Endogenous Communication Networks with Boundedly Rational Agents*

S. Angus†

Feb 2006

Abstract

This paper concentrates on non-cooperative communication network formation processes as introduced analytically by Bala and Goyal (BG, 2000), and subsequently investigated experimentally by Falk and Kosfeld's (FK, 2003) faithful human implementation. Key to BG's model was the rationality of agents, whereas FK's work emphasised the break-down of this assumption in certain cases. In the present work, the perfect rationality assumption of BG is relaxed, and instead, agents are implemented who possess a variety of 'abilities' at the outset, but who are also able to learn from each other over time. Furthermore, since each agent's decision calculus is completely transparent to the inquirer, the present framework provides the ability to test the FK hypthotheses, whilst also permitting the observation of extended concepts such as 'complexity', 'diversity' and 'cooperation'.

I find that traditional learning of payoff optimization is a poor proxy for human decisions-making, with an alternate heuristic based on a combined strategic efficiency and agent-to-agent reciprocity measure providing a more realistic rendering of human decision outcomes. Moreover, the supposed desire for payoff equality and strategic inertia does not appear causally related to improved agent performance, but rather seems to be an emergent property of the learning system.
JEL codes: C72, C73, D82, D83, D85

1 Introduction

Communication networks are an apparently ubiquitous feature of many business and inter-personal contexts. In each, depending on the costs of information access and benefits of information content, entities (e.g. firms,

*My thanks to the members of the Theory Workshop for helpful discussions and feedback on this work, and to the Australian Government's APA scholarship scheme for the generous support that made the work possible.

†*School of Economics, John Goodsell Building, The University of New South Wales, 2052, NSW, Australia.; e-mail:* `s.angus@unsw.edu.au`.

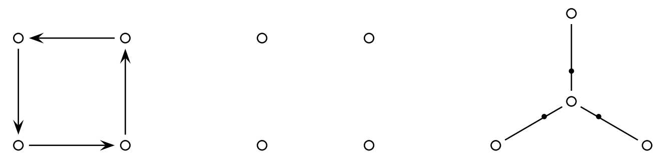


Figure 1. Example Strict Nash equilibrium networks, as in BG2000: (left) Circle (or 'wheel'); (middle) Empty; and (right) Centre-sponsored Star (two-way information flow; smaller dots near centre node indicate sponsorship cost of centre node).

individuals) face a strategic problem of who to engage in mutual or unilateral information sharing partnerships given the actions of other entities that together comprise the information 'landscape'. Although several communication network models have been reported under varying specifications[1] one influential model of *non-cooperative* communication network formation is that of Bala and Goyal (2000)[2] who identified *equilibrium* information network structures under various treatments, including cost of edge sponsorship and direction of subsequent information content flows. Namely, they identify *minimally connected* Nash equilibrium structures in both one- and two- way specifications and using the Strict Nash equilibrium refinement[3] obtained specific equilibrium structures, namely: *empty*, *circle*[4] and *centre-sponsored star*, under given treatments (see Fig. 1).

In their model, agents are assumed to be able to observe all previous period network sponsorship decisions of their opponents, i.e. they observe G_{t-1}, and when given the opportunity to update their strategy, choose the (myopic) *best-response* to G_{t-1}. Convergence is guaranteed by in-built strategic *inertia*, at least one agent plays the sponsorship strategy of period $t-1$ in period t, a process that searches the space in an incremental fashion, eventually arriving at one of the Nash structures.

Whilst promising, subsequent human trials of the BG2000 noncooperative network formation set-up ($n = 4$), conducted by (Falk and Kosfeld, 2003)[5] find that the BG2000 results receive mixed support in the field. Specifically, in the one-way information flow case, under both low and high costs of edge sponsorship, agents do discover Nash outcomes a majority of the time; especially so towards the end of each treatment where order ef-

[1] See for example, Chwe (1995, 2000); Comellas et al. (2000); Jackson and Watts (2002); Slikker and van den Nouweland (2000).

[2] To be referred to as BG2000 hereafter.

[3] In this case, the Strict Nash refinement implies that not only will each player play a utility maximising strategy in response to each other player's utility maximizing strategy, but each player will only have one such strategy.

[4] That is, arrange all nodes on a circle, connect one to another in a 'daisy-chain' procedure such that each agent's in- and out- degree is unity. NB: BG2000 actually refer to this case as the 'wheel', we shall use the terms interchangeably.

[5] Subsequently referred to as FK2003.

fects are clearly evident. However, in the two-way information case, subjects performed remarkably poorly with respect to finding minimally-connected networks (Nash), barely finding the Strict Nash (centre-sponsored star) outcome *at all*. Such evidence demands further explanation.

In their discussion, FK2003 point to various problems for subjects with the two-way Strict Nash equilibrium structure including the serious assymmetry in both a strategic and a payoff sense, conjecturing that agents dislike such conformations. Further, by use of regression analysis, FK2003 show a strong explanatory correlation between *inertia* (as explained above) and experienced payoff equality in the previous period (controlling for previous best-response play) to support their claim. That is, if other agent payoffs are roughly similar (either low or high) in the previous period, then a subject is more likely to exhibit strategic inertia in the present period.

Such considerations will clearly require a review of the Best Response decision rule to model these contexts. Indeed any form of boundedly rational behaviour due to irrational preferences or other, is not well catered for by the standard best-response decision making rule. It is quite likely that agents are employing (consciously or not) a diverse set of heuristics to determine their strategic play. The problem, of course, is to determine which rules to include in the model. Whilst, for instance, the insightful work of Matros (2004) finds that for generic n-player games, so long as all players are using decision rules from within the union-set of *weakly rational* rules[6] and the best-response rule, then in the short-run the outcomes are 'identical' to that if individuals were constrained to just the best-response rule, giving implicit support to the BG2000 frame-work, it is not clear from FK2003 that such conditions do prevail in reality. We may conclude that individuals draw their decision-making rules from outside of such 'minimal curb sets',[7] at least in the present example.

Moreover, the FK2003 results show that agents undergo clear learning processes as they play the game, with performances in the one-way information case improving appreciably between mixing stages (see below). Hence, the challenge of such economic modelling is two-fold: first, to model a diverse range of feasible decision-making heuristics for each agent, allowing for varying apparent 'abilities' and initial insights; and second, to model some kind of learning process, such that selection between such heuristics occurs in a realistic manner.

In the present paper, such considerations are dealt with in a novel way. First, the enormous complexity of decision-making in the network formation

[6] Informally, a set of simple rules is weakly rational if the application of these rules on a sub-set of states (a *minimal curb configuration*) will return one of the member states of this sub-set; the rules are then "consistent with an equilibrium".

[7] Compare the suggested decision-making influences above – those of symmetry and equity.

game described by BG2000,[8] is reduced to a tractable (and implementable) human-decision making process by allowing individuals to recognise and respond to equivalent structures (under relabelling). Second, agent cognition is dealt with by constructing a managable series of response-rules to each structure under a common reference frame-work but without biasing the relative 'intelligence' of any individual agent. And third, in consequence and in the spirit of Arthur's (1994) artifical bar attendees, since agent cognition architecture is common to all agents, between-agent learning and experimentation or mistake-making is afforded and directly observable.

The results of the current approach can be summarised as follows: first, pure payoffs (e.g. monetary, as in FK2003) appears to be an unsuccessful candidate for the objective criterion used by agents to judge 'good' plays by their opponents; second, and in the place of payoffs, a combination of a kind of 'efficiency' measure (to be explained below), and a strategic cooperation or reciprocity measure gives rise to comparable outcomes to that of the experimental data; and third, although equality and best response considerations are found to be significantly associated with strategic inertia as found by FK2003, these effects appear to be *emergent* in nature, rather than being the driving dynamic in the learning environment as they suppose. Indeed, by explicitly incorporating either into the objective learning measure as may be thought reasonable, agents perform *worse* over time than without them.

The rest the paper includes a summary of the pertinent results from both BG2000 and FK2003 (§2), a description of the model (§3) and its implementation (§4), which is followed by computational modelling results and analysis (§5), before several concluding comments are drawn (§6).

2 Background

2.1 BG2000 Model Predictions

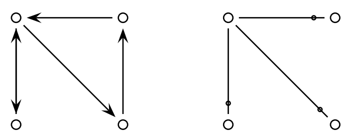
The key predictions from the BG2000 model are summarised in Table 1. The circle, wheel and centre-sponsored star have already been introduced, leaving the so-called 'minimally-connected' networks to explain. In each, all agents obtain full information, in the sense that they can observe the information of each other agent. However, for the one-way case the minimally-connected property implies that each edge is *necessary*, that is, the removal of any edge will cause some agent(s) to lose their access to the other $(n-1)$ agents' information. In the two-way case, such a criterion results in cycles being ruled out (one edge would be unnecessary), plus any case where two agents mutually sponsor a direct link to each other (again, one would be unnecessary). Such is the nature of two-way information flows. Examples of each type of minimally-connected structure can be seen in Fig. 2.

[8]For example, consider the $n-4$ case under one-way information flows. The number of distinct feasible graphs is given by $2^{n(n-1)}$, which is more than 4000 for this case.

Table 1. Predicted structures by BG2000 under each treatment.

Flow	Edge Costs[a]	Structure[b]				
		m1c	circle	empty	m2c	cs-star
One-way	Low	△	▲*			
	High	△	▲*	▲		
Two-way	Low				△*	▲
	High			▲	△*	

Notes: [a] Low $C \leq V$, High $C > V$; [b] structure *m1c* and *m2c* are minimally-connected non-empty graphs in one- and two- way information flow cases respectively. (△) non-empty nash, (▲) strict nash, (*) indicates that the structure is also *efficient* (following FK2003).



(a) m1c **(b)** m2c

Figure 2. Examples of minimally-connected structures (following FK2003): (a) the one-way information flow case; and (b) the two-way case (periphery-sponsored star).

2.2 FK2003 Experimental Findings

The authors of the FK2003 study reproduced the BG2000 network formation context in an extremely faithful manner. For a full description of their study, the reader is referred to the reference. However, to summarise the procedural details, a total of 160 subjects, 32 in each of 5 treatments – 3 under one-way information flows (costs 5, 15 and 25) and 2 under two-way flows (costs 5 and 15) – were randomly grouped into 8 mixing groups of 4 to play 5 rounds of the network formation game (we shall refer to this process as a 'stage'). After a stage, the 32 subjects would be shuffled and re-assigned to 8 groups of 4 for a further five rounds. In all, 3 stages were conducted for each treatment. Subjects interacted with a computer screen (after familiarisation) to both observe the sponsorship decisions of their opponents in the preceding round, and to enter their own sponsorship decisions in the current round. Subjects received 10 Swiss Francs for showing-up, and received monetary reward for theiry play; information observation of one node gained 10 points with all costs (as just mentioned) measured in points with 10 points representing 0.9 Swiss Francs (subjects received around 50 SF on average).

Several FK2003 results are relevant. First, under the four treatments over information flow direction and edge sponsorship costs, the proportions of each graph-type as predicted by the BG2000 results are striking (see

Table 2). Under one-way information flow, aggregated accross all (three) learning stages and all (five) rounds within each stage, a strong tendency for playing Nash is revealed. Moreover, in a large proportion of cases where a Nash structure was played, the Strict Nash structure (either just the circle-graph, or the circle and the empty graph type) were played. This result is dashed however in the case of two-way flows, where Ntrict Nash play was non-existent and the Nash structures that did result were played in a significantly lower number of rounds relative to the one-way case, with low edge costs yielding a Nash frequency of around 30% (as opposed to around 50% for one-way), reducing to less than 10% when edge costs were high. Interestingly, average agent degree reveals a roughly constant pattern between information flow regimes, indicating that *strategic*, rather than purely edge sponsorship propensity is the likely cause of failure when flows are two-way.

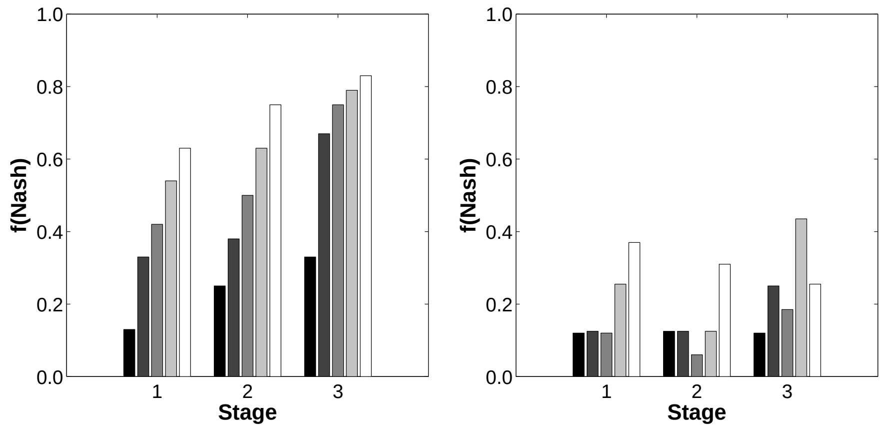
Table 2. Summary of selected FK2003 experimental results under relevant treatments. (Compare Table 1.)

Flow	Edge Costs[a]	Structure					$\langle d_i \rangle$
		m1c[b]	circle	empty	m2c[b]	cs-star	
One-way	Low (5)	0.48	0.41				1.19
	High (25)	0.59	0.49	0.10			0.76
Two-way	Low (5)				0.31	0.00	0.91
	High (15)			(nr)	0.09		0.75

Notes: [a] figures indicated in parenthesise are in 'experimental points' – 10 points equated to 0.9 Swiss Francs (about $US 0.59) and value of one agent's information constant at 10 points; [b] minimally connected values include other structures (such as Strict Nash as appropriate); (nr) indicates 'not reported'.

Proportions of Nash play both between and within each learning stage further reveals the problems that subjects had with the two-way information regime (see Fig. 3).

As can be seen, in the one-way case subjects underwent a strong learning dynamic between learning stages, and within stages achieved a similar improvement. Within stage improvements, or alternatively, low frequencies in the early rounds of each stage, are attributed to mis-coordinations due to the mixing of subject grouops that occurs between learning stages. In the two-way case, the lack of Nash structures in comparison is stark with both the between-stage and within stage improvements not nearly as evident.



(a) One-way **(b)** Two-way

Figure 3. FK2003 frequency of Nash graph structures under each information flow regime. Bars represent average frequency over edge costs 5, 15 and 25 in the one-way case, and over edge costs 5 and 15 in the two-way case. NB: refer to Table 1 for graph structures that are 'Nash' under each information and cost scenario.

3 Model

3.1 The Network Formation Game

Suppose a group of agents $N = \{1, \ldots, n\}$ are selected to play a communication network formation game.[9]Let the communication network be comprised of vertices as given by N, and let there be edges as given by the set of edge-sponsorship actions[10]$g \in \mathbf{g}^n$, $g = \{g_1, \ldots, g_n\}$, where $g_i = (g_{i,1}, \ldots, g_{i,i-1}, g_{i,i+1}, \ldots, g_{i,n-1})$ is an ordered (row) vector of pair-wise sponsorship decisions, and $g_{i,j} \in \{0, 1\}$, $\forall\ j \in N/\{i\}$. Correspondingly, let the communication network be represented by the graph $G(N, g) \in \mathbf{G}$.

For example, suppose $N = \{1, \ldots, 4\}$, and g' has the form,

$$\begin{aligned} g'_1 &= \{1, 0, 0\} \ , \\ g'_2 &= \{0, 1, 0\} \ , \\ g'_3 &= \{0, 0, 1\} \ , \\ g'_4 &= \{1, 0, 0\} \ , \end{aligned}$$

then, under one-way information flows, the corresponding graph $G'(N, g')$ would be a cycle, as given in Fig. 4.

[9]Nomenclature in this section largely follows that of BG2000 where possible for consistency.

[10]NB: we differ here to the description of BG2000 since g_i gives the *outcome* of an agent's strategy (that is, an action to be taken), rather than comprising the strategy itself, as described by BG2000.

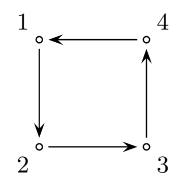


Figure 4. Example graph G'.

The communication network confers benefits on agents by gaining them access to the information of other agents. In the terminology of BG2000, each agent is able to 'observe' the information of an agent they sponsor a link to. Importantly, however, indirect 'flows' of information are allowed. Hence, if $(i \rightarrow j) \equiv (g_{i,j} = 1)$ then the graph where $(a \rightarrow b)$ *and* $(b \rightarrow c)$ is true implies that a can not only observe herself (by convention), they can also observe b (by direct sponsorship) *and* c (by indirect information flow). In the linear benefits specification of BG2000, followed in both FK2003 and in the present paper, let each agent's information be of value $V \in \mathbb{R}^+$, and let $\{\mu_i(G) \in \mathbb{N} \mid \mu_i > 0\}$ be the number of agents i observes, and further let $C \in \mathbb{R}^+$ be the cost associated with sponsoring one edge, and $\{\delta_i(G) \in \mathbb{Z}^+\}$ be the count of edges that i sponsors (that is, the *degree* of vertex i). Then define each agent's payoff function $\pi : \mathbf{G} \mapsto \mathbb{R}$ as

$$\pi_i(G) = \mu_i(G)V - \delta_i(G)C \ . \tag{1}$$

Further, let a *strategy* (decision-making rule) for an agent i in some stage $t \in \{1, \ldots, T\}$ be a mapping $\mathcal{S}_i^t : \mathbf{G} \mapsto \mathbf{g}$, and denote the set of all such rules $\mathbf{S}$. Within such a stage, each agent will take part in a number of rounds, enumerated by $r \in \{1, \ldots, R\}$. And hence, an agent's strategy $\mathcal{S}_i^t$ will be current for all R rounds of some stage t. Indeed, we can define a communication graph in period r as $G^r(N, \{g_1^r, \ldots, g_n^r\})$, since it will be constructed from each agent's link sponsorship decisions, g^r.

The timing of the game is simple in nature. In the first stage, each agent simultaneously reveals their initial sponsorship decisions to form the first round graph $G^1(N, g^1)$. Given the resultant communication network, payoffs are awarded to each agent as in (1). After observing the structure of the graph, each agent then applies their decision-making rule, $S_i^1 : G^1 \mapsto g_i^2$ and then reveals their new sponsorship decisions which forms the new round two communication graph $G^2(N, g^2)$, with payoffs again being awarded accordingly, and so on. This process continues until R rounds have been played, at which point, agents enter a new stage (stage 2) and the process is repeated in full until stage T has been completed. Importantly, however, whilst the retention of an edge sponsorship between rounds is costly (incurring C), the severring of an edge between rounds is not (apart from the potential loss of any information value that will result from the break).

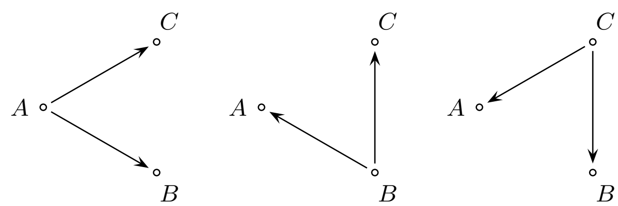


Figure 5. A 3-member graph equivalence set.

3.2 The Decision-making Rule **S**

In the present work,[11] $\mathbf{S}$ is implemented so as to permit full inspection of all decision-making processes at any time, and to model possible learning, and/or, decision-making hypotheses. Hence, let $\mathbf{S}$ be of the form (for some agent i), $\mathcal{S}_i = \{s(\mathcal{T}_1), \ldots, s(\mathcal{T}_k)\}$ where $\mathbf{T}(n) = \{\mathcal{T}_1, \ldots, \mathcal{T}_k\}$ is the set of all *minimal absentee graphs (types)* for given n (to be explained below), and $s(\cdot) : \mathbf{T} \mapsto \mathbf{g}$ – the engine room of the decision-making process.

Suppose $\mathcal{G}$ is the set of all n-person graphs, then $\mathbf{T} \subset \mathcal{G}$ is the set of all non-equivalent graphs in $\mathcal{G}$. Infact, all n-person graphs can be constructed from $\mathbf{T}$ under a simple re-labelling of nodes. That is, $\mathbf{T}$ represents the *minimal set of non-equivalent structures*, such that if we define a relabelling operator $\mathfrak{R} : G(N, g) \mapsto G''(N'', g)$ where N'' is simply a re-ordering of the elements of N, then

$$\{\mathbf{T} \mid \mathfrak{R} : \mathbf{T} \mapsto \overline{\mathbf{T}}\} \ , \tag{3}$$

or in other terms,

$$\{\mathfrak{R}^x(G^a) = G^b, \ \forall \ G^a \in \overline{\mathbf{T}}\} \ , \tag{4}$$

where $\overline{\mathbf{T}}$ indicates the compliment of $\mathbf{T}$, $G^b \in \mathbf{T}$, and $\mathfrak{R}^x$ implies an application of $\mathfrak{R}(\cdot)$ a finite number $x \in \mathbb{N}$ times. Taken together, (3) means that any relabelling operation applied to some $\mathcal{T} \in \mathbf{T}$ must result in an equivalent, and therefore, un-represented graph in $\mathbf{T}$, whilst (4) implies that any graph in $\mathbf{T}$ can be reached (exactly) via a finite series of re-labelling operations to some graph in $\mathcal{G}$ (but not already in $\mathbf{T}$).

For example, Fig. 5 is the complete set of equivalent (under-relabelling) 3-node graphs of the form $\left\{G(N^3, \tilde{g}) \mid \tilde{g} = \{(0,1),(0,1),(0,0)\}\right\}$. Consequently, the set $\mathbf{T}$ ($n = 3$) must contain exactly one of these (the choice is arbitary).

The importance of the minimal graph set $\mathbf{T}$ to the present work is the following. First, we shall assume that each agent is able to recognise when

[11] It is to be noted that for BG2000, $\mathbf{S}$ has a single member, being the best-response decision making rule, such that g_i^{r+1} (for all $i \in N$) solves the profit maximisation problem,

$$\max_{g_i \in \mathbf{g}} \left[\pi_i(g_i \cap g_j^r)\right] \quad \forall \quad j \in N/\{i\} \tag{2}$$

Or, in the case of FK2003, $\mathbf{S}$ is none other than the decision-making rules that reside in each subject's mind (the locus of our inquiry).

two graphs are equivalent under re-labelling, as formulated in A. 1,

A. 1 (Type Recognition) *Given k un-identical graphs*

$$\{G_1(N_1^n, g), \dots, G_k(N_k^n, g)\}$$

differing only in the ordering of elements in N^n *(e.g.* $N_1^4 = \{1,2,3,4\}$ *and* $N_2^4 = \{2,3,1,4\}$*), then any agent* $i \in N$ *will recognise* $\{G_1, \dots, G_k\} \equiv \mathcal{T}_j$*, where* $\mathcal{T}_j \in \mathbf{T}(n)$*.*

Such an assumption means that given any graph $G(n)$, an agent will be able to recognise which minimal graph type $\mathcal{T}$ she has infront of her.

Second, as **S** has been defined above, the agent must be able to decide on an edge-sponsorship decision g that *applies to the instance*, that is, to the graph G. For this reason, we shall make the second cognitive assumption as below,

A. 2 (Context Invariance) *Given any instance of an information network* G *which corresponds to a minimal graph* $\mathcal{T}$*, any agent* $i \in N$ *is able to apply the resultant edge sponsorship decision* $s(\mathcal{T})$ *to the context, and thus arrive at* g_i *that accords to the instance* G *before her.*

The two assumptions given above may seem obvious since they accord with what appears intuitive to any human faced with the network formation game as described above. However, their adoption has important ramification for the modelling of such network formation environments – principally by resulting in a vastly reduced parameter space that an agent must cognitively face, and thus, providing a tractible (an intuitive) method of modelling such complicated decisions.

Finally, we can observe that for some agent $i \in N$, her decision making process will only ever need to consider the graph $G(N/\{i\}, g/\{i\})$ since first, adjacent edges – sponsored by another agent to i – do not increase i's observation set, and second, for any decision on some edge $g_{i,j} (j \neq i)$, whether to set $g_{i,j} = (1 \,\text{or}\, 0)$, the state of the link in the previous period is unimportant.[12]Hence, for the n-player game, each agent's strategy vector $\mathcal{S}$ need only prepare responses for the set $\mathbf{T}(n-1)$. For example, if $n = 4$, then the full set of graphs they need to respond to will be comprised of the agents $|N/\{i\}| = 3$, and be drawn from the canonical set $\mathbf{T}(3)$ as shown in Fig. 6.

[12]Consider, if $g_{i,j}^r = 0$, then establishing a link will incur cost C, which is the same as if $g_{i,j}^r = 1$, whilst leaving the link un-sponsored has neutral cost impact, which is the same as if $g_{i,j}^r = 1$, as before, since under this specification, there is no added cost for severring a link.

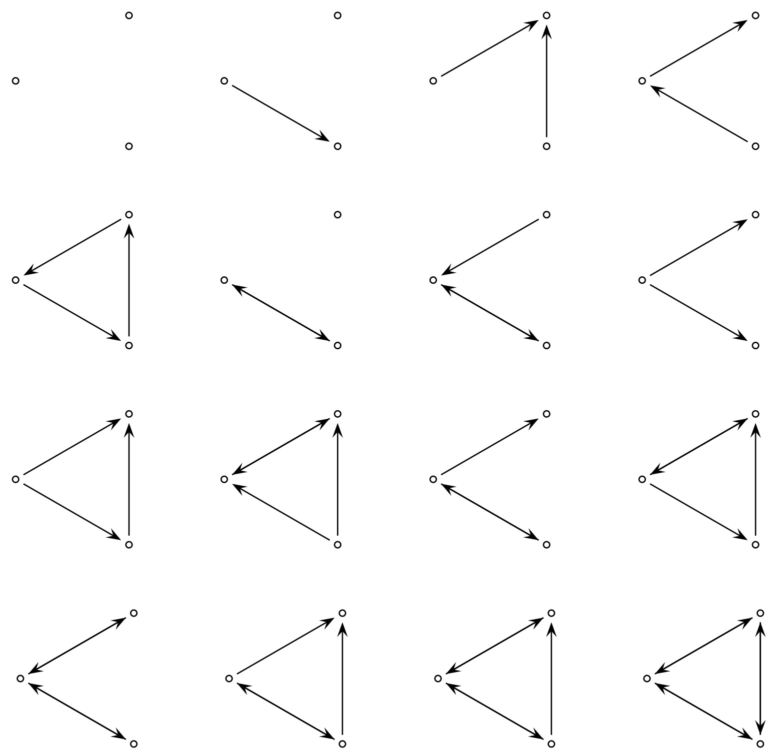


Figure 6. Full set of minimal absentee graph types, $\mathbf{T}(3)$.

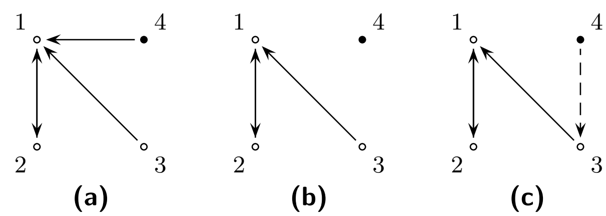


Figure 7. Example decision process: (a) The original graph G^r; (b) the absentee graph, or type graph $G^r/\{4\} \equiv \mathcal{T}_7$; and (c) the resultant graph, with 4's decision incorporated $G^r/\{4\} \cup g_4^{r+1}$ (assuming $g_j^{r+1} = g_j^r \ \forall \ j \in N/\{i\}$).

3.2.1 Example Application of **S**

Consider a 4-player network formation game, and suppose that in some stage and round, the current network is of the form $G^r(N, g^r)$, where $g_1^r = (1,0,0)$, $g_2^r = (1,0,0)$, and $g_3^r = (1,0,0)$, whilst $g_4^r = (1,0,0)$ (Fig. 7(a)). Consider agent 4 (currently sponsoring one link), they must form a decision response to the absentee graph $G/\{4\}$ (Fig. 7(b)). Now according to A.1, suppose $G/\{i\}$ is recognised as a member of type $\mathcal{T}_7$, and thus, suppose that she considers her response to such a type (say, g_4^*), and then by A. 2 judges that this response implies the sponsorship decision $g_4^2 = (0,0,1)$ in the given context (sponsor a link to 3, as in Fig. 7(c)).

We might summarise this process as,

$$G^r \longrightarrow G^r/4 \overbrace{\longrightarrow}^{A.\ 1} \mathcal{T}_7 \longrightarrow s(\mathcal{T}_7) \longrightarrow g_4^* \overbrace{\longrightarrow}^{A.\ 2} g_4^{r+1} \ . \tag{5}$$

3.3 Properties of the Decision-making Rule **S**

Although formal in nature, such a process is directly equivalent to the informal, 'recognise the graph type, and make a response that fits the given instance of that type.' In this way, the process is intuitive and somewhat obvious, but the reader will realise that such a decision-making process has some pleasing properties.

First, as has been mentioned, the present implementation of **S** greatly reduces the cognitive expectation on the model agents. It is trivial to show that the total number of graphs that are possible for a given number of agents is $2^{n(n-1)}$ and $2^{\frac{n(n-1)}{2}}$ for directed (one-way) and undirected graphs respectively. Which, for example, under one-way information flow equates to a total of 64 and 4096 directed graphs for $n = 3$, and $n = 4$ respectively. However, in the present formulation, since we focus on the absentee graphs, and then on equivalent *classes* of such graphs only, the total number of distinct structures $|\mathbf{T}(n)|$ is 3 and 16 (the count of structures in Fig. 6) accordingly; a dramatic reduction in the complexity of the modelling problem.

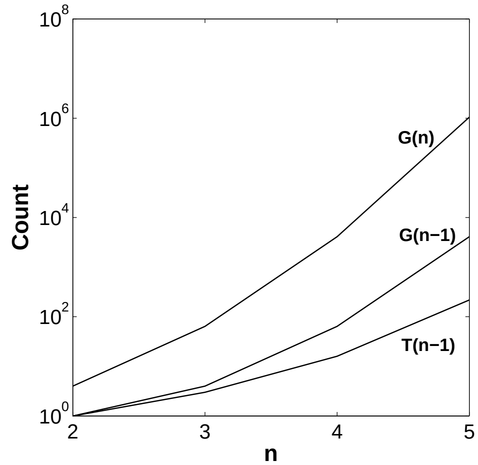


Figure 8. Count of distinct graphs in possible response sets: the full graph set $\mathcal{G}(n)$, the absentee full set $\mathcal{G}(n-1)$, and the minimal absentee set $\mathbf{T}(n-1)$.

Thus, we have a strong reason to suppose that the present formulation accords with what an average person could cognitively manage. Obviously, some people will be able to recognise more structures than others, and the numbers given should be treated as the upper-bound of what is required. By way of example, if the full graph set were considered for $n = 5$, a total of $|\mathcal{G}(5)| = 2^{20} \simeq 1 \times 10^6$ graphs would feature, or reducing this to just the absentee set, $|\mathcal{G}(n-1)| = |\mathcal{G}(4)| = 2^{12} = 4096$, whilst by constraining this set to just the *minimal absentee graphs*, that is $|\mathbf{T}(n-1)| = |\mathbf{T}(4)| = 218$, a four order of magnitude reduction is achieved. Figure 8 shows these counts for $n = \{2, \ldots, 5\}$. Second, as is the intention of this model, the specification of **S** in this way clearly allows for all manner of updating strategies. It should be pointed out that the 'best response' function (see footnote above) is accomodated in this framework – the best response is none other than the solution to a profit maximization problem contingent on the absentee graph as described above. However, in the present specification, by constructing **S** in the current way, we allow for all kinds of response functions. For example, strategies 'always sponsor-none',

$$s(\mathcal{T}_k) = (0, \ldots, 0) \ \forall \ \mathcal{T}_k \in \mathbf{T}(n) \ ;$$

'always sponsor-all',

$$s(\mathcal{T}_k) = (1, \ldots, 1) \ \forall \ \mathcal{T}_k \in \mathbf{T}(n) \ ;$$

and various strategies in between (e.g. 'uniform random':

$$s(\mathcal{T}_k) = (a, \ldots, a) \ \forall \ \mathcal{T}_k \in \mathbf{T}(n) \ ,$$

where $E[a] = \frac{1}{2}$) are all possible. Thus we have a truly rich environment, whereby via inoculation, or some other process (e.g. learning, see below), diverse agents can be modelled and interact.

3.4 *Learning & the Decision-rule* **S**

It is clear from the experimental work of FK2003 that subjects engaged in this non-cooperative decision problem undergo a process of learning throughout the stages of the game (as indicated by the positive gradient to Nash networks as reported in their paper, see Fig. 3). This kind of activity is natural in such a complicated setting since it employs largely foreign terms of reference (subjects would rarely play, or be aware that they are playing, such a game) and so, the game itself provides a forum whereby 'good' plays are revealed to agents over time.

Whilst it is noted that BG2000 (for example) were largely interested in the equilibrium (long-run) outcome of the non-cooperative network formation process, the interest of this paper is on how networks that are realised by human decision making come about. Hence, it is incumbent on the modeller that some attempt at learning is made.[13] Therefore, in the present framework we allow agents to learn from other agents in the following manner.

In the first stage (phase, $t = 1$), agents are endowed with some decision rule (e.g. random allocation as described in the foregoing paragraph), which they employ for all rounds (for all $r = \{1, \dots, R\}$) in that stage. At the end of this stage, a measure of the successfulness of each agent strategy is applied (e.g. $\overline{\pi} = \frac{1}{R}\sum_{r=1}^{R} \pi^r$) and used to rank each agent's performance. The highest ranked agent (or agents, if more than one agent shares the highest rank) shall then be called 'teachers', and the remainder of the agents (assuming a non-empty remainder set) the 'students'. The learning phase then ensues, with each student taking on a *public* part of a teacher's strategy (chosen equiprobably if more than one). That is, since not all graph types will necessarily be observed during a stage, each student has access to only those responses actually employed in the stage. This process occurs (with replacement of the teacher) until all students have been considered.

It is to be noted that so far, although agents might enter the model with a diversity of decision-making rules, if the model only ever allows for imitation within those rules, then long-run behaviour will be drawn only form the support of the initial rule distribution. Clearly, subjects are prone to discover new ways of solving the network formation problem that are

[13] It is to be noted that BG2000's solution process (best-response updating with inertia) might be construed as a type of 'learning' process, but infact, the device employed provides an excellent search method for finding the Nash outcome (as it does in each information flow case), but does not, as explained earlier mimic the way that real agents come to terms with the network formation problem.

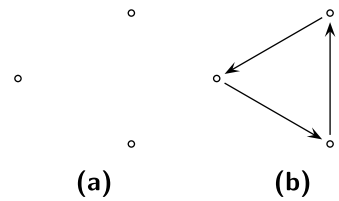


Figure 9. Learning example graphs: (a) $\mathcal{T}_1$; and (b) T_5

drawn from outside of their counterparts' rules. Therefore (and as is natural in modelling such *artificial adaptive agents* for precisely this reason), during the learning phase, let $m(\mathcal{S}, e)$ where $m : \mathbf{S} \times [0,1] \mapsto \mathbf{S}$ be a mistake-making or 'innovation' filter which is applied such that each agents' graph-type response decisions are reversed with vanishing probability ($e \simeq 0.01$) (e.g. from 'sponsor this link', to 'do not sponsor this link'). In the present model, $m()$ is applied to any student's learnt behaviours.[14]

3.4.1 Example of the Learning Process

Suppose two agents a and b with strategies $\mathcal{S}_a$ and $\mathcal{S}_b$ respectively have played the 4-player network formation game for R periods, and a ranking measure

$$\overline{\pi}_i = \frac{1}{R}\sum_{r=1}^{R} \pi_i^r \ , \quad i = \{a, b\} \tag{6}$$

has been applied with outcome $\overline{\pi}_a > \overline{\pi}_b$ .Then further suppose that a has responded to just two minimal absentee graphs in the previous stage (her *public* or *observable* plays), $\mathcal{T}_1$ and $\mathcal{T}_5$ (as shown in Fig 9) and in each case, her strategy response has been $s_a(\mathcal{T}_1) = (1,1,1)$ and $s_a(\mathcal{T}_5) = (0,0,1)$ respectively. Learning then proceeds as follows: first, as a result of the *imitative* part of the learning process,

$$\|s_b(\mathcal{T}_1)\ s_b(\mathcal{T}_5)\| = \|s_a(\mathcal{T}_1)\ s_a(\mathcal{T}_5)\|$$

where $\| \ldots \|$ indicates that a contiguous *sub-vector* of the horizontally concatenated public/observed plays is learnt; and second, the *innovation* filter is applied to the learning agent's decision-making rule, $m(\mathcal{S}_b, 0.01)$.

For example, this might have the outcome for b that

$$\mathcal{S}_b = \left\{(1,1,\overset{\downarrow}{0}), \ldots, s_b(\mathcal{T}_4), (0,0,1), s_b(\mathcal{T}_6), \ldots, s_b(\mathcal{T}_k)\right\} \ ;$$

[14] The qualitatitive naming given to such a filter is somewhat arbitary. Importantly, what $m()$ brings is the capacity for a broader solution-space than was present initially to be searched by the agents, and thus, they are able to realise solutions not previously known to them, or even known *in-part* to a member of their playing group.

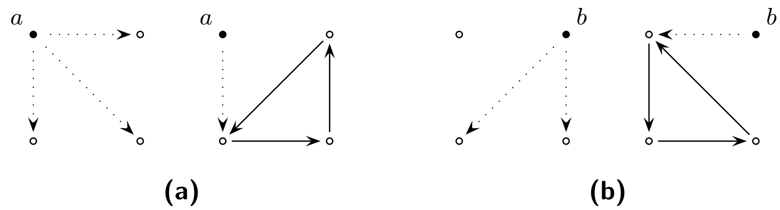


Figure 10. Outcome of learning process ([—] absentee graph edges, [$\cdots$] agent responses): agent b (b,left) learns a's (a,left) response to the graph $\mathcal{T}_0$ imperfectly; but agent b (b,right) successfully imitates a's (a,right) response to graph type $\mathcal{T}_5$.

in this example, b has learnt from a by imitation, but has made a single mistake in this imitation which has caused the generation of an altogether new decision-making rule over the two structures (summarised in Fig. 10).

3.5 Summary of Model

The present model aims to incorporate reasonable assumptions of human decision-making to investigate and explain the analytic claims and experimental observations of agents in a particular non-cooperative network formation problem. To do this, agents are given a strategy of action $\mathcal{S}$ that equips them to respond to the network-formation decisions of others through the assumead ability to recognise the type of network before them ($\mathcal{T}$). As the game progresses, agents who are initially endowed with a mixture of 'good' and 'poor' plays, as measured in action by some objective function, will observe the plays of their counterparts and mimic successful strategies at least in part. Further, between rounds, agents are able to update their own decision-making rules unilaterally through a process of low-level search of the decision space ($\mathbf{S}$).

4 Implementation

4.1 The Main Procedure

In order that the results from the present model could be compared effectively with those of FK2003, the subject-oriented experimental design was matched as closely as possible.[15] The total population of 'subjects' was spilt into a number of *mixing groups* (count `Mgrps`), so-called since over the subsequent phases, the *experimental group* (count `Xgrps`)that an agent will be allocated to (to play the n-player network formation game) will be drawn only from members of her mixing group. Hence, the mixing group will share 'playing information' over time with other members within the group, but not, therefore, with members outside of this group.

[15]Refer to Algorithm 1.

At the beginning of a modelling run, each sponsorship decision in the decision-making rule $\mathcal{S}$ is set (via RAND) to an equiprobably (sponsor)/(not-sponsor) element for all players,

$$s(\mathcal{T}_k) = (a, \ldots, a) \;\forall\; \mathcal{T}_k \in \mathbf{T}(n) \;,$$

where

$$E[a] = \frac{1}{2} \;,$$

and at the beginning of each network-formation game ($r = 0$), the graph is formed from each agent's first sponsorship action vector $s(\mathcal{T}_0)$. Thereafter, the network is defined by the round-wise application of each agent's decision making rule as described above.

Additionally, a sub-routine (GTYPES) is invoked to find the identity, and total number, of minimal type graphs that are possible for a given n. The outcome of this procedure defines the number of distinct graphs that an agent will have to respond to (in their decision-making rule $\mathcal{S}$) and also provides the necessary information to carry out the recognition- and application- steps in the subsequent round-based play.

After allocation of agents first, into their mixing groups, and then at the beginning of a stage into their actual experimental groups (count n), each experimental group is addressed in turn, with a total of R rounds being played. Here, as described in the model above, agents respond to the current minimal absentee graph ($\mathcal{T}$) as it appears to them from the present full graph (G), subsequently receiving some measure of successfulness (via OBJ). Additionally, the sponsorship decisions played by each agent are recorded as 'public' (since they are now revealed) and can then be imitated in the later learning phase, and various graph-properties are also recorded (via GSTATS).

After all rounds for an experimental group have been played, summary objective measures are compiled (here, via MEAN), and agents are designated to be either students or teachers (via RANK), with the students copying something of the public plays of the teachers as revealed through the fore-going rounds (via LEARN). This process continues for all experimental groups, and all stages, ending the modelling procedure.

4.2 Sub-Procedures

4.2.1 Identifying Minimal Types, GTYPES(.)

A procedure for identifying unique graph types, equivalent under re-labelling, was constructed in a two part process.[16] In the first case, each agent's feasible sponsorship strategies were counted and encoded, with all possible $n-1$ combinations with repeats enumerated. However, due to the presence of symmetries in the strategic plays, a second process was required to find

[16]Refer to Algorithm 2.

which of these graphs were actually equivalent under relabelling. To do this, a relabelling matrix was constructed and sequentially applied to each possible basic graph, with the outcome structures (aliases) being encoded and stored. A simple comparison could then be made of the alias 'finger-prints' for each of the possible type structures, with duplicates subsequently removed.

This two-step process, although somewhat exhaustive, does not give rise to large time-delays in modelling since it need only be carried out once.

4.2.2 Agent Performance, Obj(.)

Measuring agent performance has a two-fold interpretation in the implementation of the network formation game.[17] First, it provides the mechanism of incentives for agents to play (and solve) the network formation game, and thus gives rise to efficient or Nash play. However, secondly, since we are attempting to also model a method of agent learning, this measure will determine what strategy decisions are imitated by others in subsequent phases of play. Here, we shall consider two methods of agent performance measurement.

The first and most obvious choice is the *net payoff* to each agent, as conferred by the communication network itself and is simply calculated for a round by (1) and for a whole stage by (6).

However, it is possible that subjects do not rely on payoffs alone in determining outcomes. For example, consider the graphs presented in Fig. 11. For agent a, it is trivial to calculate π_a under (1) (denoting graph (a) L, and (b) R),

$$\begin{aligned} \pi_a(L) &= 4V - 3C; \\ \pi_a(R) &= 3V - C; \end{aligned}$$

which are equal for $V = 2C$ (a specific case studied in FK2003). It is reasonable to assume that the situation on the right (b) displays a better use of existing edge sponsorships for a out of the two, and would thus be more likely to receive imitative behaviour. It is relatively straight forward to note that infact, the non-cooperative network formation game actually gives rise to an *externality* generating environment. Principally due to the assumption of information flows (sharing) between agents, each agent's edge sponsorships do not just have an effect on their own information observations, but actually provide an information gathering structure for other agents.

Thus, we develop a measure of agent *efficiency* as an alternative objective measure. In line with the above observations, the measure should capture the degree to which an agent has been able to exploit any externalities in the graph structure. Intuitively, a simple ratio of obervations (μ) and

[17] Refer to Algorithm 3.

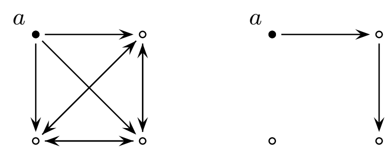


Figure 11. Graphs L (a) and R (b) giving rise to equal net payoffs for agent a.

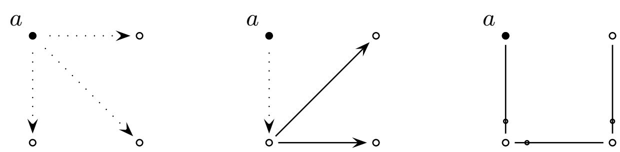


Figure 12. Efficiency measure example networks ($n = 4$): (a) minimum; (b) maximum (one-way flows); and (c) maximum (two-way flows). Agent sponorship strategy represented by dotted lines. NB: in (c), (following BG2000), sponsorship is indicated by small dot on edge sponsored by closest agent.

agent-degree (δ) ought provide such a measure. However, the convention of self-observation ($\mu = 1$) with zero link-sponsorship ($\delta = 0$) must be taken into account, thus we form the following measure,

$$f_i(\mu_i, \delta_i) = \frac{\mu_i V + C}{C(\delta_i + 1)} \ , \tag{7}$$

which has several pleasing properties. First, as desired, $f_\mu > 0$ and $f_\delta < 0$, but moreover, by rearranging, we have,

$$f_i(\mu_i, \delta_i) = \left(\frac{1}{\delta + 1}\right) \left[\left(\frac{V}{C}\right) \mu + 1\right] \ ,$$

which indicates that the ratio of observation benefits and edge sponsorship costs V/C weights the relative importance of the number of observations, versus the number of sponsorships. Hence, when V/C is large (i.e. low cost-regime), information observed is favoured, and numbers of edges sponsored is not, whereas in the alternate cast when V/C is small (i.e. high cost-regime), edge-sponsorships become more 'costly', having a greater relative effect on efficiency. Hence, although the measure captures the main interest by rewarding externality exploiting strategies, it also retains a connection to the actual cost of edge-sponsorship.

By way of example, and to normalize f on $[0, 1]$ the maximum and minimum strategies can be identified by considering the best- and worst-case externality scenario in each of the one- and two- way conditions. In the one-way case, the worst-case scenario is any strategy in which $\mu_i = n$ and $\delta_i = (n-1)$ (sponsor-all); such a case can be seen in Fig. 12(a). The best-case efficiency for an agent in the one-way case with low edge sponsorship costs is when they get the highest externality for their strategy, which can be easily verified to be where $\mu_i = n$ and $\mu_i = 1$; this situation is shown in Fig. 12(b). However, it can be shown[18] that in the one-way case, the optimal efficiency depends on the relation $C \gtrless V(n-2)$ and that in fact, if $C > V(n-2)$, then the maximal efficiency for an agent is obtained by sponsoring no links at all. That is, the return to sponsoring one link, *even if it returns full information*, is lower than the cost of sponsoring that link. Clearly, this is a natural phenomena and will have important ramifications in modelling.

Table 3. Efficiency measure $f(\mu_i, \delta_i)$ for all feasible combinations of μ_i and δ_i with $n = 4$. One- and two- way information flows indicated by $\rightarrow$ and $\leftrightarrow$ respectively.

		$\mathbf{C} = 5$		$\mathbf{C} = 25$	
μ_i	δ_i	$\rightarrow$	$\leftrightarrow$	$\rightarrow$	$\leftrightarrow$
1	0	0.33	0.11	**1.00**	0.38
2	0	-	0.41	-	0.59
3	0	-	0.74	-	0.80
4	0	-	**1.00**	-	**1.00**
2	1	0.11	0.04	0.33	0.13
3	1	0.56	0.19	0.60	0.23
4	1	**1.00**	0.33	0.87	0.33
3	2	0.04	0.01	0.11	0.04
4	2	0.33	0.11	0.29	0.11
4	3	*0.00*	*0.00*	*0.00*	*0.00*

In the two-way case, although the minimum is unchanged, the best-case strategy becomes 'sponsor-none' (as shown in Fig. 12(c)), since it is possible to obtain full-information and yet provide no support for the information network. This is akin to a truly non-cooperative ('defect') strategy since all benefits are retained, with costs born by others. The full results for

[18] We ask, is it possible that sponsoring no links, i.e. gaining $f(1,0)$, could ever be greater than the externality exploiting strategy that yields $f(n,1)$,

$$\begin{aligned} f(1,0) &\geq f(n,1) \;, \\ \frac{V+C}{C} &\geq \frac{nV+C}{2C} \;, \\ C^* &\geq V(n-2) \;. \end{aligned}$$

normalized efficiencies calculated by (7) under $n = 4$ and information flow of both types is given in Table 3 for comparison. It can be seen that subtle changes in preference ordering do occuring within columns due to the cost of link sponsorship; ultimately these will be reflected in learning outcomes when $f(\mu, \delta)$ is the objective measure.

4.2.3 Graph Analysis, GSTATS(.)

Of interest in the network formation game is the type of structure that the agents construct (calculated by GSTATS in Alg. 4, line 15). Table 1 summarises the graph qualities that are of interest under each formulation of the network formation game, as given by BG2000.[19] Tests for these structures were implemented with the following criteria:

M1C(.) In FK2003, *one-way minimally connected* graphs are defined to be where, 'all individuals are connected with each other and the removal of any direct link destroys this property.' (p.6) The following criteria was therefore used:

1. $\mu_i = n, \quad \forall\, i \in N$; and
2. Define the in-degree and out-degree of a node i to be d^i_{in} and d^i_{out} respectively, then for all $\{g_{i,j} = 1\} \in G$ either $(d^i_{out} = 1)$ or $(d^j_{in} = 1)$.

WHEEL(.) Here, the criteria for $m1c$ is true, and in addition, there does not exist two edges $g_{i,j} = 1$ and $g_{j,i} = 1$ for $(i \neq j) \in N$ (no two edges directly link the same nodes).

EMPTY(.) Trivially, the cardinality of the set $\{g_{i,j} : g_{i,j} = 1,\ \forall\ (i \neq j) \in N\}$ must be zero.

M2C(.) Three criteria apply:

1. $\mu_i = n, \quad \forall\, i \in N$;
2. There exists no sequence $a \to b \to \ldots \to a$ for information flows (that is, no cycles of at least 3 members);[20] and
3. There does not exist two sponsorships $g_{i,j} = 1$ and $g_{j,i} = 1$ for $(i \neq j) \in N$ (no two agents sponsor a direct link to each other).

CS_STAR(.) The center-sponsored star requires that there be exactly one central agent (sponsoring exactly $n - 1$ edges),

$$\left|\{i : \delta_i = n - 1,\ \forall\ i \in N\}\right| = 1 \ ,$$

[19] For proofs of these structures under each condition, the reader is referred to BG2000.
[20] To test for cycles, the fast algorithm of Tiernan (1970) was utilised.

$$
\begin{aligned}
\mathcal{S}_t &= \Big(s(\mathcal{T}_1), \ldots, \overbrace{000, 110, 001}^{\substack{\textit{section to be}\\ \textit{imitated}}}, 101, \ldots, s(\mathcal{T}_k) \Big) \\
\mathcal{S}_s &= \Big(s(\mathcal{T}_1), \ldots, 011, 010, 011, 001, \ldots, s(\mathcal{T}_k) \Big) \\
&\Downarrow \\
\mathcal{S}_s^* &= \Big(s(\mathcal{T}_1), \ldots, 000, 11\underline{1}, 001, 001, \ldots, s(\mathcal{T}_k) \Big)
\end{aligned}
$$

Figure 13. Example of a teacher-student learning process where $n = 4$ (thus each $s(\mathcal{T})$ comprises a sponsor-decision over $n - 1 = 3$ edges; $\{0, 1\}$ represent 'not-sponsor' and 'sponsor' respectively; $\underline{1}$ indicates a mistake.

and that this agent (say, i^*) is the only agent with a non-zero sponsor-set:

$$\delta_j = 0 \ \forall \ j \in N/\{i^*\} \ .$$

4.2.4 Agent Learning, Learn(.)

The present modelling prescription allows for simple treatment of learning.[21] Each agent retains a vector of sponsorship-decisions towards the $n-1$ other agents for the total number of minimal absentee graph types they might face (as enumerated by GTYPES). Futher, since we assume that agents are able to both recognise graph types (from the current instance) and then apply (via manipulation) their chosen response to the instance, inter-agent imitation is quite easily handled by simply reading a teacher's response vector and writing it to the appropriate section of the student's response vector (see Fig. 13).

The only additional treatment is given by a mistake-making process, applied with vanishing probability to each learnt decision, allowing agents to innovate, or mistake-make, their way to possible ways of playing not present in the original strategy space.[22]

[21] Refer to Algorithm 5.

[22] The present methodology might be recognised as a modifid *genetic algorithm* process, however, in this case, the traditional crossover operator is not a bi-genetic operation, but rather a single-direction transfer from teacher to student.

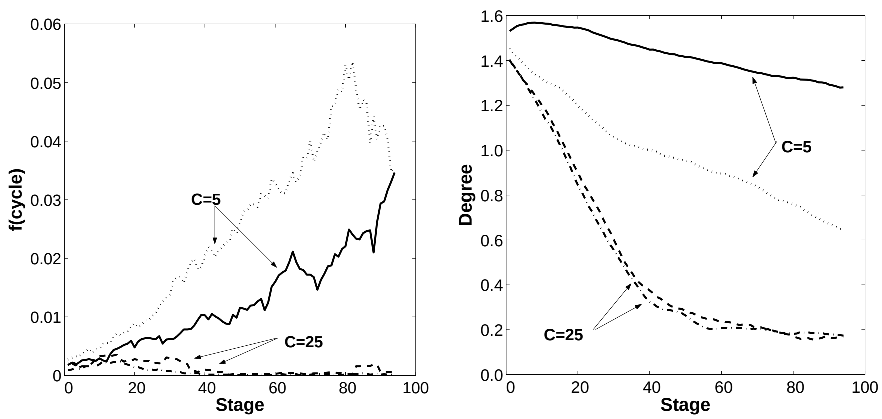


Figure 14. One-way information flows: (left) frequency of the 'wheel' (cycle) graph; and (right) mean agent degree. Cost of link sponsorship as indicated by C; and lines (C, objective measure): — (5,π); $\cdots$ (5,f); $--$ (25,π); and $-\cdot-$ (25,f).

5 Results & Discussion

5.1 Objective Measures & Learning

Initially, the one-way information flow case was analysed over a range of parameterisations.[23] Specifically, both high- and low- edge sponsorship costs were considered together with the two objective measures as given above in (1) and (7). Results for these experiments are given in Fig. 14.

As can be seen, both objective measures gave apparently good learning dynamics across stages in low-cost conditions, but the opposite is true of the high-cost regime. The average agent degree plot is revealing, indicating that as expected, in the high cost regime agents quickly chose to sponsor few, if any edges, whereas in the low cost regime, a distinction is apparent. The payoff objective (top) gave far more stable edge sponsorship strategies, whereas the efficiency measure, although producing mean agent degree of around 1 for much of the middle section, continues to fall, and finishes below 0.7; a trend reflected in the late dip in the $f(cycle)$ trace.

Although promising – agents are able to learn from each other towards realistic outcomes in some cases as observed in the laboratory – the model fails for the high-cost regime, and at least one objective measure is apparently brittle under even the better performing low-cost regime. Clearly the human trials indicate that despite the inherent risks to agents sponsoring links at high cost, this is exactly what happens, and more-over, is the learnt

[23]Results given for 10 independant trials, each comprising of 40 agents, split into 10 mixing groups of 4 for each stage play, over 100 stages, playing a 20 round network formation game. Experiment parameters as follows: $V = 10$, $m_i = 0.050$, $m_f = 0.005$, others given in caption to Fig. 14.

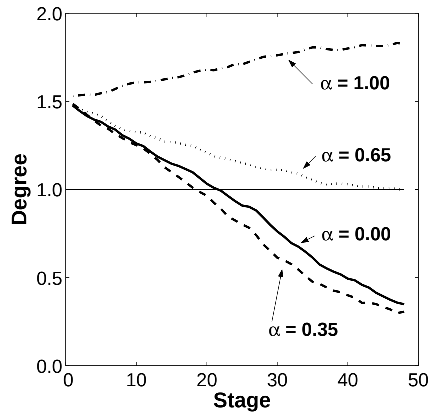


Figure 15. Combined (reciprocity,efficiency) objective measure used to determine learning outcomes. NB: line at $\delta = 1.0$ given to guide the eye only.

outcome across phases (in FK2003, average agent degree for one-way flows and $C = 25$ was 0.76).

5.2 Reciprocity & Cooperation

To address this problem, a further learning measure was introduced that gave part- or whole- priority to agents who showed a kind of 'cooperative' behaviour (network building despite risks). Specifically we suppose that for each revealed sponsorship decision within a round, agents are assigned a kind of 'reciprocity' measure r_i as follows,

$$r_i = \begin{cases} 0 & \text{for} \quad \delta_{in} = \{0,1\}, \quad \delta_{out} = 0\,, \\ 1 & \text{for} \quad \delta_{in} = 1, \quad \delta_{out} = 1\,, \\ 2 & \text{for} \quad \delta_{in} = 0, \quad \delta_{out} = 1\,, \end{cases} \tag{8}$$

where δ_{in} and δ_{out} are an agent's in- and out- degree respectively. For each agent over R rounds, the mean value of r is calculated and contributes to the overall success measure used to determine 'teachers' and 'students' by simple convex combination,

$$\Omega_i = \alpha\langle r_i\rangle + (1-\alpha)\{\langle \pi_i\rangle, \langle f_i\rangle\} \tag{9}$$

where $\alpha \in [0,1]$, and $\langle\cdot\rangle$ indicates arithmetic mean.

Figure 15 gives $f(cycle)$ for a 50 stage run of the high-cost regime ($C = 25$) using (9) with the efficiency measure (since it provided more stable results in the first experiment) and $\alpha \in \{0.00, 0.35, 0.65, 1.00\}$.[24] Values for α of 0.00 and 0.35 appear to do little to change the overall down-ward trend

[24]Experiment parameters as per previous experiment, with $T = 50$.

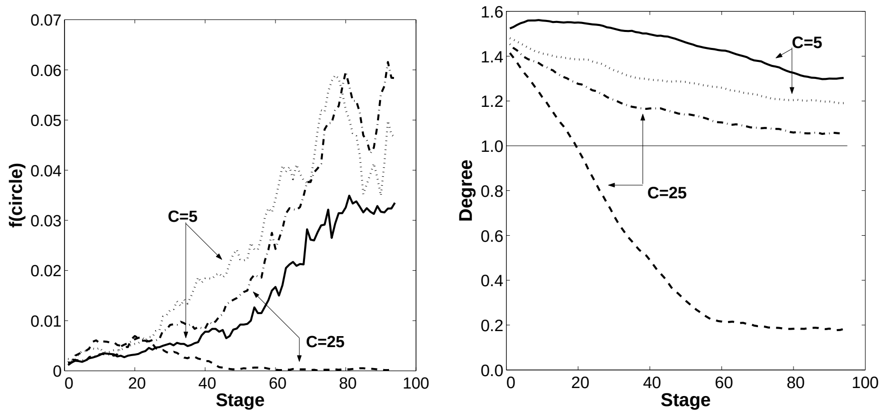


Figure 16. One-way information flows under *combined* objective measure conditions: (left) frequency of the 'wheel' (cycle) graph; and (right) mean agent degree. Cost of link sponsorship as indicated by C; and lines (C, objective measure): — (5,π); $\cdots$ (5,f); $--$ (25,π); and $-\cdot-$ (25,f).

in edge sponsorship, however, a non-linear movement for $\alpha \geq 0.65$ yields the desired result. Taking this new *combination* objective measure ($\alpha = 0.7$) to a longer study gives rise to more realistic learning and behavioural outcomes (see Fig. 16). In three of the four cases (both measures under $C = 5$, and for efficiencies under $C = 25$), agents reduce average edge-sponsorship numbers towards 1.0 or above, with resultant strategic outcomes also displayed in relatively good $f(cycle)$ growth over time.

5.3 Learning Dynamics

To study the dynamics of the learning process for comparison with the subject data, two long-run experiments were conducted. In the first, the conditions as determined above were implemented as is: one-way flow, low- and high- costs of edge sponsorship. In the second, the two-way information flow condition is introduced for the first time, to see if the seemingly realistic conditions chosen above are applicable to replicate the poor subject results of the two-way flow game.

For ease of comparison with FK2003, the long-run studies are split into meta-stages so that both play *within* and *between* learning junctions can be compared.[25] To this end, in each bar-chart that follows, the 200 explicit model stages were grouped under 5 meta-stages,

$$t = (1, \ldots, 50), \ldots, (151, \ldots, 200)$$

and within each of these, average data for play within specific rounds were

[25] Experiment parameters as given in footnote 23, with $T = 200$, combined (reciprocity–efficiency) objective measure used ($\alpha = 0.70$).

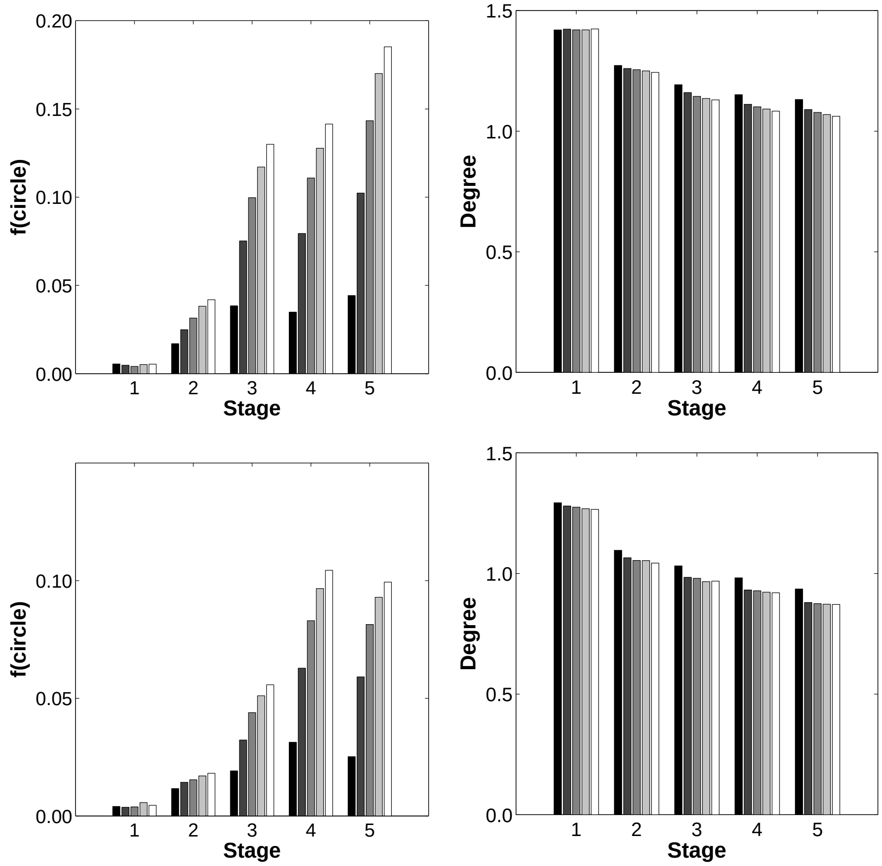


Figure 17. Meta-stage bar-charts showing learning dynamic within and between learning blocks (explained in text). Top: $C = 5$; bottom: $C = 25$.

also split into 5 groups (that is, $r = (1, \ldots, 5), \ldots, (16, 20)$). The outcome for the one-way case is very interesting (see Fig. 17), with improvements clearly visible both within and between phases. This phenomena is exactly as observed in the FK2003 subject trials; the authors hypothesised that in the early rounds of a stage (recall: groups are re-formed between stages), mis-coordination occurs due to the unfamiliarity with the playing partners, but this goes away over time.

Given the seemingly strong correlation between average degree and propensity of groups to find Nash structures, it could be hypothesised that the strong learning results between stages and improvement dynamic within stages are simply the result of random occurence aided by players lessening the number of edges that they sponsor. To test this hypothesis, the same data treatment is performed on mean agent degree for the one-way case (see right of Fig. 17). As can be seen, although there is a slight reduction of

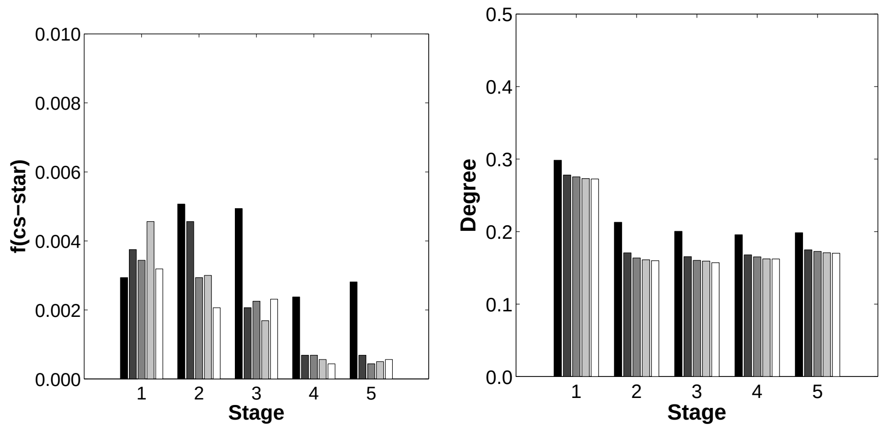


Figure 18. Learning within and between meta-stages in the two-way information case, $C = 5$; frequency of centre-sponsored star (left) and average agent degree (right).

average edge sponsorships over the 5 meta-stages which may indeed account for some of the overall learning dynamic between stages, within each stage, there appears to be a vanishingly small difference in agent edge sponsorship numbers, whereas the probability of finding the Strict Nash structure (in this case, the cycle-graph) grows markedly. Hence, we can conclude that the apparent improvement dynamic within stages of the artificial agents is not by chance alone under lower edge sponsorships, and is very likely due to strategic learning as modelled.

Interestingly, the higher edge-sponsorship regime (bottom charts Fig 17) do not reveal a similar improvement to Strict Nash play as was evident with human subjects. However, it is quite possible that a stronger form (higher α) of learning esteem for playing partners who play cooperatively (support edges even if others aren't) is needed in the $C = 25$ case to replicate human decision-making under these conditions. From the previous discussion and analysis above, it is clear that use of the efficiency measure (as opposed to straight payoffs) *and* the reciprocity component are necessary conditions to prevent the kind of mean edge sponsorship decline that can occur.

Turning to the two-way information flow case, it is interesting to see if the present specification of learning modelled on the one-way outcomes gives rise to problematic play in the two-way case, as was the result in human trials. Figure 18 certainly seems to suggest this. The figure gives results for the low-cost regime, showing the difficulty that agents had with finding the efficient outcome which is likely a reflection of the efficiency measures given to each strategy outcome as presented previously in Table 3; the 'sponsor-none' regime is simply far too attractive, although it is self-defeating for all concerned. Indeed, for the high-cost regime, not a single efficient graph (centre-sponsored star) was arrived at, as was found in the human trials.

Hence, we find that the present results with artificial agents are very close to that of the human subjects.

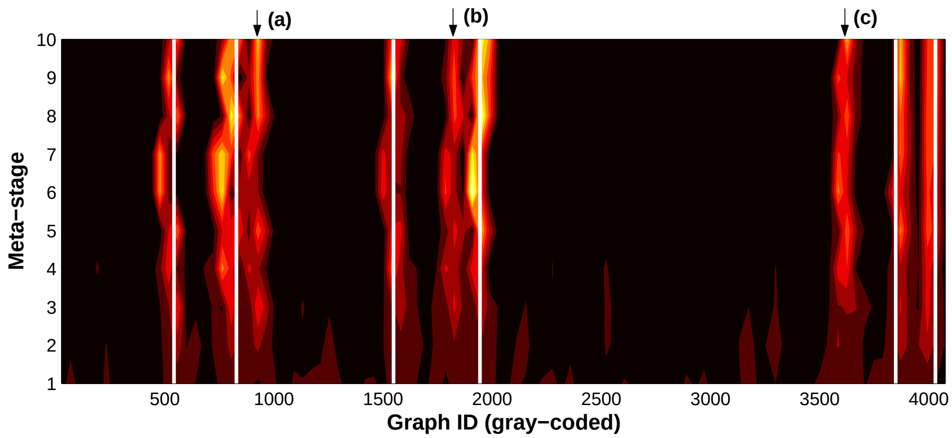
5.4 Graph Outcomes, a Closer Look

Given the findings above, that the artificial agents behave in a similar manner to human subjects when they jointly learn efficient and quasi-cooperative play, it is of interest to study what strategies are being played the rest of the time; that is, during the rounds were efficient play is not observed. To this end, all network outcomes for every playing group in every round were first Gray-coded[26] and then a frequency distribution of the resultant Graph IDs was generated, grouping the graphs of the last ($r = R$) round in each case within 10 distinct 'meta-stages' – representing a grouping of 20 individual stages together (there were 200 individual stages in each trial). Recall that learning occurs between stages, hence, by grouping in this way, the relative abundance of different graph types throughout aggregate learning stages can be determined.

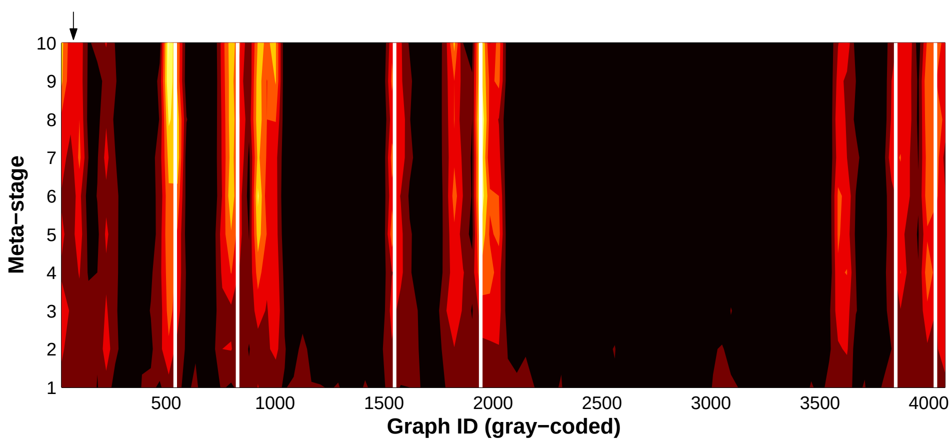
Figure 19 shows the outcome of this analysis for the one-way information flow cases, with the low- and high- cost regime represented at the top and bottom of the figure respectively. In both cases, the narrowing in on just a few graph types (or close to these types) can be seen, with the Strict Nash outcomes – the *cycle* case, for one-way flow) – represented by vertical white lines giving a very good prediction of learning outcomes. (Recall, that by the Gray-coding procedure, graphs that differ by just one *bit*, i.e. one sponsorship decision between two nodes, will lie next to each other in the encoded Graph ID space.) It is to be noted that in the high-cost regime, for one-way flow, an additional Strict Nash outcome – that of the empty graph (Graph ID = 0) – becomes prevalent (indicated by the arrow in the bottom plot).

In the top plot, three prominent and persistent peaks are identified by markers (a-c) with arrows. A closer inspection of these areas indicate that just three graph IDs are responsible, namely, IDs 924, 1820 and 3614. These structures are represented in Fig. 20). Interestingly, whilst the third case (ID 3814) appears to be a pre-curser to the cycle graph, requiring just D to correct their play (sponsoring the link $D \to A$, the first two graphs (IDs 924 and 1820), which are actually analogs of one another (under relabelling) have the potential for cyclic behaviour. Taking case ID 1820 for example, notice that the present information sets for each of A and D are $\{A, B, C\}$ and $\{D, B, C\}$ respectively, and so, if B and C were to play a best-response with the restriction that they only sponsor one link, they would sponsor links

[26] So named after Frank Gray, patenter of the encoding technique. From Press et al. (2002), 'A Gray code is a function $G(i)$ of the integers i, that for each integer $N \geq 0$ is one-to-one for $0 \leq i \leq 2^N - 1$, and that has the following remarkable property: The binary representation of $G(i)$ and $G(i + 1)$ differ in *exactly one bit*.' (emphasis retained)

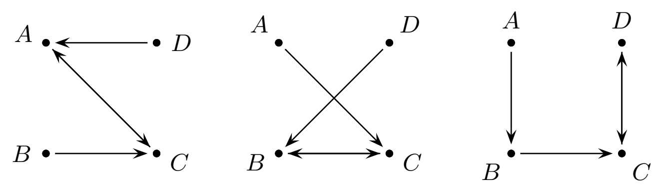


(a) $C = 5$



(b) $C = 25$

Figure 19. Relative frequencies of Gray-coded graph IDs in the one-way information flow case. Each Meta-stage represents the aggregate data taken from the last round in each of 20 grouped stages (e.g. the first meta-stage comprises $t = \{1, \ldots, 20\}$. Lighter shading indicates higher relative frequency. Vertical lines represent Strict Nash outcomes in each case. Other features explained in text.



(a) 924 **(b)** 1820 **(c)** 3614

Figure 20. Alternate persistent graph IDs (Gray-coded) observed during one-way information flow at both high- and low- edge sponsorship costs (see Fig. 19.

to A and D respectively. Whilst on A and D would sponsor links to each other, and thus expect to gain full-information. However, the graph that would result would be exactly the same conformation as ID 1820 under relabelling. Hence, if each player plays without inertia[27] then an infinite cycle would result. Indeed, upon closer inspection of the relative frequency plot at ID 520 (the resultant Gray-coded graph ID), a small peak is discernable, though not with nearly the same intensity as the former structure.

In the case of two-way information flows and low the outcome is stark, with the vertical lines falling in areas associated without spectrum mass. Even in the detailed section (IDs 0-100) shown in the bottom plot, where relative frequencies are amplified, the strict-nash structures are not arrived at, with the first structure (at ID 4) the only one with some interest. In the top plot, three prominent and persistent structures are identified and marked by arrows (a-c). Upon closer inspection of these areas, it was found that 5 graphs contributed to the peaks: 511, (1008,1009), and (2032,2033). These structures are shown in Fig. 22.

As before, the propensity for cyclic behaviour is again visible, with graph ID 511 a possible pair with IDs 1008 and 1009 (although in these cases, not in terms of best-response play). Recall that in the two-way case, information can be gained 'passively', that is, by simply receiving the information flows that arrive as a result of a partner agent's edge sponsorship. On the whole however, the collection of prominent graphs, other than displaying low edge-sponsorship, do not appear to yield to similar intuitive explanation as in the one-way case. This is consistent with the lack of strong learning behaviour in this case however, and it is quite likely that these graphs are prominent purely as low-edge sponsorship cases alone, rather than representing clear strategic play.

[27] i.e. $g_i^{t+1} \neq g_i^t \ \forall \ i \in N$.

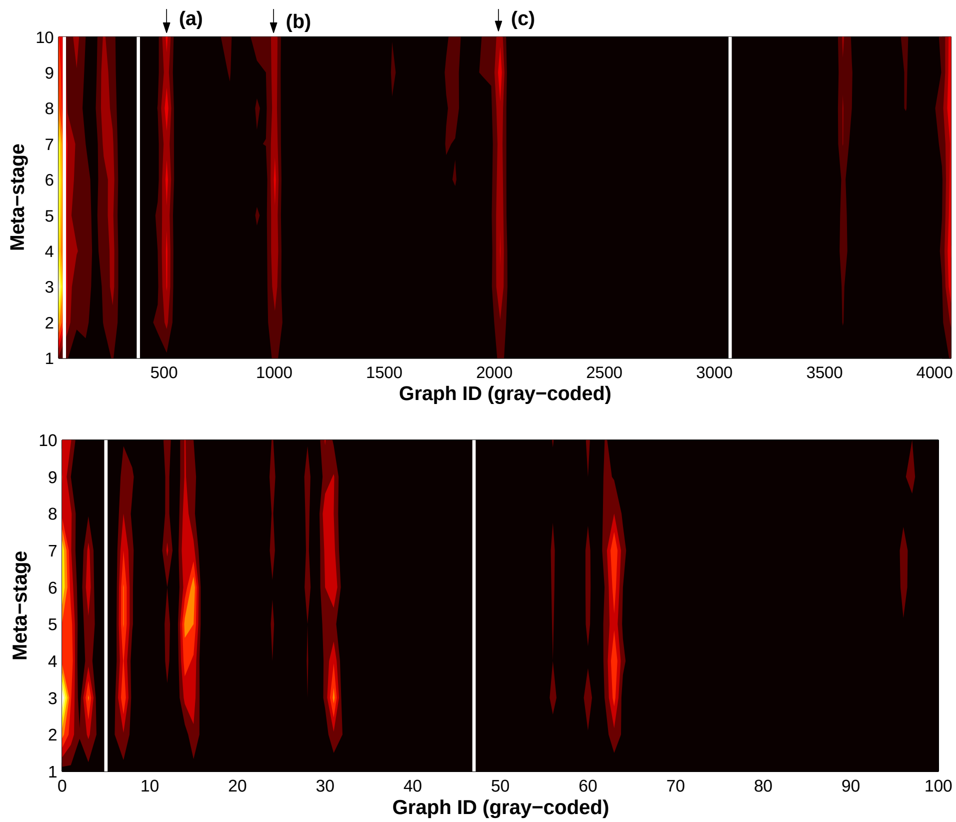
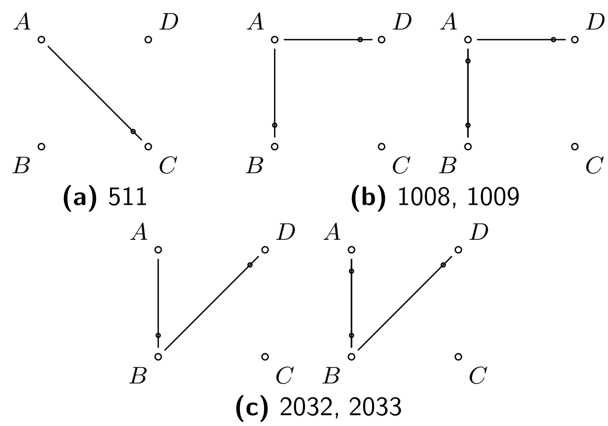


Figure 21. Relative frequencies of Gray-coded graph IDs in the two-way information flow case: (top) the full spectrum; and (bottom) detail from graph ID region 0-100, note that relative frequencies are greatly amplified in the lower plot due to the smaller sample range. In both plots, $C = 5$; vertical white lines indicate the position of strict-Nash structures (centre-sponsored star type). Other features and notes to construction explained in caption to Fig. 19.



(a) 511 **(b)** 1008, 1009 **(c)** 2032, 2033

Figure 22. Alternate persistent graph IDs (Gray-coded) observed during two-way information flow at low edge sponsorship costs (see Fig. 21).

5.5 Inertia: Best Response, Equality & Emergence

FK2003 found that where an agent played the same strategy in the present round as in the previous round, i.e. displaying strategic *inertia*, there was a significant explanatory effect of payoff *equality* in the preceeding round (controlling for the previous play being a best-response play). In other words, an agent who found their payoff different to other player's payoffs in the previous period would be more likely to *change* their strategy in the current period.[28] For the purposes of the present study, agent inequality (following Fehr and Schmidt (1999), in FK2003) is defined as follows,

$$e_i = \sum_j |\pi_i - \pi_j| \ , \tag{10}$$

for all $j \neq i$; the sum of absolute differences between an agent and all other players in the game.

To investigate this correlation, a plot of strategic inertia frequency versus learning stage was prepared (see left of Fig. 23), together with a plot of the same versus measured agent inequality (see right of Fig. 23).[29]

Taken together, in the one-way information case, there does appear to be a good correlation between inertia and inequality, with both experiments finishing towards the top-left quadrant of the plot (low inequality, high inertia). However, from the plot on the left, it is clear that this correlation is an *emergent* property of the system, being significantly affected by the

[28]Note that this is not the same as asking whether or not an agent chose to *update* their strategy or not since they could possibly playing the same strategy as the previous round, and hence would not display strategic inertia.

[29]Data represents mean measures over all trials and rounds $\{2, \ldots, R\}$, NB: lines are smoothed (10 points) for clarity.

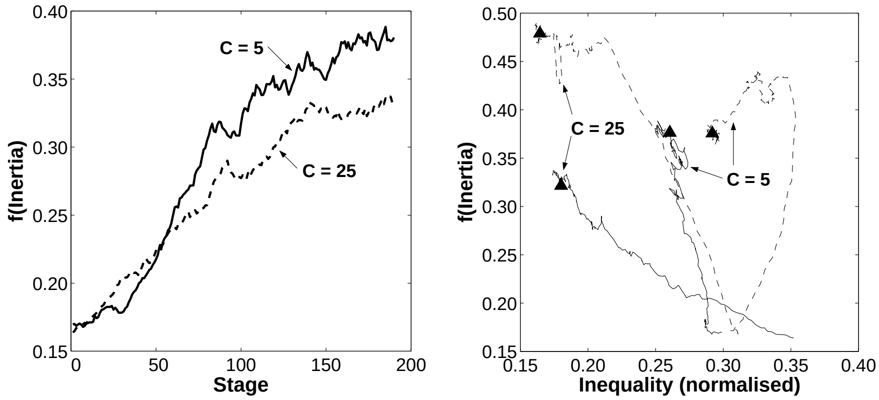


Figure 23. Strategic inertia frequency. (Left) for each learning stage, one-way information flows; and (right) for mean agent inequality measures, (▲) indicates the final experiment stage; both one- (solid) and two- (dashed) way information flow cases are plotted.

successive learning stages, rather than seeming to be foundational for the rise of Nash play (compare Fig. 17). In the two-way case the results are similar, whilst it is clear that the connection between inertia and inequality is more complicated, with the low cost case showing severe non-linearities in its system trajectory, however, it can be seen that from the initial condition through to the end of the experiment (in each cost-scenario), the level of inertia moves positively with time, and is not necessarily connected to inequality in all cases. Hence we might conclude that the present experiments urge caution in attributing human subject solution techniques to social preferences such as equality of payoffs, and instead note that such preferences are a feasible emergent property of the learning environment, in this case through the accumulation of efficient and/or cooperative/reciprocal plays.

6 Conclusions

Whilst much is still to be understood in terms of how individuals go about solving complex problems such as the communication network formation problem studied in this contribution, the present analysis offers several insights. First, by applying a novel graph-equivalence criterion and mapping technique, we have shown that the apparently vast and seemingly unfeasible strategy space can be modelled with artificial adaptive agents in an intuitive manner. Whilst this approach is likely useful for $n \lesssim 10$, it is in this region that the problem retains strategic interest given the limits of human functionality.

Second, we have observed that simple payoff learning optimization does not give rise to realistic behaviour since agents are apparently more willing to venture edge sponsorships than 'good' strategy would dictate. Instead, a mixture of a kind of strategic *efficiency* and agent-to-agent reciprocity appears a reasonable fit to observeed experimental behaviour. However, as is apparent with human subjects, such activity does very poorly in the two-way information environment with not only the payoff optimal play, but also the seemingly efficient play being to severely limit the number of edges sponsored, resulting in poor Nash or Stric Nash equilibria statistics.

Finally, the link between previous payoff equality and strategic inertia was investigated, appearing to be less of a causal relationship, than an emergent phenomena due to the learning process as previously above. Although, the FK2003 considerations concerning social-preferences could well be expressed through the learning dynamic (via the reciprocity measure) rather than at the point of round-based play itself.

References

Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *American Economic Review, Papers and Proceedings*, 84(2):406–411.

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

Chwe, M. S.-Y. (1995). Strategic reliability of communication networks. Technical report, University of Chicago, 1126 East 59th Street, Chicago, Illinois 60637, USA.

Chwe, M. S.-Y. (2000). Communication and coordination in social networks. *Review of Economic Studies*, 67:1–16.

Comellas, F., Ozon, J., and Peters, J. G. (2000). Deterministic small-world communication networks. *Information Processing Letters*, pages 83–90.

Falk, A. and Kosfeld, M. (2003). It's all about connections: Evidence on network formation. Discussion Paper 777, Institute for the Study of Labor (IZA). available at http://ideas.repec.org/p/iza/izadps/dp777.html.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, 114:817–868.

Jackson, M. O. and Watts, A. (2002). The evolution of social and economic networks. *Journal of Economic Theory*, 106:265–295.

Matros, A. (2004). Simple rules and evolutionary selection. IZA Bonn research seminar paper, available at http://www.iza.org/.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2002). *Numerical Recipes in C: The Art of Scientific Computing*. Cambrindge University Press, Cambridge CB2 1RP, 2nd edition.

Slikker, M. and van den Nouweland, A. (2000). Network formation models with costs for establishing links. *Review of Economic Design*, 5:333–362.

Tiernan, J. C. (1970). An efficient search algorithm to find the elementary circuits of a graph. *Communications of the ACM*, 13(12):722–726.

Algorithm 1 Pseudo-code for the main program.

```
procedure MAIN
    Xgrps ← Mgrps × XperM;                          ▷ Total experiment groups
    N ← Xgrps × n;                                  ▷ Total agents in model
    T, nT ← GTYPES(n);
    S[0] ← RAND(N, (n − 1) × nT);
    allocate agents into Mgrps mixing groups;
    for t ← 1 . . . T do                            ▷ For all stages
        form experimental groups (exp_groups);
        for ex ← 1 . . . Xgrps do                   ▷ For all exp. groups
            G ← EMPTY; P[ex] ← [ ];
            for r ← 1 . . . R do                    ▷ For all rounds
                G ← g* ← S(T(G));                   ▷ Apply strategy rules S
                P[ex] ← P[ex] ∩ g*                  ▷ Record public strats
                O[r] ← OBJ(G);
                st[r] ← GSTATS(G, C, V, dirF);
            end for
            calculate and store summary graph stats;
            O* ← MEAN(O);                           ▷ Summary objectives
            rnk ← RANK(O*)                          ▷ Rank agents in exp. group
            S[t] ← LEARN(ex, S[t − 1], rnk, P);
        end for
    end for
end procedure
```

Algorithm 2 Procedure for enumerating **T**

```
procedure GTYPES(n)
... First get all distinct sponsorship graphs
    a ← n − 2;                                    ▷ Agents to link to
    n_sp ← 2^a;                                   ▷ No. of distinct strategies
    T ← (1, . . . , n_s p);                       ▷ The first agent
    for c ← 1, . . . , a do                       ▷ For subsequent agents
        t ← [ ];
        r ← 0;
        l ← ROWS(T);
        for i ← 1, . . . , l do                   ▷ For each previous situation
            for j ← MAX(s(i,:)), . . . , n_sp do  ▷ For each distinct strategy
                r ← r + 1;
                t(r, .) ← [s(i, :) j];
        T ← t;
    nT ← ROWS(T);
... Now purge the list of equivalent graphs due to symmetry
    R ← PERMS(n − 1);                             ▷ Node relabelling options
    for i ← 1, . . . , nT do
        g_id ← [ ];
        for all lbl ∈ R do                        ▷ For each labelling
            g_id ← [g_id   RELABEL(T(i, :), lbl)];
        G(i, :) ← g_id;                           ▷ Store the alias graphs
    Gvec ←ROW2INTEGER(G);                         ▷ Convert each set to unique identifier
    u_i ←UNIQUE(Gvec);
    T ← T(u_i)                                    ▷ Take only unique types
    nT ← ROWS(T);
    return T, nT
end procedure
```

Algorithm 3 Pseudo-code for Objective measure procedure

```
procedure Obj(G, C, V, pF)
    D[1 ... n][1 ... n] ← MIN_PATH(G)
    for all i ← 1 ... n do
        μ[i] ← 1 + |{j : D[i][j] ≠ ∞, ∀ j ≠ i}|;        ▷ Count(nodes observed)
        δ[i] ← SUM(G^T)                                  ▷ Node degree
    end for
    if pF = 1 then                                       ▷ Payoffs
        π ← μ ∗ V − δ ∗ C;
        return π
    else if pF = 2 then                                  ▷ Efficiencies
        φ ← (μ ∗ V + C)/(C ∗ (δ + 1));
        return φ
    end if
end procedure
```

Algorithm 4 Pseudocode for graph statistics procedure

```
procedure GStats(G, C, V, dirF)                  ▷ (NB: end statements omitted)
    c, e, m, w ← 0;
    if C ≤ V then                                ▷ Low cost case
        if ISEMPTY(G) then
            e ← 1;
        else
            switch dirF
            case 1
                w ← WHEEL(G);
                m ← M1C(G);
            case 2
                c ← CS_STAR(G);
                m ← M2C(G);
    else if C > V then                           ▷ High cost case
        if ISEMPTY(G) then
            e ← 1;
        else
            switch dirF
            case 1
                m ← M1C(G);
                w ← WHEEL(G);
            case 2
                m ← M2C(G);
    return c, e, m, w
end procedure
```

Algorithm 5 Pseudocode for Learning procedure

```
procedure LEARN(ex, S[t − 1], rnk, P)
    T ← {i ∈ ex : rnk[i] = 1}
    L ← {i ∈ ex : rnk[i] ≠ 1}                ▷ Identify teachers (T) & students (L)
    for all i ∈ L do
        j ← CHOOSE(T)                         ▷ Choose (equiprob.) teacher
        k ← P[j]                              ▷ Ids of public strats
        S[i][k] ← MISTAKE(S[j][k])
    end for
    S ← MISTAKE(S)                            ▷ Apply innovation to all strats
end procedure
```