a task-oriented network formation game of partial information (unknown incentives or types of the other players, unknown and evolving global network structure) and asychronous, repeated moves with finite termination time. We note that formal analysis of even vastly simplified versions of this game appears to be quite challenging, but might be an interesting avenue for future work.

3. BACKGROUND ON PRIOR FIXED-NETWORK EXPERIMENTS

As mentioned above, we have conducted experiments similar to those described here since 2005, but always designing and exogenously imposing the network structures mediating interaction. The tasks we have given to subjects are diverse, and include graph coloring [Kearns et al. 2006], consensus [Judd et al. 2010], networked trading [Kearns and Judd 2008], networked bargaining [Chakraborty et al. 2010], independent set [Kearns et al. 2011], and biased voting [Kearns et al. 2009]. While direct comparisons across tasks can be difficult, there is one general and easily measured metric of collective performance, which is the *efficiency:* in any given experiment, we can compute the configuration of play that would have maximized the total payments to the subjects, and then compute the fraction of that maximum payoff the population actually realized. We can then average this quantity across all experiments ever conducted, regardless of task, network structure, and other design variables. The resulting value is 0.88 — in other words, over the lifetime of the project, subjects have extracted almost 90% of the value that was available to them in principle. We conclude that humans are quite good at solving a variety of challenging tasks from only local interactions in an underlying network.

In the previous biased voting experiments on exogenous networks, 55 of 81 experiments resulted in unanimity and therefore some payoff to all subjects, again giving fairly strong collective performance. On the subset of experiments in which the network and incentives had what we called a *minority power* structure, performance was even stronger, with 24 of 27 experiments reaching consensus. We shall contrast these findings with those for biased voting with network formation.

4. RESULTS

4.1. Overall Performance

Our experiments were structured in three separate sessions with different subject pools; while the conditions in the first two sessions were similar and designed in advance, the third session was designed in response to the findings of the first two, and shall be discussed separately below.

Session 1 consisted of 99 short experiments, with 63 of these being unseeded; Session 2 consisted of 72 experiments, 27 of which were unseeded. Across these 171 experiments, various other conditions varied as well, including the cost per edge, the fraction of players with higher payoffs for red, and the relative strengths of the incentives for players of different types. We shall discuss the effects of these design variables later, and for now focus on the collective performance across all these conditions in Sessions 1 and 2.

Compared to our long series of prior fixed-network experiments, that performance was surprisingly poor: Session 1 produced only 47% successful (unanimous) outcomes, and Session 2 only 39%, for an overall success rate of 44%. This is in sharp contrast to the aforementioned efficiency across all tasks of 88% — approximately double that of the current experiments — and the 68% success rate of the fixed-network biased voting experiments, more than 20% higher than in the network formation game. It appears that allowing the subjects to control the creation of the network significantly *worsened* collective performance [1].

There are at least two plausible explanations for this degradation in performance that do not simply entail that subjects built "bad" networks for the task. The first explanation is one of cognitive *overload*: perhaps subjects built "good" networks, but simply ran out of time to solve the biased voting problem on those good networks. The second explanation is one of *stubbornness* due to modified incentives: perhaps the subjects built good networks, but due to the edge purchases, some players had reduced the net payoff of their less preferred color to such a small amount that they might be very resistant to acquiesce to the majority color, resulting in stalemates. This hypothesis is made more plausible by the significant amount of "free riding" that occurred with respect to edge purchases, discussed later.

To investigate the overload and stubbornness hypotheses, we designed and conducted a third session of experiments with fresh subjects. In Session 3, each game was seeded with a network that was the *final* network constructed by the subjects in an unseeded Session 1 experiment. In these Session 3 experiment, the subjects *only* played the biased voting game — *no edge purchases* were allowed by the system. The subjects were thus back in the setting of our earlier, exogenous fixed-network experiments, but this time using networks *built by previous human subjects*. We deliberately chose a subset of the final networks from Session 1 on which the performance there was particularly poor — namely, 18 final networks on which the success rate was only

[1] In the games that failed to converge, the average size of the minority was actually smaller than it was for the failed biased voting games, but not to the level of statistical significance.

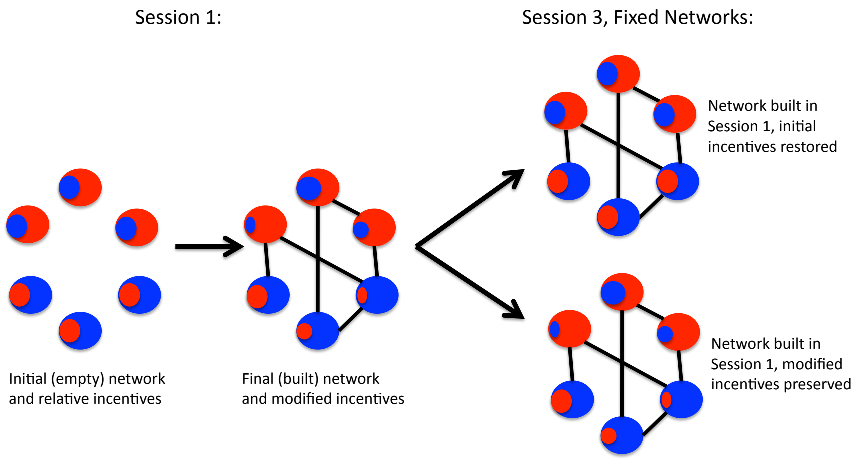

Fig. 2. Design of Session 3 experiments. At the start of an unseeded Session 1 experiment (left), the network is empty and some players (the upper 3) have higher payoffs for red than blue, schematically represented by the ratio of the red to blue areas in each vertex. At the completion of the experiment (middle), a network has been built by the players, and some players may have more extreme relative preferences due to the reduction of their payoffs by edge purchases. In Session 3 (right), we took the final networks built by Session 1 experiments, and exogenously imposed them as the networks of a pure biased voting game without any edge purchases. We did so using both the original, restored incentives of Session 1, and the post-edges modified incentives.

17%. We ran each Session 3 network under two different incentive conditions: one in which the incentives the players were the same as at the *beginning* of the corresponding Session 1 experiment, and one in which they were the same as at the *end* (after edge expenditures). See Figure 2 for a description of Session 3 design.

Together these conditions allow us to investigate the validity of the explanations above: if subjects were simply running out of time in Session 1 (the overload hypothesis), they should fare much better in Session 3, since now the network formation task is removed, and they can focus only on the biased voting task; and if the difficulty in Session 1 was due to stubbornness after edge purchases, the Session 3 experiments in which the Session 1 networks are used but the incentives are restored to their starting values should be more successful.

The stubbornness and overload hypotheses are strongly refuted by the Session 3 results. Success rates were slightly higher than in Session 1, but not significantly so. The strong signal was that Session 3 success rates in games using the original payoffs, and in games using residual payoffs, were both significantly lower than in our earlier fixed-network biased voting games [Kearns et al. 2009], with $P < 0.01$ in both cases. The success rate (57/81) during all games in the fixed-network biased voting session was significantly higher than that of all games (100/243) in the three network formation sessions ($P < 0.0001$). More pointedly, the games in each of the three network formation sessions are individually lower than the fixed-network biased voting games (all with $P < 0.01$).

We are thus led to the conclusion that our subjects *built networks on which it was simply difficult to accomplish the very task they were being paid to accomplish.* They appear to have had enough time and incentive, but built inherently poor networks. As per the earlier discussion, this is the first task of the many we have investigated in which human performance was so low.

4.2. Effects of Seed Networks

Recall that a subset of the Session 1 and 2 experiments explored network formation in which the subjects were provided an initial seed network as free infrastructure. Our goal was to examine whether having this seed network, which might facilitate communication and coordination at no cost, would allow the players to build better networks and yield stronger performance.

We examined three types of seed network structure: a 2-dimensional grid or torus network, which provides global connectivity with relatively few edges; a network of 6 cliques of size 6, which groups the players into small highly-connected communities; and preferential attachment networks that were also the focus of the minority power experiments we discuss shortly. Again, in each of these experiments the players were free to purchase additional edges that would be dynamically added to the seed network.

Again somewhat surprisingly, none of these seed network structures seemed to improve collective performance much, with the completion rates on seed torus experiments being 33% (6/18), and on seed clique experiments being 33% (9/27) — neither of which is significantly different from the unseeded networks of Session 1, but both of which are significantly lower for the previous fixed-network biased voting networks ($P < 0.01$). Whatever network the subjects were given to start, they seem to have turned it into a poor network for the task.

A major finding of our original biased voting experiments focused on experiments in which networks were generated according to preferential attachment, which results in a heavy-tailed degree distribution, and where we gave a (sometimes very small) minority of the players a higher payoff for red, with the majority preferring blue. The twist was that the red minority consisted of the highest-degree vertices in the network; we were thus investigating whether a small but well-connected minority could systematically impose its preference against the majority's. The answer was resoundingly positive: 24 of 27 (89%) such fixed-network experiments ended in consensus, every one of them on the minority preference [Kearns et al. 2009].

In the current experiments, we were interested in how this finding might be changed if subjects could add edges to the minority power networks. We thus ran a number of experiments in which both the seed network structure and arrangement of incentives were identical to those in the earlier minority power experiments. The success rate was 61% (22/36) — higher than for torus or clique seeds, but still much lower than the original exogenous network minority power rate of 89%. Once again, permitting the subjects to modify the network has harmed collective performance. However, now 35% of the successful experiments ended with the *majority* preference — compared with none in the exogenous network case, a dramatic change. One interpretation is that permitting the purchase of edges allows the majority players to better realize they are in the majority — which may have been difficult in the exogenous network case, especially for the preponderance of low-degree vertices — and causes them acquiesce to the minority less readily. This could account both for the lower overall success rate, and the increased rate of majority victories.

4.3. Effects of Edge Costs and Incentives

Recall that across the many experiments, we varied the cost per edge purchase, and the absolute and relative incentives across the population. Edge costs were low ($0.01), medium ($0.10), or high ($0.25), with the proviso that the edge purchases of a player must always be strictly less than $1 (which was always the payoff of the less preferred color).

Not surprisingly, the cost per edge had a strong effect on the resulting network density (as we shall see in the following section), but also on collective performance. The overall success rates for unseeded Session 1 experiments were 67% for low-cost experiments, 38% for medium-cost experiments, and 14% for high-cost experiments; the differences between these quantities are all significant at $P < 0.05$. We note that although there was a clear relationship between edge costs and performance, with higher costs resulting in worse performance, in no case did the subjects collectively approach the maximum allowed edge expenditures; the fraction of possible edge purchases in unseeded Session 1 experiments was 64% for low edge cost, 42% for medium, and 59% for high. Thus subjects could have built considerably denser networks in all cases, but chose not to. We also note that the seeded experiments provided the subjects with considerably denser networks for less edge expenditures, yet failed to significantly improve performance.

The incentives given to players also had a pronounced effect on subject performance. Recall that in each experiment a subject always desired unanimity (no payment was distributed unless unanimity was reached), but had a preference for one color over another. For example, a subject might receive 4 if all chose blue, and only 1 if the consensus was to red. We maintained the smaller incentive at 1 for all experiments, and varied the payoff of the preferred color, from 4 (strong preferences), 2 (weak preferences), and 1 (indifference between the color choices). The overall success rates for Session 1 experiments were 58% (7/12) when the subjects were indifferent between the two colors, 53% (19/30) when they had a weak preference for one of the colors, and only 17% (4/24) when they had a strong color preference. While there was no statistically detectable difference between the impact of weak and no preference, strong preference had a clear detrimental effect on solution rate ($P < 0.01$, comparing against weak preference).

4.4. Network Structure and Centrality Skewness

Both as a general matter, but especially in light of the overall poor performance of the subjects, the structure of the networks built during the experiments is a topic of central interest. How do the built networks compare to more naturally evolved social networks, and what properties of the built networks might account for the difficulty they posed for the biased voting problem? Here we initially restrict our analysis to the majority of experiments where there was no seed network, so that all structure was built by the subjects themselves.

We begin by establishing that the built networks actually do share a number of structural "universals" that appear frequently in real-world networks. The first remark worth making is that in every unseeded experiment, the subjects built a connected network — there were never two or more disconnected components. Thus regardless of the edge costs and other parameters, the subjects always bought enough edges to establish global communication.

The diameters (pairwise average shortest path distances) of the built networks were generally quite small compared to the population size; while the diameters show a strong dependence on the network density and therefore the edge costs, they averaged 1.32 for all low-cost experiments (standard deviation 0.17), 1.87 (standard deviation

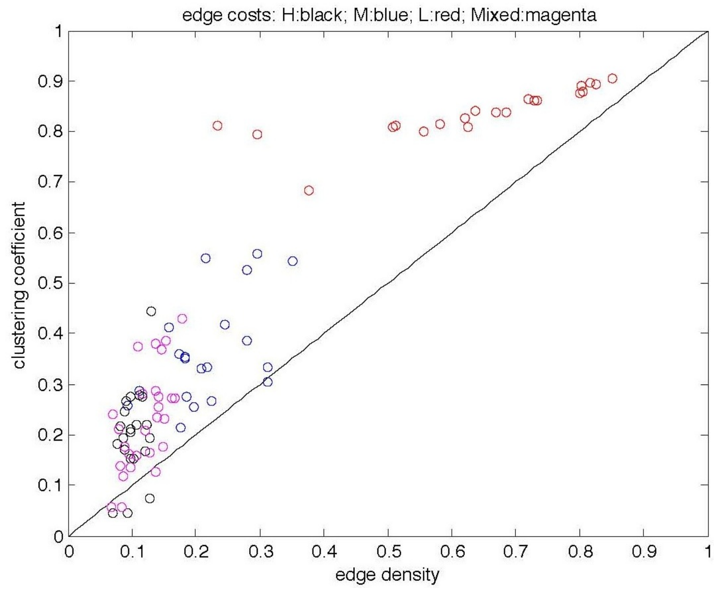

Fig. 3. Clustering coefficients vs. edge density for networks built in all unseeded experiments. For each built network, the x value indicates the fraction of possible edges present, and the y value indicates the clustering coefficient of the network. Random networks of the same densities would have clustering coefficients equal to their density, as suggested by the diagonal line. We see that clustering in the built networks is uniformly higher except at very low densities, where presumably the subjects may be been primarily concerned with establishing global connectivity. There are strong effects of edge costs (coded by color); higher edge costs consistently lead to lower densities and clustering. The mixed experiments had variable edge costs for the players, but always either medium or high.

0.19) in medium-cost experiments, and 2.38 (standard deviation 0.16) in high-cost experiments. Also as is for typical social networks, the clustering coefficients of the built networks were generally much higher than for random networks of the same density; see Figure 3. Furthermore, examination of the degree distributions of the built networks reveals the presence of "connector" vertices whose degrees are several times larger than the mean, another commonly cited property of natural social networks.

Given that the structural properties cited so far were generally present in our earlier, fixed-network experiments — in which the subjects performed much better — what can account for the difficulty of the built networks? While it is impossible to be certain, due to the complexities of both the networks and subject behavior, a strong candidate is the overreliance on very few vertices or subjects for connectivity and communication. In Figure 4 we demonstrate this reliance visually by showing, for three of the built networks, the effects on structure and connectivity of deleting a few vertices with the highest degrees. In each case, the network quickly becomes highly fragmented, a property generally true of the built networks.

We can make this analysis more systematic and rigorous by considering the quantity known as *betweenness centrality* (centrality in the sequel). Centrality is designed to measure, for each vertex u, the extent to which the rest of the network relies on u for its global connectivity and communication. More formally, we define

$$C_B(u) = \sum_{v,w \in V : v \neq u, w \neq u, v \neq w} \frac{n^u_{v,w}}{n_{v,w}}$$

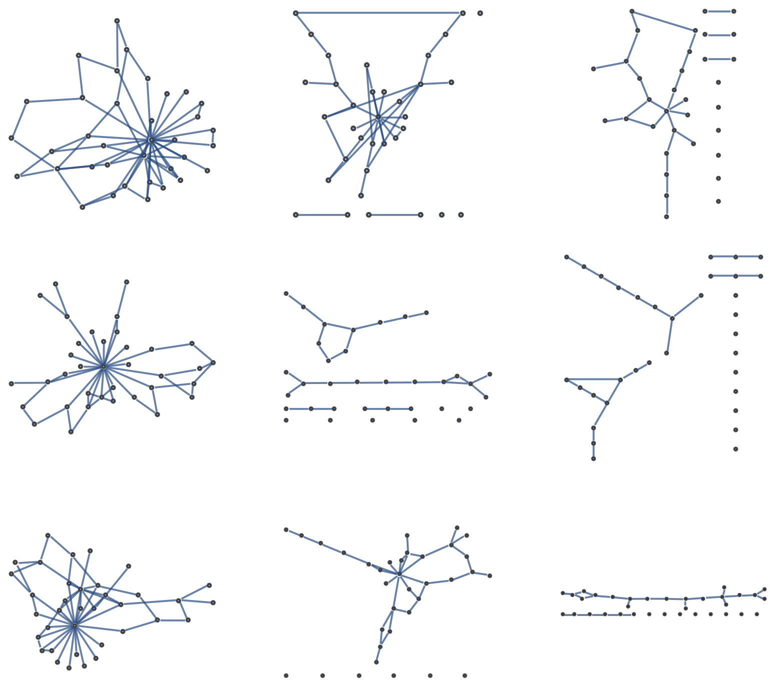

Fig. 4. Visualization of built networks in three unseeded Session 1 experiments with high edge costs. The first column shows a visualization of each of the built networks. Subsequent columns show the networks after repeated deletions of the highest-degree vertex remaining. In each case, the first deletion already shatters the network into multiple connected components, and subsequent deletions yield a large number of isolated vertices.

where V is the set of all vertices, $n_{v,w}$ is the number of shortest paths between v and w, and $n^u_{v,w}$ is the number of shortest paths between v and w that pass through vertex u. Thus $C_B(u)$ is a global measure of how often vertex u appears on shortest paths between all pairs of other vertices; it is a common metric of influence on communication and connectivity in social networks.

Echoing the analyses above, it turns out that the subject-built networks systematically differ from naturally occurring networks, and the ones we imposed on subjects in our earlier experiments, in their distribution of centrality. In particular, as with degrees, the built networks display a considerably more *skewed distribution* of $C_B(u)$: compared to natural network models at the same edge density, there are more vertices with very high and very low C_B , and fewer with intermediate values of C_B. See Figure 5.

There are a number of obvious reasons why overreliance on a few high-centrality vertices might make the biased voting task difficult. If a large fraction of the population implicitly relies on high centrality vertices to be effective aggregators of global information (such as the current majority color), noisy or selfish behavior by these individuals can impede collective performance. In the successful Session 1 experiments, the correlation between centrality and whether a subject received their higher payoff was both positive (0.18) and highly significant ($P < 0.001$), suggesting that high-centrality players may have implicitly used their position to influence outcome rather

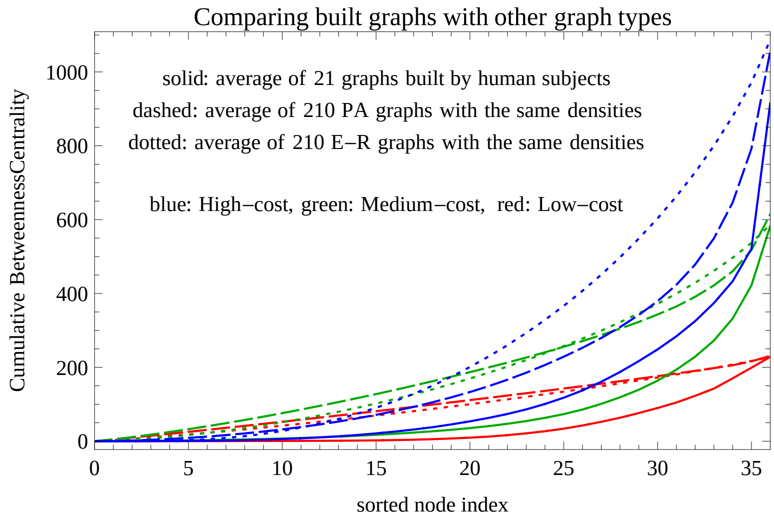

Fig. 5. Comparison of the distribution of centrality C_B between subject-built networks and standard generative models. For a given network, we sort vertices in order of increasing C_B values, and then compute the sum of all C_B values through a given rank in the ordering. We then average such curves over many built networks or many sampled networks from the generative models. We group the averaged curves by edge costs (color coded in the figure) for the built networks, and compare them to Erdös-Renyi and preferential attachment networks of the same average density. At each edge cost, we see that the built networks (solid lines) have much more skewed distributions of C_B than Erdös-Renyi (dotted lines) and preferential attachment (dashed lines): while the sum of all C_B values (rightmost point) is comparable for all three classes of networks at each density, much more of the cumulative centrality is accounted for by the final few, most central, vertices in the built networks, whose curves are considerably more convex than for the models.

than coordinate behavior — potentially contributing to the great majority of failed games.

We note that one might be tempted to think that the starting seeds would dissipate the propensity for skewness in the C_B values. For instance, in the torus networks, there already are 4 edges uniformly assigned to each node, so one could imagine new edges would have less of a biasing effect than when starting from an empty network. However, this intuition is misleading. When we examined the C_B values in networks that were seeded, we found final values that were both higher and lower than the starting values. The variance in people's buying behavior injected a variance into the C_B values, and they were spread out in both directions.

4.5. Purchasing Behavior and Free Riding

Thus far we have offered evidence that subjects built poor networks for the given task. It is natural to ask what particulars of subject behavior accounted for this. In this section, we examine the distribution of edge purchases in the population, which sheds some light on this question. Here the most striking aspect is the preponderance of free-riding: at all edge costs and in each experiment, there is a significant fraction (roughly 20% or more) of players who purchase *no* edges, and another large group who purchase very little compared to the average. Thus the vast majority of the cost of building the networks was undertaken by only a small fraction of the population. This variance in behavior is what we believe generated the aforementioned variance in C_B values.

For low and medium edge costs, fully 50% of the population contributed less than 10% of the total edge expenditures; at high edge costs, where no player could afford to purchase more than 3 edges and thus variability of expenditure should be reduced, the free riding 50% still purchase less than 20% of the edges. Furthermore, free riding was an economically beneficial policy: at the level of individual human subjects, the correlation between the number of edges purchased in all Session 1 experiments and their total payoff is -0.72 ($P < 0.001$). The primary builders of the networks were thus apparently not financially favored by their resulting positions in it. Networks tended to be built rapidly, with the vast majority of edge purchases coming in the first half of the allowed time.

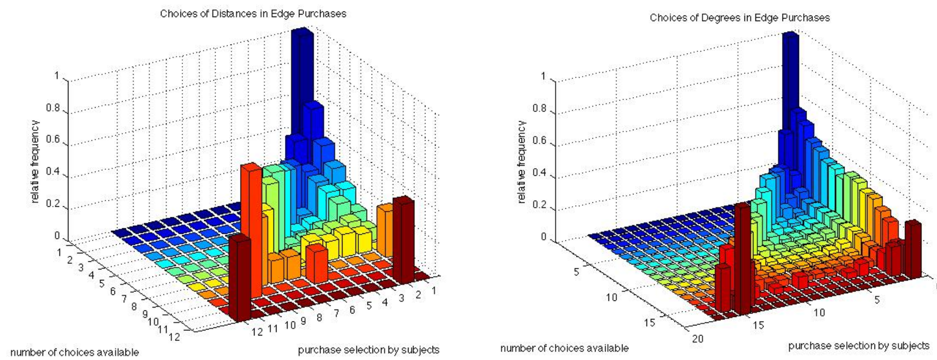

Fig. 6. Frequencies of edge-purchasing decisions with respect to non-neighbor vertex distances (left panel) and degrees (right panel). For each edge purchase, the left axis represents how many distinct choices the purchaser had, and the right axis represents which of these ranked options they selected. The vertical axis then shows the relative frequency they made each ranked choice. Thus the diagonals indicate cases where they purchased an edge to the most distant, or highest-degree, non-neighbor, respectively. We thus see that while there is a tendency to purchase edges to the most distant and highest degree vertices, there is also considerable mass at low and intermediate distances and degrees. Color is for visualization clarity only.

So far our emphasis has been on the structural properties of the built networks and the distribution of expenditures; we next examine the criteria subjects seemed to use in edge-purchasing decisions, within the constraints of the degree and distance information they were provided about current non-neighbors. Normalization is an issue here since (for instance) what constitutes a relatively "high degree" vertex is different near the start of an experiment than towards the end. Instead we can simply ask, at each moment an edge purchase was made, how many choices the purchaser had in each dimension, and which one they made: that is, how many different degree values, and how many different distance values, were populated by at least one non-neighbor on the edge-purchasing GUI grid. The results are summarized in Figure 6, and they show that while subjects most often chose to buy edges to vertices with the highest available distance, or the largest available degree, they also frequently chose deliberately low values in both dimensions as well, and there is significant mass on intermediate values as well. Thus purchasing behavior is not easily consistent with standard generative models such as Erdös-Renyi (which would induce uniform distributions in both dimensions) or preferential attachment (which would not generate the observed tendency to connect to low-degree vertices). Despite the simplicity of the informational interface, subject purchasing behavior does not easily fall into simple models.

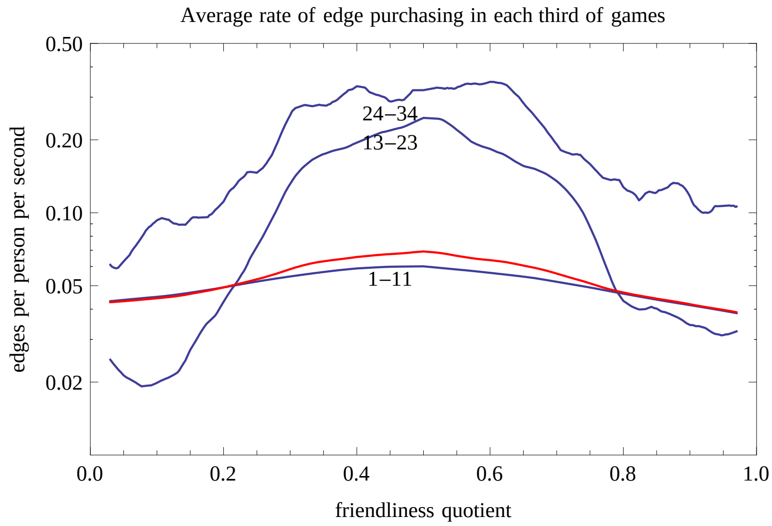

Fig. 7. Edge purchasing rate (edges purchased per subject per second) as a function of the friendliness quotient (see text) of the purchaser at the moment of purchase. The red curve shows the aggregate across all Session 1 and 2 purchases, and the peak near 0.5 is 60 or 70% higher than at the edges. If we condition on the purchaser also having a current degree in some range (blue curves annotated by degree range), we see that this tendency to purchase more when there is local indecision and conflict becomes greatly pronounced at higher degrees. The two higher degree curves are 5 to 10 times higher in the middle than at the sides.

While the preceding analysis examines how subjects used degree and distance in edge purchasing decisions, it is also of interest to investigate how such decisions were influenced by the local state of play — in particular, whether most neighbors were playing the subject's preferred (higher payoff) color or not. Figure 7 shows the rate at which subjects purchased edges in Session 1 and 2 experiments as a function of the "friendliness quotient" of their neighborhood at the moment of purchase. The friendliness quotient is the fraction of current neighbors who are playing the purchaser's higher-payoff color. We see that there is a marked increase in the proclivity of a player to buy an edge if she finds herself with an approximately equal number of friends and enemies (friendliness quotient 0.5). In situations where her neighbors are largely colored the same (whether of the higher or lower payoff color), she refrains from buying. This tendency becomes even more pronounced if we condition on just those purchases made by players whose current degree is higher. It may be that this tendency to buy more edges when there is a large amount of local disagreement (high degree, friendliness quotient close to 0.5) only worsened the indecision for everyone, thus leading to poor convergence performance.

5. CONCLUDING REMARKS

The results presented here have shown that human subjects, given the opportunity and incentives to collectively build a network in service of a competitive coordination task, did so poorly, creating networks inherently difficult for the task. This occurred despite an edge purchasing interface that permitted, in principle, the creation of networks known from earlier experiments to be much easier, such as random or preferential attachment networks. Our findings are in contrast to some case studies and theories about the abilities of human populations to effectively self-organize in service of collective goals. This contrast clearly deserves further scrutiny and controlled study.

REFERENCES

ALBERS, S., EILTS, S., EVEN-DAR, E., MANSOUR, Y., AND RODDITY, L. 2006. On Nash equilibria for a network creation game. In *Symposium on Discrete Algorithms (SODA)*.

BARABASI, L. AND ALBERT, R. 1999. Emergence of scaling in random networks. *Science 286*, 509–512.

BOLLABAS, B. 2001. *Random Graphs*. Cambridge University Press.

BORGS, C., CHAYES, J., DING, J., AND LUCIER, B. 2011. The hitchhiker's guide to affiliation networks: A game-theoretic approach. In *Innovations in Theoretical Computer Science (ITCS)*.

BRAUTBAR, M. AND KEARNS, M. 2011. A clustering coefficient network formation game. In *Symposium on Algorithmic Game Theory (SAGT)*.

BURNS, T. AND STALKER, G. 1994. *The Management of Innovation*. Oxford University Press.

CHAKRABORTY, T., JUDD, J. S., KEARNS, M., AND TAN, J. 2010. A behavioral study of bargaining in social networks. In *ACM Conference on Electronic Commerce*. 243–252.

EVEN-DAR, E. AND KEARNS, M. 2006. A small world threshold for economic network formation. In *Neural Information Processing Systems (NIPS)*.

EVEN-DAR, E., KEARNS, M., AND SURI, S. 2007. A network formation game for bipartite exchange economies. In *Symposium on Discrete Algorithms (SODA)*.

FABRIKANT, A., LUTHRA, A., MANEVA, E., PAPADIMITRIOU, C., AND SHENKER, S. 2003. On a network creation game. In *Principles of Distributed Computing (PODC)*.

JACKSON, M. O. 2005. A survey of models of network formation: stability and efficiency. In *Group Formation in Economics: Networks, Clubs and Coalitions*. Cambridge University Press.

JACKSON, M. O. 2010. *Social and Economic Networks*. Princeton University Press.

JUDD, S., KEARNS, M., AND VOROBEYCHIK, Y. 2010. Behavioral dynamics and influence in networked coloring and consensus. *Proceedings of the National Academy of Sciences 107,* 34, 14978–14982.

KEARNS, M. AND JUDD, J. S. 2008. Behavioral experiments in networked trade. In *ACM Conference on Electronic Commerce*. 150–159.

KEARNS, M., JUDD, S., TAN, J., AND WORTMAN, J. 2009. Behavioral experiments in biased voting in networks. *Proceedings of the National Academy of Sciences 106,* 5, 1347–1352.

KEARNS, M., JUDD, S., AND VOROBEYCHIK, Y. 2011. Behavioral conflict and fairness in social networks. In *Workshop on Internet and Network Economics*.

KEARNS, M., SURI, S., AND MONTFORT, N. 2006. An experimental study of the coloring problem on human subject networks. *Science 313*, 824–827.

KLEINBERG, J. 2000. Navigation in a small world. *Nature 406*, 845.

NISHIGUCHI, T. AND BEAUDET, A. 2000. Fractal design: Self-organizing links in supply chain management. In *Knowledge Creation: A Source of Value*, G. von Krogh et. al., Ed. St. Martin's Press, 199–230.

TARDOS, E. AND WEXLER, T. 2007. Network formation games and the potential function method. In *Algorithmic Game Theory*. Cambridge University Press, 487–513.

WATTS, D. AND STROGATZ, S. 1998. Collective dynamics of small-world networks. *Nature 393*, 409–410.