Short communication

On the difficulty of characterizing network formation with endogenous behavior☆

Benjamin Golub [a,b], Yu-Chi Hsieh [a,*], Evan Sadler [c]

[a] *Northwestern University, United States of America*
[b] *Stanford University, United States of America*
[c] *Columbia University, United States of America*

ARTICLE INFO

ABSTRACT

Bolletta (2021) studies a model in which a network is strategically formed and then agents play a linear best-response investment game in it. The model is motivated by an application in which people choose both their study partners and their levels of educational effort. Agents have different one-dimensional types – private returns to effort. A main result claims that (pairwise Nash) stable networks have a *locally complete* structure consisting of possibly overlapping cliques: if two agents are linked, they are part of a clique composed of all agents with types between theirs. A counterexample shows that the claimed characterization is incorrect. We specify where the analysis errs and discuss implications for network formation models.

Bolletta (2021) studies the joint strategic choice of network links and productive effort, motivated by applications in the economics of education. The theoretical findings are used in interpreting empirical studies on peer effects, such as Carrell et al. (2013). This fits into a literature on network games[1] and games on endogenous networks – see, e.g., König et al. (2014), Hiller (2017), and Badev (2021).

The paper examines a stylized network formation process. At each time, starting from a status quo network, two (random) agents i and j can form a new link between them by mutual consent; alternatively, either can unilaterally sever any of their own existing links. After each such stage, agents play the static Nash equilibrium of a coordination game in the current network: each wants to match an average of a private ideal effort (its "type") and average behavior in its neighborhood. Higher types are inclined to make higher efforts in the effort game, holding spillovers fixed. In assessing the consequences of their decisions in the formation stage, players anticipate equilibrium behavior in the action-choice game on the resulting network. The process is said to reach an equilibrium if no pair of agents wish to change their linking decisions at the formation stage. In the leading application, links are study-partner relationships and the action is educational efforts.

A basic theoretical question is: which networks can be equilibrium outcomes? A main claim of Bolletta (2021), stated in its Proposition 1, is that equilibrium networks have a specific structure, called *locally complete*. This means that, at any equilibrium, if two agents are connected, they and all agents with types between theirs form a clique – a completely connected subnetwork.

The purpose of this note is to point out that this proposition is false. We do this by providing an explicit counterexample: an equilibrium outcome that is not locally complete. We also show that this outcome can be reached by the dynamic process studied in Bolletta (2021) starting from the empty network. It is thus a natural outcome reachable from a "neutral" status quo. Our note also examines arguments that underlie the proposition in order to highlight where the analysis errs.

We give a brief roadmap. In Section 1, we establish key notation and review Bolletta's model as well as some definitions relevant both to Bolletta's and our analyses. In Section 2, we present our counterexample to Proposition 1 in Bolletta (2021); we also discuss the key errors in the proof. Section 3 shows that the counterexample network to Bolletta's Proposition 1 is a natural outcome of the dynamic process, not just an equilibrium of it. Section 4 concludes.

1. The model and a main result of Bolletta (2021)

In this section, we review Bolletta (2021)'s model setup and state the result that we focus on. We mostly work within the same notation, but in some cases, we make slight modifications.

1.1. Model

There are N agents. All networks discussed are undirected, unweighted graphs without self-edges on the nodes $[N] := \{1, 2, \ldots, N\}$.

Agent i has a type θ_i, a neighborhood $g_i \subseteq [N]$, and an action y_i. Given a network g, the degree of player i is $d_i = |g_i|$, the number of links i has. The vector of efforts for all players is denoted by $y = (y_1, \ldots, y_N)$. The payoff of player i is given by

$$U_i = u\left(\mathbf{y}, \theta_i, g\right) = (1 - \alpha)\theta_i y_i - \frac{y_i^2}{2} + \sum_{j \in g_i} \left(\delta + \frac{\alpha y_i y_j}{d_i}\right)$$

if player i has a non-empty neighborhood, and $U_i = \theta_i y_i - \frac{y_i^2}{2}$ otherwise. The value δ represents an exogenous positive benefit that a player receives from each link. The parameter $\alpha \in [0, 1)$ captures the weight that players put on matching their own type versus the average of their neighbors' actions. We assume that θ_i can take on any positive real value. By standard results on network games, if $\alpha \in [0, 1)$, then fixing a network g, the game of effort choice has a unique equilibrium (Bolletta, 2021, Lemma 1).

The paper proceeds to introduce a dynamic process of network evolution and an equilibrium concept, which defines points at which the process can come to rest. The main claims of Bolletta's paper apply to the equilibrium outcome, so we can present our counterexample using only the definitions above. Later, in Section 3, we discuss the dynamic process leading to equilibrium outcomes and discuss our counterexample in light of it.

1.2. Equilibria

Given a specification of payoffs in each possible network, an equilibrium network is defined to be one in which no pair of agents mutually wishes to create an additional link, and no agent wants to unilaterally sever any subset of her existing links. This equilibrium concept corresponds to the standard notion of pairwise Nash stability. Crucially, the payoffs of any network g' (contemplated by agents at the status quo and in their deviations) are determined by equilibrium effort choices $y^*(g')$ given the network g'. We denote these payoffs by $U_i(g')$.

Definition (*Pairwise Nash Stability (Jackson (2010))*). A network g is pairwise Nash stable if

1. For all $ij \notin g$, if $U_i(g + ij) > U_i(g)$ then $U_j(g + ij) < U_j(g)$.
2. For all $i \in N$ and any set L of edges involving i, we have $U_i(g) \geq U_i(g - L)$.[2]

The specification of what payoffs agents anticipate when contemplating deviations from g is an important modeling choice. We have made the same one as Bolletta (2021, Definition 1) (and we have confirmed in correspondence with the author that our interpretation is the intended one). An alternative assumption would be that effort choices y are fixed at the current levels (equilibrium efforts given g).[3] Some further discussion of this can be found in Section 2.1, where we show that some of the key arguments of Bolletta (2021) do not hold under either modeling choice.

Next, we define a locally complete network, which is a notion central to Proposition 1 of Bolletta (2021). This definition is adapted to the present setting from Dutta and Jackson (2013).

Definition (*Locally Complete Network*). Let $i \leftrightarrow j := \{k \in [N] \mid \theta_i \leq \theta_k \leq \theta_j\}$ be the set of all agents k such that θ_k is in the interval $[\theta_i, \theta_j]$ (including i and j themselves). A network g is called *locally complete* if, whenever $\theta_i \leq \theta_j$ and $ij \in g$, the graph g contains the complete graph on $i \leftrightarrow j$.

1.3. Result

The main result of Bolletta (2021) asserts a characterization of pairwise Nash stable outcomes as defined above. The result is stated as Proposition 1:

Every stable network is a locally complete network.

2. A counterexample to Bolletta's proposition 1

In this section, we provide a counterexample to Proposition 1 in Bolletta (2021). We consider three-player networks with the parameters, agents' equilibrium actions (y_i^*, in black) and utilities (in blue, italicized) shown in Fig. 1. We will show network $g = \{\{1,2\}, \{1,3\}\}$ is pairwise Nash stable. Note this network is not locally complete: agent 2 with the middle type is not linked to agent 3.[4]

To show pairwise stability, we consider all networks in $\mathcal{G}(\{1,2,3\})$ that can be reached by a deviation considered in the definition of pairwise Nash stability. As shown in Fig. 1, no deviation that can be effected unilaterally or by a pair of agents is profitable. Thus, g is a pairwise Nash stable network but, as we have said, not a locally complete one. This contradicts Bolletta's Proposition 1.

It is worth commenting on the proof of Proposition 1 and where it goes wrong when applied to the counterexample. A key step in the proof asserts (without argument) that if there are three agents with distinct types ordered $\theta_3 < \theta_2 < \theta_1$, and agent 1 is connected to both agents 2 and 3, then agent 2 must also want to link with agent 3. It may seem intuitive that if even the high type wants to link with the low type, then the medium type would also want to. Our counterexample shows that this need not be so: in the move $g \to g'$ in Fig. 1, agent 2 is connected with agent 1 but prefers not to link with agent 3. Intuitively, the flaw in the claim is that it overlooks the high- and medium-type agents' different neighborhoods and the impact of adding link $\{1,3\}$ on actions.

2.1. A key monotonicity lemma and a counterexample

Beyond the error we have just discussed, a key building block in Bolletta's proof of Proposition 1 is a lemma – potentially of independent interest – asserting that agents' preferences over partners are monotonic in their peers' types. More precisely, the paper's Lemma 2 states:

Consider the well-ordered set of agents $\Theta := \{\theta_1, \ldots, \theta_i, \ldots, \theta_n \mid \theta_1 < \cdots < \theta_i < \cdots < \theta_n\}$. For all agents i, $U_i(y, g + ik) > U_i(y, g + ij)$ for all $\theta_k > \theta_j$. Moreover, $y_i^*(g + ik) > y_i^*(g + ij)$, for each $\alpha \in [0, 1)$.

In the proof, a hypothetical violation of local completeness branches into two cases, with the erroneous analysis we have discussed above dealing with one case, and Lemma 2 dealing with the other case.

It is intuitive that this lemma should be useful in deducing a lot of structure in equilibrium networks: each agent wants to link with all those above a certain exogenous type, which helps in ruling out "gaps" in neighborhoods. However, in this section, we show that this lemma is also false. As we will discuss more later, this suggests that imposing the right amount of order on the formation process is trickier than it may appear.

To give the counterexample, we examine an equilibrium (called the original equilibrium) of a four-player network with the parameters and equilibrium actions shown in Fig. 2.

Consider agent 5, an outside agent with no pre-existing links and a type of $\theta_5 = 19$. Using the notation from Lemma 2, let $j = 2$ and $k = 3$, and compare the two equilibria depicted in Fig. 3. Agent 5 would have a higher utility if it linked only with agent 2 ($y_5^* \approx 15.5$) rather than if

---

[2] Equivalently, the network g is a Nash equilibrium of the Myerson link announcement game, in which all agents simultaneously announce a desired set of links, and the link ij forms if both i and j announced it.

[3] There are also other possibilities, involving more farsighted anticipation of where further deviations may lead, but Bolletta (2021) does not consider this possibility, and we will not, either.

[4] More precisely, note agent 1 and agent 3 are linked and $\theta_2 \in (\theta_1, \theta_3)$, while agent 2 and agent 3 are not linked.
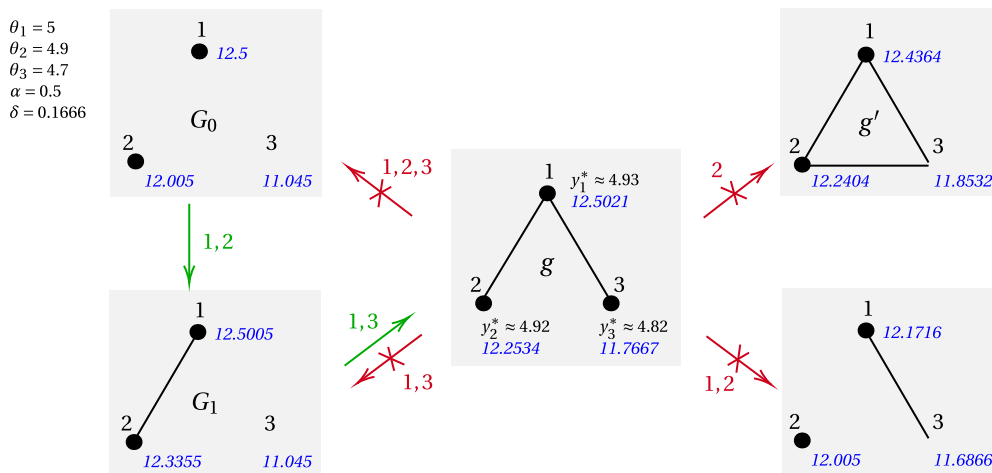
Fig. 1. A counterexample to Bolletta's Proposition 1. The payoffs of all agents (under equilibrium actions) are shown in blue, italicized text. Crossed-out, red arrows indicate moves that at least one pivotal agent (indicated in the arrow label) does not benefit from. These are used to check stability. Green arrows that are not crossed out are moves (link creations) from which the involved agents both weakly benefit. They are important to the reachability of g and are discussed in Section 3.2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

it linked only with agent 3 ($y_5^* = 15$). This contradicts the statement of the lemma.

We believe that the issue in the proof of Lemma 2 lies in the final three lines (p. 9).[5]

Now call $\overline{y}_i'$ the term such that $k \in g_i$ and $j \notin g_i$ and $\overline{y}_i$ the term such that $j \in g_i$ and $k \notin g_i$. The simple observation that $\overline{y}_i' > \overline{y}_i$ completes the proof.

In this proof, $\overline{y}_i'$ and $\overline{y}_i$ refer to the average equilibrium actions in i's neighborhood in the new networks; in more detail, these averages are equal to $\frac{1}{d_i} \sum_{j \in g_i} y_j$, for different specifications of the neighborhood g_i. According to this interpretation, the claim that "$\theta_k > \theta_j$ implies $\overline{y}_i' > \overline{y}_i$ and thus $y_i^*(g + ik) > y_i^*(g + ij)$" is shown to be false by our counterexample.

2.1.1. An alternative specification

The model we have worked with throughout defines U_i by positing that agents correctly anticipate equilibrium actions in the new network following any deviation. We have shown the lemma is false under this specification of U_i. But another specification that seems reasonable is one where agents make linking decisions *myopically*, assuming others' actions remain the same as before the link change. (This interpretation is suggested by the quoted lines of proof above, which do not seem to consider changes in equilibrium actions after link changes.) Our counterexample disproves the lemma under this interpretation, as well. Consider the original equilibrium in Fig. 2: Agent 5 would prefer to link with agent 2 ($y_2^* = 14$) rather than with agent 3 ($y_3^* = 11$), even though $\theta_3 > \theta_2$. Thus, linking incentives are not monotonic in opponent types.

3. A dynamic process

The main claims of Bolletta (2021) can be considered solely in the context of stable outcomes, as we have done so far. However, the paper also spends considerable time on a dynamic process where a network is formed by successive rounds of (pairwise consensual) link formation and (unilateral) link deletion. Natural outcomes of such a process are ones that are not only stable once formed, but reachable through the dynamic process. This section formalizes a notion of reachability and shows that our counterexample network to Bolletta's Proposition 1 satisfies it. This strengthens our main result.

---

[5] Note N_i is used in Bolletta's proof for g_i; we have replaced N_i by g_i in the quotation to avoid confusion.

Fig. 2. The original equilibrium (with $\alpha = \frac{2}{3}$ and $\delta = 75$).

3.1. A network formation process

Bolletta (2021) adopts a process from Watts (2001), as described in Definition 2:

Players meet over time $T = 1, 2, \ldots, t, \ldots$. At each period, a pair ij is selected to decide whether to form a link or sever one that already exists. Players selected this way can simultaneously sever any existing links with all $k \in g_i$ and $h \in g_j$.

The timing of each stage of the process (as presented in the paper) is:

1. *Agents form the network,*
2. *Outcomes are determined as the solution of a system of best responses.*

We remark that there is some ambiguity in the definition of the formation process because it does not specify *which* network agents think will be in effect for the purpose of determining effort choices. Will it be the network formed immediately after their deviation, or do they anticipate further deviations by others? We now fully specify the network formation process, choosing the same specification of U_i as in the stability analysis. This resolution is based on our correspondence with the author of Bolletta (2021).

Network formation process. Time proceeds in discrete instants indexed by the nonnegative integers, and $(G_t)_{t=1}^{T'}$ is a sequence of N-player networks indexed by time t. We let uppercase G denote networks appearing in the formation process, whereas lowercase g stands for arbitrary networks. We set the initial network G_0 to be the empty network. At time t, starting from network G_t, the following steps occur.

1. A pair of players i and j is (randomly) selected to act. If either player can strictly benefit by unilaterally severing some subset of
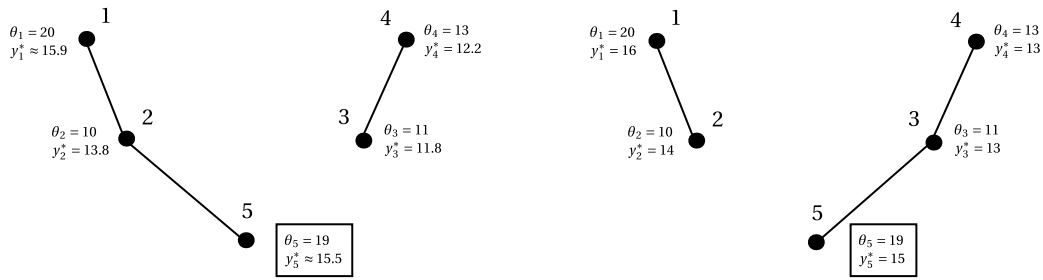
Fig. 3. A counterexample to Bolletta's Lemma 2.

her links, she does so.[6] In this case, the round ends. Otherwise, i and j form a link if doing so weakly benefits both and strictly benefits at least one.

2. A Nash equilibrium effort profile in the network G_{t+1} is played and players receive payoffs $U_i(G_{t+1})$; these are the payoffs used in determining which deviations are profitable in (1).

We say the sequence of networks $\{G_t\}$ *reaches* an equilibrium g at some finite time T if $G_t = g$ for all $t \geq T$. We call such equilibrium networks *reachable*.

3.2. The counterexample network is reachable

In order to prove that the counterexample network g is reachable, we reason as follows. Starting from the empty network, both agent 1 and agent 2 are better off by forming the link $\{1, 2\}$. Then, similarly, both agent 1 and agent 3 are better off by forming the link $\{1, 3\}$. Therefore, the network $g = \{\{1, 2\}, \{1, 3\}\}$ is reachable through the network formation process $G_0 = \{\} \rightarrow G_1 = \{\{1, 2\}\} \rightarrow G_2 = g$, illustrated by the two green arrows in Fig. 1.

4. Concluding discussion

Bolletta (2021) examines a model of endogenous network formation combined with subsequent effort choice, and the paper's Proposition 1 claims equilibria necessarily involve locally complete networks. In this note, we have presented a counterexample to that claim and identified an error in a crucial lemma used in the proof of the proposition.

More broadly, our analysis highlights that in a setting with endogenous network formation and endogenous effort choice, clean characterizations of equilibrium outcomes are difficult to obtain. Sadler and Golub (2022) makes progress on this problem, identifying conditions that are sufficient to ensure locally complete networks. In fact, some of the key arguments there work in a similar way to the above-discussed proof of Bolletta's Lemma 2, but the analysis has to be different in several ways to make these arguments valid. For example, the ordering in Sadler and Golub's definition of locally complete networks is based on agents' equilibrium *actions*, not their types.

Equally importantly, the qualitative properties of incentives in Bolletta (2021) are very different from those used by Sadler and Golub to deduce a locally complete structure. Under the separable spillovers studied in the latter paper, externalities on an agent from a neighbor's effort do not depend on the agent's degree. In Bolletta's coordination game, spillovers are averaged and do depend on one's degree.[7]

That degree-dependence makes it hard to deduce a locally complete structure. Even if we make the definition of local completeness based on equilibrium actions, rather than types (sidestepping the lack of a monotonic relationship between actions and types that we have discussed in connection with Lemma 2), Fig. 1 nevertheless shows that in Bolletta's model, stable networks are not locally complete. That is, two linked agents need not be part of a clique consisting of all the agents taking actions between theirs.[8]

In short, when links and behavior are both endogenous, the relationship between payoff structure and network structure is subtle. Local-average settings present considerable complications, while progress can be made under different assumptions. Though some of the results in Bolletta (2021) are incorrect, in our view his paper highlights both the importance and the difficulty of the kind of problem it studies.

Data availability

No data was used for the research described in the article.

References

Badev, A., 2021. Nash equilibria on (un)stable networks. Econometrica 89 (3), 1179–1206.

Bolletta, U., 2021. A model of peer effects in school. Math. Social Sci. 114, 1–10.

Boucher, V., Rendall, M., Ushchev, P., Zenou, Y., 2022. Toward a General Theory of Peer Effects. CEPR Discussion Paper No. DP17315.

Bramoullé, Y., Kranton, R., 2016. Games played on networks. In: The Oxford Handbook of the Economics of Networks. Oxford University Press.

Carrell, S.E., Sacerdote, B.I., West, J.E., 2013. From natural variation to optimal policy? the importance of endogenous peer group formation. Econometrica 81 (3), 855–882.

Dutta, B., Jackson, M.O., 2013. Networks and Groups: Models of Strategic Formation. Springer.

Galeotti, A., Goyal, S., Jackson, M., Vega-Redondo, F., Yariv, L., 2010. Network games. Rev. Econom. Stud. 77 (1), 218–244.

Hiller, T., 2017. Peer effects in endogenous networks. Games Econom. Behav. 105, 349–367.

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

Jackson, M.O., Zenou, Y., 2015. In: Young, H.P., Zamir, S. (Eds.), Games on networks. In: Handbook of Game Theory with Economic Applications, vol. 4, Elsevier, pp. 95–163.

König, M., Tessone, C., Zenou, Y., 2014. Nestedness in networks: a theoretical model and some applications. Theor. Econ. 9 (3), 695–752.

Liu, X., Patacchini, E., Zenou, Y., 2014. Endogenous peer effects: local aggregate or local average? J. Econ. Behav. Organ. 103, 39–59.

Sadler, E., Golub, B., 2022. Games on endogenous networks. arXiv preprint arXiv:2102.01587.

Watts, A., 2001. A dynamic model of network formation. Games Econom. Behav. 34 (2), 331–341.

---

[6] In contemplating deletions, the player assumes the network is otherwise held fixed. If there are multiple subsets that a player can profitably sever, one can be chosen arbitrarily.

[7] The distinction between total and average spillovers shows up in other related problems as well. For instance, Galeotti et al. (2010) study network games of strategic complements and substitutes, and they obtain much stronger comparative statics results in examples with additive spillovers. For more on distinctions between "local aggregate" and "local average" network games – see Liu et al. (2014) and Boucher et al. (2022).

---

[8] Agent 1 and 3 are linked, and agent 2's action is between theirs, but the three do not form a clique.