LSE Research **Online**

Hongzhong Deng, Peter Abell, Ofer Engel, Jun Wu, Yuejin Tan

The influence of structural balance and homophily/heterophobia on the adjustment of random complete signed networks

Article (Accepted version)
(Refereed)

http://eprints.lse.ac.uk

The Influence of Structural Balance and Homophily/Heterophobia on the Adjustment of Random Complete Signed Networks

Hongzhong Deng[1], Peter Abell[2], Ofer Engel[3], Jun Wu[1], Yuejin Tan[1]

[1] Department of Management, College of Information Systems and Management, National University of Defence Technology, China

[2] Managerial Economics and Strategy Group, Department of Management, London School Economics and Political Science

[3] Centre for the Analysis of Time Series, London School Economics and Political Science

ABSTRACT:

The mixed empirical support for balance theory is often explained by recourse to competing mechanisms that reduce the total degree of balance in the network. These mechanisms (such as differential popularity and subgroup hostility) may depend on exogenous properties of the nodes. This paper offers an alternative explanation to violations of balance, according to which the local/myopic nature of sign adjustment in accordance with balance may reduce the global degree of balance and impede the formation of groups, whereas other mechanisms that rely on exogenous node properties (e.g., homophily) facilitate these processes. The paper describes a set of simulations designed to study the evolution of complete signed networks under a local sign-change regime, induced by structural balance, homophily and heterophobia. Tolerance for local violation of balance and homophily is allowed to vary with a consequent impact upon the global degree of balance and group formation processes. We find the conditions under which the pressure towards local homophily and balance operate against each other so that homophily adjusts towards group formation but balance undermines this process.

KEY WORDS: Balance theory, Homophily, Heterophobia, Signed networks, Structural balance, Simulation, Evolution

1 Introduction

Social structures, which may conveniently be depicted as signed graphs or di-graphs

with positive (P) and negative (N) edges on n labelled nodes, may experience the joint

impact of both structural balance (Cartwright and Harary, 1956) and homophily/heterophobia

(McPherson et al 2001, Wimmer and Lewis, 2010). These are just two amongst other

mechanisms influencing the adjustment dynamics and eventual disposition. This paper is

concerned with such an impact in the context of an ongoing search for a parsimonious model

of group formation. It builds upon several previous papers, which analyse the role of balance

in the adjustment of randomly generated complete and incomplete structures under the

auspices of both sign change and link deletion processes (Abell and Ludwig, 2008, Deng and

Abell, 2010, Deng et al, 2012).

The general conclusions which may be drawn from these papers are first, that the local adjustment[1] of links in accordance with balance does not necessarily increase the overall degree of balance[2] in the network, this result anticipated by Dorean and Mrvar (1996) when they note that under the auspices of the balance mechanism, "as each actor is involved in many triads, changes towards balance in one triad may move other triads into imbalance". A second conclusion is that the degree of balance is not materially altered by: (1) whether the initial structure is generated randomly or rather by preferential attachment, (2) the size of the network (n > 20), (3) the degree of completion (density) when above about 20%.

The theory of balance leaves open the conditions under which balance will surrender either a single or two plus-sets (Cartwright and Harary 1956, Davis 1967). In addition though balance theory has largely been treated as deterministic, the issue arises as to how tolerant near-balance is to a small number of imbalanced 3-cycles. A theorem due to Easley and Kleinberg (2010) addresses this issue and the theory of generalized balance, formulated within the framework of block modelling (Doreian and Mrvar 2009, 2014) allows for departures from balance under the auspices of other mechanisms (see below). Davis (1967) proposed that all negative 3-cycles may well be relatively stable and if so then a completely balanced network can consist of more than two plus-sets ($k > 2$). Incomplete structures will allow for null relations internal to and between plus-sets and usually models unreciprocated (i.e. non symmetric) positive and negative links as equivalent to symmetrically related pairs.

The universal instability of one negative link 3-cycles has also been questioned (Doreian and Mrvar, 2009) permitting groups consisting of either mutual positive and negative related "mediators" with positive links to other groups which are themselves predominately negative linked to each other. This licences "relaxed balance" which in the framework of the cluster minimisation algorithm produces generalized block modelling where positive blocks can appear off and negative blocks on the main diagonal of the associate matrix. Extending this line of analysis, Doreian and Mrvar (2014) offer an explanation of relaxed balance called "Structural Balance with Differential Popularity and Mutual Dislike Block Model (SB-DP-MD Model). They found empirical support, even with limited Newcomb (1961) and Sampson (1964) data, for differential popularity (positive related mediators) and mutual hostility (negatively related mediators). This brings into prominence the issue of multiple mechanisms in the evolution of signed structures which in a rather restrictive framework of the interaction of balance and homophily is the subject matter of this paper. Indeed, Doreian and Mrvar (2014) remark, "Increasing concentration of receiving both positive and negative ties could rest on clearer perceptions of actor attributes...", i.e., mechanisms which depend on properties that are exogenous to the network structure, just like homophily and heterophobia which we study below.

Since adjustment to more than one possible stable state is predicted by the interaction of balance theory with other mechanisms (and, indeed, by balance in isolation) the conjoint theory cannot be regarded as complete without a further specification of the adjustment

processes (Abell, 2014). The dynamics could be determined by the proportion of positive and negative links formed in the early period of the evolution of the structure when balance would constrain the transitive closure of 3-cycles (Doreian and Krackhardt 2001). This would, however, require actors to anticipate the stress of imbalanced cycles and consequently to avoid them. It may be useful to conceive this early period as one where the dynamics are solely governed by pair relations progressively giving way to adjustment according to the wider structural precepts of balance. We adopt this viewpoint, and conceive balance as an adjustment mechanism of a structure, evolved for other reasons. The stress is experienced after the links are established. Clearly, dynamics could involve both adjustment and initial constrained transitive closure but in this paper we focus upon "later" adjustment. Separation in this respect is not entirely satisfactory but we attempt to justify it below.

The dynamics of both weighted and binary symmetric positive and negative links have been studied, combining a mixture of simulation and analytical results. Anatal et al (2005, 2006) have shown that in complete structures with symmetric binary links, the globally constrained adjustment of links can sometimes fail to reach balance, becoming "jammed" whilst still imbalanced. Nevertheless, simulations of a local adjustment rule (globally unconstrained) which allows that a randomly selected imbalanced 3-cycle flips the sign of a constituent link, without regard to the impact on other cycles, can produce a system in a (non-absorbing), stationary state, in which half the 3-cycles are balanced and half of the links are positive, assuming the probability of flipping any link in an unbalanced 3-cycle is

constant. Moreover, the system will stay in this unbalanced, stationary state for a period of time that increases with the number of nodes n like e^{n^2}.

Kulakowski et al (2008) studied weighted links in continuous rather than discrete time under global adjustment and found a sharp transition from two plus-sets to one, with increasing initial proportion of positive links. Using weights and a global adjustment rule, Marvel et al (2011) developed a closed form solutions over a wide range of initial proportions of positive and negative links, finding that balance is achieved in finite time though unrealistic weightings are required.

Deng and Abell (2010) studied local adjustment as a Markov process with varying transition probabilities between the four types of 3-cycles. The density of links, initial proportion of positive links and network size were allowed to vary. Local adjustment was not found to increase the degree of balance beyond a stationary state in which 50% of the 3-cycles were balanced, unless the transition probabilities reflected a strong bias towards positive links when one plus-set (k=1) is favoured (Anatal et al 2006). These results are immune to a range of network sizes above n=20 and densities larger than 0.2. However, below these values the size of the fluctuations is relatively large thanks to the discrete nature of the system. Groups are formed according to the structural theorem and the system grinds to a halt upon reaching global balance, which is an absorbing state.

The picture that emerges is one whereby balance is far from an open and shut matter and where if it is operative it will likely be attended by other mechanisms. However, many of the

empirical studies are cross-sectional in nature with limited over time observations (Newcomb 1961, Sampson 1968) and it is somewhat unclear whether, at the point(s) where the structure is studied, it has achieved a steady state or not. Doreian and Mrvar (2009, 2014) demonstrate that the global degree of balance is over time increasingly less dominant over the other structures they trace. Whether in the longer term these trends are monotonic is an unresolved issue.

Although it falls well beyond the ambitions of this paper, the current state of the literature suggests that the interaction of all of, at least, the following mechanisms will eventually need to be incorporated into both the formation and adjustment of signed networks: reciprocity, mutuality, completion (density), preferential attachment, balance homophily/heterophobia, differential popularity/unpopularity and the extent to which groups become cognised by network participants. Only when multiple mechanisms are in place can we hope to achieve in simulations structures similar to those observed in empirical settings. Furthermore, the size of the network and its initial state (e.g., the proportions of positive and negative links) may also be pertinent. Although it is not our intention to re-produce real networks at this initial stage, this preliminary exploration of simulations of balance may allow us to do this in the future.

Homophily (Mcpherson et al., 2001) disproportionately generates positive links between pairs of identically exogenously and fixed labelled nodes, though the literature rarely designates the sign of the links. In the context of signed structures we may add that heterophobia will generate negative links between exogenously labelled nodes which are

differently labelled. The exogeneity could be dropped allowing for the co-evolution of node attributes and links.

A homophilic and heterophobic (where no confusion is caused we shall use the term homophily to cover both tendencies) structure is balanced, though a balanced structure is not necessarily "h-stable" since it can contain groups, predominately positively linked nodes, which can be differently labelled – that is to say, heterogeneous groups. It is an elementary matter to show that homophily/heterophobia entails balance but not the reverse. Thus, a drive towards homophily is also in the direction of balance but a drive to balance does not necessarily move a structure towards homophilic groups.

Given the above acknowledged possible complexity of interacting mechanisms that can be found in the evolution of signed networks and the limitations of received data sets it is rather difficult know how to proceed. Designing an empirical study could prove prohibitively expensive except perhaps for the smallest of groups. We have therefore, almost in the spirit of conducting experiments that control for various potential causes, opted for a simulation approach. We choose to start with two mechanisms which apparently are entirely compatible in terms of the first structural theorem and introduced various dynamic adjustments.

We follow the procedure adopted in previous papers, distinguishing between a relatively "fast" random generation of a structure (where only dyads are relevant, not triads) followed by "slow" adjustments, attributable to either the local balancing of 3-cycles or homophily/heterophobia of dyads, in both cases, incorporating sign adjustments of the

relations/links. In this exploratory paper we limit our attention to complete structures with 50 nodes and symmetric binary positive and negative links. The size is thus well within the Dunbar number (i.e. 150), thus giving some support to studying complete structures[3].

Starting with randomly generated structures may appear a little controversial. If complete (or indeed incomplete structures) were to be generated under the auspices of balance theory then, as we noted above, the generation process would imply that people anticipate the stress implied by imbalanced 3-cycles and consequently avoid them, leading to clear groupings (i.e. either a single or two plus-sets). However, balance theory and to a lesser extent homophily are often construed as adjustment mechanisms experienced by structures generated by other means. It is this adjustment interpretation that we here adopt. In this respect we initially assume the slow adjustment dynamics from a least constrained random initial condition. It seems reasonable to assume that positive and negative relations will initially generate structures for a myriad of local reasons, which are best conceived as random. Also restricting the analysis to complete structures was suggested by our earlier result that local balance in structures of greater than 20% completion appear, under the auspices of local adjustment, to reach a stationary state which is indistinguishable from complete structures. It is our intention to eventually study network dynamics from initial conditions containing some homophily in sparse networks. This paper is also limited to symmetric positive and negative links; that is to say, structures exhibiting complete reciprocity.

We also introduce an exogenously varying threshold of intolerance to violations of local

balance and homophily/heterophobia, the threshold varying between 0 and 1 such that the former value implies that balance/homophily is not operating and the latter it is totally operative. The literature on balance has largely been motivated by the, almost invariably implicit idea, whereby the adjustment to balance is an all or nothing phenomenon. Here, following earlier papers, we envisage that social systems are characteristically tolerant to degrees of imbalance and departures from homophily/heterophobia but which may, under the auspices of macro shocks or the cognition that groups are emerging (not modelled in this paper), become less tolerant. We, thus, explore the adjustment of randomly generated structures in a state space where both the balance and homophily/heterophobia thresholds independently vary from complete to incomplete tolerance and search for the effect on the degree of balance $\beta(3)$ of the structure at each point in this state space. Although not explicitly concerned with group formation Wimmer and Lewis (2010), find that "reciprocity and triadic closure... amplify racial homophily effects....". This important result is discovered, however, in the absence of any consideration of the role of negative links (heterophobic choices) which are, of course, rather difficult to study empirically.

The paper proceeds as follows. Section 2 introduces the random complete network model. Section 3 explores the adjustment of structures under a local threshold sensitive balance rule. Section 4 carries out a similar analysis for homophily/heterophobia. Section 5 explores the alternating impact of balance and homophily/heterophobia and section 6 their joint impact, with two exogenously labelled groups. Section 7, studies the impact of group

size. Section 8, introduces distributed balance thresholds. In section 9 multiple groups are briefly studied. Finally, the paper concludes with some remarks about the significance of the studies for a theory of group formation in signed structures.

2 A Complete network

A complete structure on n nodes will have $n(n-1)/2$ symmetric edges. The edges are initially composed of a proportion (α_0) of randomly distributed symmetric positive edges and $(1-\alpha_0)$ negative edges. In a complete structure, the initial expected proportions (p_-) of each type of 3-cycle is given by:

$$\begin{cases} p_0 = (1-\alpha_0)^3 \\ p_1 = 3(\alpha_0)(1-\alpha_0)^2 \\ p_2 = 3(\alpha_0)^2(1-\alpha_0) \\ p_3 = (\alpha_0)^3. \end{cases}$$

Then the expected degree of balance in a randomly generated structure is given by:

$$\beta(3)_0 = p_1 + p_3 = 3\alpha_0(1-\alpha_0)^2 + (\alpha_0)^3. \ (1)$$

As the adjustment of the structure takes place, and α varies, $\beta(3)$, (the proportion of balanced 3-cycles in the structure) provides a suitable measure of the degree of balance and the propensity to group formation (courtesy of the first and second structural theorems, Doreian and Mrvar,1996)[4].

3 A local adjustment rule with varying tolerance to imbalance

Local, rather than global, adjustment rules are probably more consistent with the assumed cognitive limitations of individuals. By local we mean that the adjustment of links in imbalanced 3-cycles (either by link deletion or sign change) occurs in specific 3-cycles

without taking into account the impact upon the other cycles which the chosen link subtends. Thus, adjustments in a particular cycle, whilst relieving local stress, can have the effect of increasing the total imbalance in a structure. Such a rule does not necessarily increase the degree of balance nor to a group formation (Anatal et al 2006, Radicchi et al 2007, Deng et al 2010), but rather to a stationary state where $\beta(3) = 0.5$ and $\alpha = 0.5$, a state that is therefore inconsistent with the structural theorems. Deng et al (2012) have, however, explored conditions in which individuals take into account a growing number of their links (i.e., rules of an increasingly global nature) and discovered that under certain conditions, the structure adjust into a state which fully supports the first structural theorem (creating either one or two groups, depending upon the initial proportion of positive links). Nevertheless, these conditions would still, in all probability, over-tax the cognitive limits of individuals. We, therefore, retain a local rule which might appropriately be described as one of global estimation and local adjustment, as follows: we envisage that individuals (nodes) are capable of making a rough estimate of whether or not their global tolerance to imbalance is exceeded and, if so, they adjust but only in a most local and incremental manner.

Furthermore, since the local adjustments are likely to be very idiosyncratic they are best modelled as appearing at randomly selected vertices/nodes. Although it may be possible for actors to impose a less incremental adjustment rule in their response to stress, it seems analytically sensible to start with the least demanding adjustment. The process we are envisaging does of course require actors/nodes to make some global estimate of their

imbalance (to compare with their thresholds). This is akin to an estimate of a cross level contextual effect in multilevel analysis (Snijders and Bosker, 2012). Accordingly, in this paper, we model the adjustment of randomly generated complete signed structures in the following manner:

B1. Set all the nodes with an identical balance threshold, $B_{threshold}$. Where $B_{threshold} = 0$ when the node is entirely tolerant to imbalance and $B_{threshold} = 1$ when it is entirely intolerant to imbalance.

B2. Select at random a single node from the network. Calculate the node's balance ratio as follows:

$$B_i = \frac{\Delta_i^+}{\Delta_i^+ + \Delta_i^-} \quad (2)$$

Δ_i^+ and Δ_i^- are the numbers, respectively, of balanced and imbalanced 3-cycles running through the vertex i. In a complete connected network, the total number of triangles is

$$\Delta_i^+ + \Delta_i^- = \frac{(n-1)(n-2)}{2}$$
$$\Rightarrow B_i = \frac{2\Delta_i^+}{(n-1)(n-2)} \quad B_i \in [0,1]$$
$$(3)$$

B3. If $B_i \geq B_{threshold}$, for the selected node i go back to step 2.

B4. If $B_i < B_{threshold}$, select one 3-cycle, $\Delta_{selected}$, randomly from all those that run through the vertex i.

B5. If $\Delta_{selected}$ is a balanced 3-cycle, return to step 4

B6. If $\Delta_{selected}$ is an imbalanced 3-cycle, randomly select one edge from the two edges incident at vertex i and change its sign.

Thus, the individual, who possesses an exogenous tolerance level to imbalance (which could (see below) be allowed to vary across both individuals and time) is able to make an estimate as to whether or not her point imbalance exceeds her threshold. There is consequently an assumed ability on the behalf of individuals to make these sorts of rough calculations[5].

3.1 The simulation of network adjustment under the imbalance tolerant threshold rule

In order to study the effect of the threshold upon the resulting stationary state of the network, we simulate the adjustment process for varying thresholds. Fig 1 gives typical results[5] where the initial value of α_0 takes, respectively, selected values ranging from 0.1 to 0.9 and in each case n = 50 .
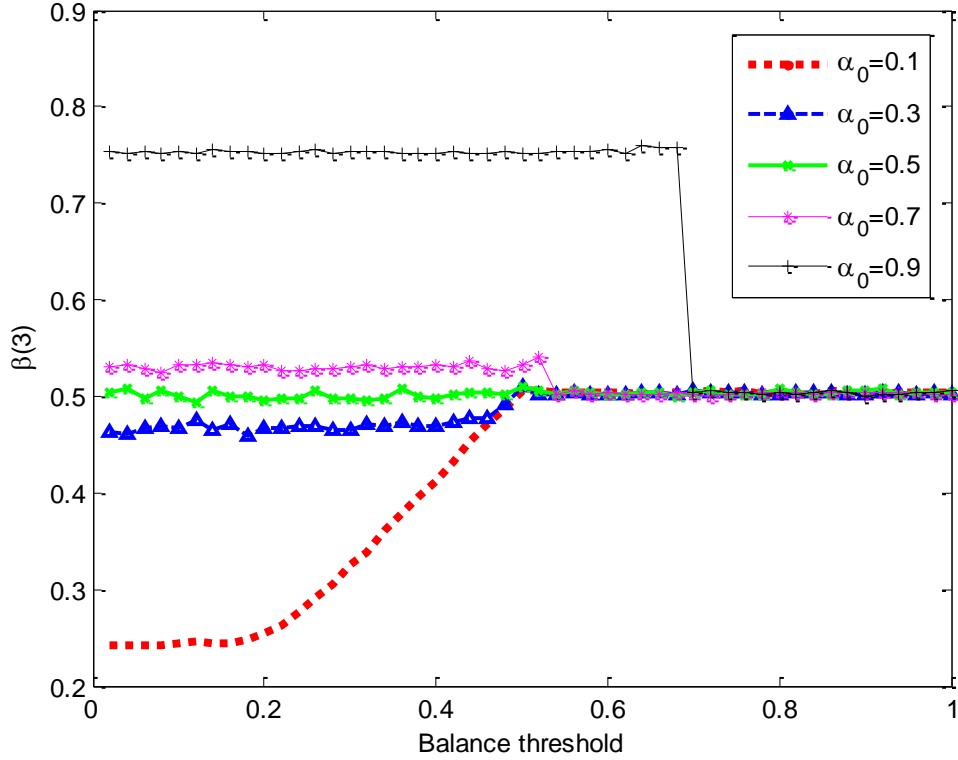
Figure 1: The stationary degree of balance $\beta(3)$ in complete networks $(n = 50)$ with varying balance thresholds

The results[5] in Figure 1 appear to demonstrate that when the initial proportion of positive links is small (i.e. $\alpha_0 = 0.1$) and the threshold is less than approximately 0.2 then the proportion of balanced cycles is close to:

$$\beta(3)_0 = 3\alpha_0(1-\alpha_0)^2 + (\alpha_0)^3 = 0.244.$$

The threshold is of a low enough value so as not to disturb the initial random proportions of the various types of 3-cycles expected given the initial proportion of positive links. With increasing value of the threshold, however, nodes become labile and the degree of balance in the stationary state increases steadily to $\beta(3) = 0.5$ when the threshold reaches about 0.45.

This degree of balance at the stationary state is familiar from previous studies (Anatal et al 2006, Deng et al, 2010, 2012). When the initial proportion of positive links is high (i.e. $\alpha_0 = 0.9$) the initial proportion of balanced cycles is about:

$$\beta(3)_0 = 3\alpha_0(1-\alpha_0)^2 + (\alpha_0)^3 = 0.756.$$

Despite the increasing threshold (intolerance to imbalance), the overall degree of balance is maintained until a threshold of about 0.7 is reached, above which there is an abrupt decline of the degree of balance, once again procuring the value $\beta(3) = 0.5$. The reason for this is that with the threshold in excess of about 0.7, the adjustment of a given node will probably involve the introduction of a negative link causing many other nodes to become imbalanced.[6]

This result is we believe of some significance. We have a situation where the intolerance to imbalance increases under the auspices of a local adjustment rule, yet the global balance at the stationary state shows a precipitous decline. This is the opposite of what the proponents of balance theory might have anticipated, though note that a global adjustment rule would have produced an increase in balance (Deng and Abell 2010). But as we have argued global rules are entirely unrealistic in largish structures. The lesson we draw is that not only, are adjustment dynamics prone to exhibit strong discontinuities but also that under some circumstances (i.e. high proportions of positive links), increases in the propensity to balance can reduce the global degree of balance at the stationary state. Note that by definition, the stationary state is not static but dynamic implying a rotation of 3-cycle types.

When α_0 lies between the extreme values so far described the threshold has little impact

upon the dynamics of the structure, which approaches $\beta(3)=0.5$ at very low values of the threshold and remains at this value as the threshold increases. This implies, of course, that as the initial proportion of positive links moves in either direction away from 0.5, abrupt changes will take place in the stationary value of $\beta(3)$ at low threshold values.

It is important to acknowledge that the results contained in Fig 1 apply to "large" structures (i.e. n=50). In much smaller structures (n in the range of 10), fluctuations are (relatively) larger and global balance can be obtained in simulations. For example, when α_0 is above 0.8 we obtain one plus-set, or two groups when below this value. We draw the conclusion that under local adjustment rule in small structures balance is achieved within a reasonable timeframe, the relatively high fluctuations across simulations guaranteeing this result. However, in larger structures, thanks to the small fluctuations, the chance of observing balance in a reasonable number of adjustments is vanishingly small.

4 Local adjustment under homophily/heterophobia

Homophily is apparently a common phenomenon relating to race gender and many other characteristics as revealed in a multitude studies (Mcpherson et al 2001 for a review). It has, however, largely been studied in the absence of signed relations/links and, therefore, not incorporating heterophobia. Separate measures of homophily and heterophobia could easily be devised but here we combine the two and still, for reasons of brevity, refer to homophily. Accordingly, complete randomly generated signed structures were allowed to adjust towards homophily with a threshold as follows:

H1. Exogenously label all nodes as either As or Bs.

H2. Set an identical homophily threshold for all nodes, $H_{threshold}$. When $H_{threshold} = 0$ then the node is entirely tolerant to lack of homophily, when $H_{threshold} = 1$ the node is entirely intolerant.

H3. Select at random one node j from the network. Calculate its homophily ratio,

$$H_j = \frac{Es_j^+ + Ed_j^-}{(n-1)} \quad (4)$$

Es_j^+ is the number of positive links to identical labelled nodes incident at node j.

Ed_j^- is the number of negative links to alternatively labelled nodes incident at node j.

H4. If $H_j \geq H_{threshold}$, the selected node j do nothing, go step 2.

H5. If $H_j < H_{threshold}$, select at random a link subtended by j which is not homophilic and change its sign.

Clearly, the impact of homophilic adjustment is fundamentally different from balance adjustment. Assuming symmetric links, any chosen node's homophilic adjustment will enhance the homophily of both the involved nodes. Local adjustment is, thus, globally optimal, unlike local balance adjustment where the adjustment of a link in a given 3-cycle can increase the imbalance in other cycles.

4.1 The influence of the homophily threshold upon network evolution

In order to study the impact of the homophily threshold upon the degree of balance

$\beta(3)$ in the stationary state, simulations were run with randomly generated complete structures containing 50 nodes and with the initial sizes of the subgroups (exogenously labelled as A nodes and B nodes) varying. We report the cases where groups A and B are of the same size and where there are 10 As and 40 Bs. The initial proportion of positive links α_0 also varies and we report respectively (low) 0.1 and (high) 0.9 values (i.e. the values that demonstrate a significant variation in impact in the case of balance). The results are depicted in Figures 2 and 3.
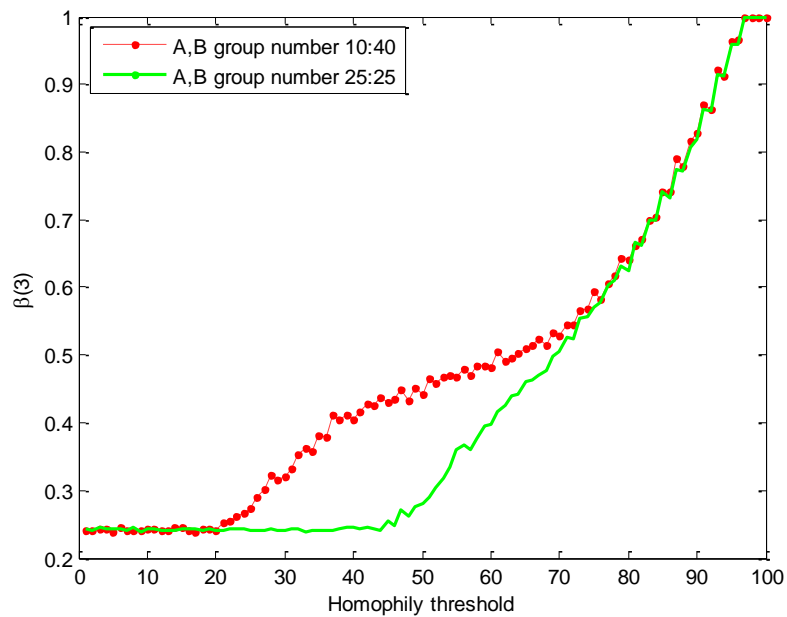

Figure 2. The influence of the homophily threshold (intolerance to deviation from homophily) upon the degree of balance in a complete network at a stationary state $(n = 50, \alpha_0 = 0.1)$.
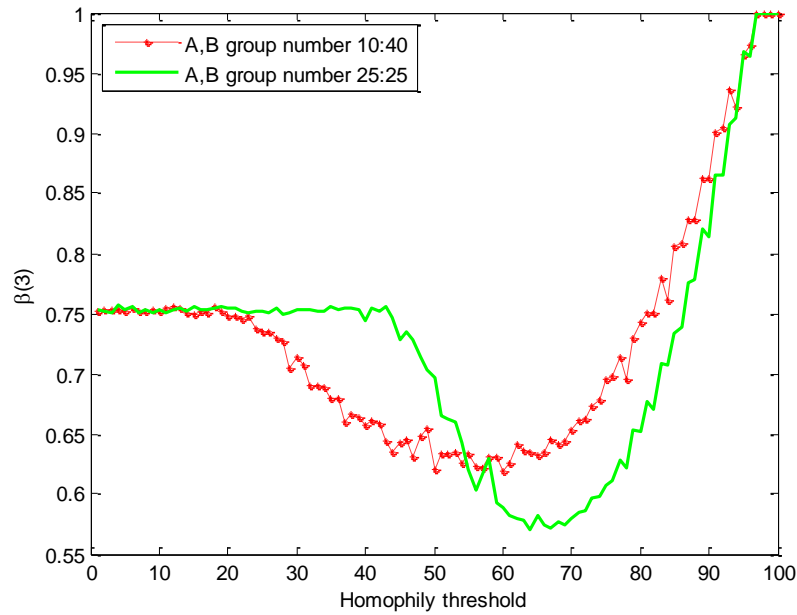
Figure 3. The influence of the homophily threshold upon the degree of balance in a complete network at a stationary state $(n = 50, \alpha_0 = 0.9)$.

The results in Figures 2 and 3 demonstrate that:

1. The curves for the differing subgroup sizes are very similar in shape.

2. When the initial ratio of positive links is $\alpha_0 = 0.1$, the groups are set at 10/40 and the value of homophily threshold is less than about 0.2, the degree of balance reaches about 25%; homophily has little impact and the degree of balance $\beta(3) = 3\alpha_0(1-\alpha_0)^2 + (\alpha_0)^3 = 0.24$, the value we expect in the initial random structure. When the subgroups are of equal size then the turning point is, by similar calculation, at about 0.50. Randomly expected homophily is sometimes dubbed (Mcpherson et al 2001) as base–line homophily and any departures as inbreeding. Here, however, our parameter covers both base-line homophily and heterophobia. Although we do not pursue the idea it

would clearly be of interest to chart base- line homophobia and heterophobia separately.

3. When the value of the homophily threshold exceeds about 0.2 (unequal group sizes) and 0.5 (equal group sizes) the degree of balance increases steadily with increase in the homophily threshold. Note, that in contrast to the local balance dynamics, the adjustment eventually, reaches complete balance. The groupings so formed are, of course, homogeneous, comprising either all As or all Bs. This property clearly follows from the consistency of local and global dynamics in the case of homophily. We can at this juncture, though, anticipate inherent complexities when a structure is adjusting under the impact of both balance and homophily/heterophobia (see below).

4. When the initial ratio of positive links in Figure 3 is set high at $\alpha_0 = 0.9$ and the value of the homophily threshold is less than about 0.23 (unequal group sizes) and 0.45 (equal group sizes), the degree of balance reaches a constant at about 75%. The reason being that the initial homophily value of the nodes in subgroup A is, from equation (1), about 0.75 and in sub group B, about 0.737. The threshold, within this range, has little impact upon a node's adjustment and the randomly generated structure will to a significant degree approximate two homogeneous groups.

When the values of the homophily thresholds exceed 0.23/0.45 but are less than 0.55/0.65, the ratio of balanced 3-cycles decreases. The reason for this can perhaps be best appreciated by taking a numerical example for unequal group sizes. For example, if the ratio of positive links in subgroup A is 0.4, the ratio of negative links between subgroups

A and B is also 0.4 and the ratio of positive links in subgroup B is 0.9 then balanced 3-cycles are of four kinds: AAA with all three nodes in subgroup A; BBB with all three nodes in subgroup B; ABB and AAB each with a positive and two negative edges. The proportions of balanced cycles of these four kinds are then as follows:

$$
\begin{aligned}
p_{AAA-balance} &= p_1 + p_3 = 3\alpha(1-\alpha)^2 + \alpha^3 = 0.496 \\
p_{BBB-balance} &= p_1 + p_3 = 3\alpha(1-\alpha)^2 + \alpha^3 = 0.756 \\
p_{ABB-balance} &= p_1 + p_3 = (0.6 \times 0.6 \times 0.9 + 2 \times 0.4 \times 0.6 \times 0.1) + (0.4 \times 0.4 \times 0.9) = 0.52 \\
p_{AAB-balance} &= p_1 + p_3 = (0.6 \times 0.6 \times 0.4 + 2 \times 0.4 \times 0.6 \times 0.6) + (0.4 \times 0.4 \times 0.4) = 0.496
\end{aligned}
\tag{5}
$$

The numbers of 3-cycles belonging to kinds AAA, BBB, ABB and AAB are:

$$
\begin{aligned}
T_{AAA} &= C_{10}^3 = 120 \\
T_{BBB} &= C_{40}^3 = 9880 \\
T_{ABB} &= C_{10}^1 \times C_{40}^2 = 7800 \\
T_{AAB} &= C_{10}^2 \times C_{40}^1 = 1800
\end{aligned}
\tag{6}
$$

So, the ratio of balanced 3-cycles in whole network is:

$$
\begin{aligned}
P &= \left(p_{AAA-balance} \times T_{AAA} + p_{BBB-balance} \times T_{BBB} + p_{ABB-balance} \times T_{ABB} + p_{AAB-balance} \times T_{AAB} \right) / C_{50}^3 \\
&= 0.63
\end{aligned}
\tag{7}
$$

Equation (7), thus, accounts for the decrease of the balance to about to 0.63 with unequal sized groups (Fig 3). Parallel reasoning accounts for the similar pattern with equal sized groups.

5. When the value of homophily threshold exceeds 0.55 the degree of balance eventually reaches unity.

5 Adjustment under the sequential presence of balance and

homophily/heterophobia

In order to study adjustment under the sequential operation of balance and homophily, each with independently varying thresholds, the following procedure was adopted:

1. All nodes are ascribed an identical balance threshold $B_{threshold}$.

2. All nodes are ascribed an identical homophily threshold $H_{threshold}$.

3. Execute the balance tolerant threshold rules (B2 - B6).

4. Execute the homophily threshold rules (H3 − H5).

5. Go to step 3.

The simulated results with the total number of nodes set at 50, the initial proportion of positive links $\alpha_0 = 0.1$, and the sizes of the two potential subgroups set at 10 and 40, are depicted in Figure 4.
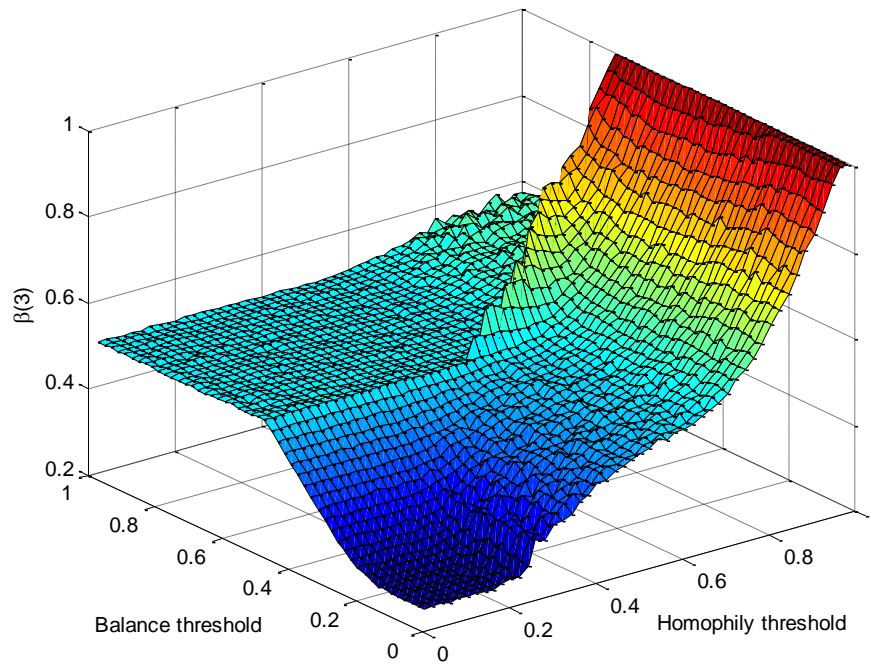
Figure 4. The stationary degree of balance in a complete network $(n = 50, \propto_0 = 0.1)$ under varying threshold regimes. The sizes of the two subgroups are 10 and 40.

The results indicate that:

1. When the values of both the balance and homophily thresholds are less than about 0.2, the degree of balance reaches about 0.24. The joint impact of the two mechanisms is modest, with little differentiation into groups.

2. Whilst the value of balance threshold remains less than about 0.5, but the homophily threshold increases beyond 0.2, the degree of balance increases steadily, eventually procuring two homogeneous groups ($\beta(3) = 1$).

3. When the balance threshold exceeds 0.5, the stationary degree of balance reaches about

0.5 as long as the threshold is less than 0.5. Balance only increases to about 60% when the homophily threshold exceeds 0.5. Furthermore, this critical value of the balance threshold coincides with a remarkably abrupt phase change. It appears, therefore, that local balance adjustment, when pitched beyond a critical threshold, actually undermines the impact of homophily preventing the formation of homogeneous groups. This is our most striking result. Whereas balance has, under the auspices of the first structural theorem, been interpreted as a mechanism which drives structures towards group formation, our results demonstrate that local balance adjustment (and it is important to emphasise the local nature of the adjustment) countermands homophily and heterophobia, and consequential homogeneous (i.e. all A or all B groups) group formation. The result is, however, established for a low initial proportion of positive links. One might reasonably conjecture that many real world structures will have a much higher initial proportion of such links.

Changing the initial ratio of positive links from 0.1 to 0.9, gives somewhat similar results, the pattern for high positive links is portrayed in Figure 5.
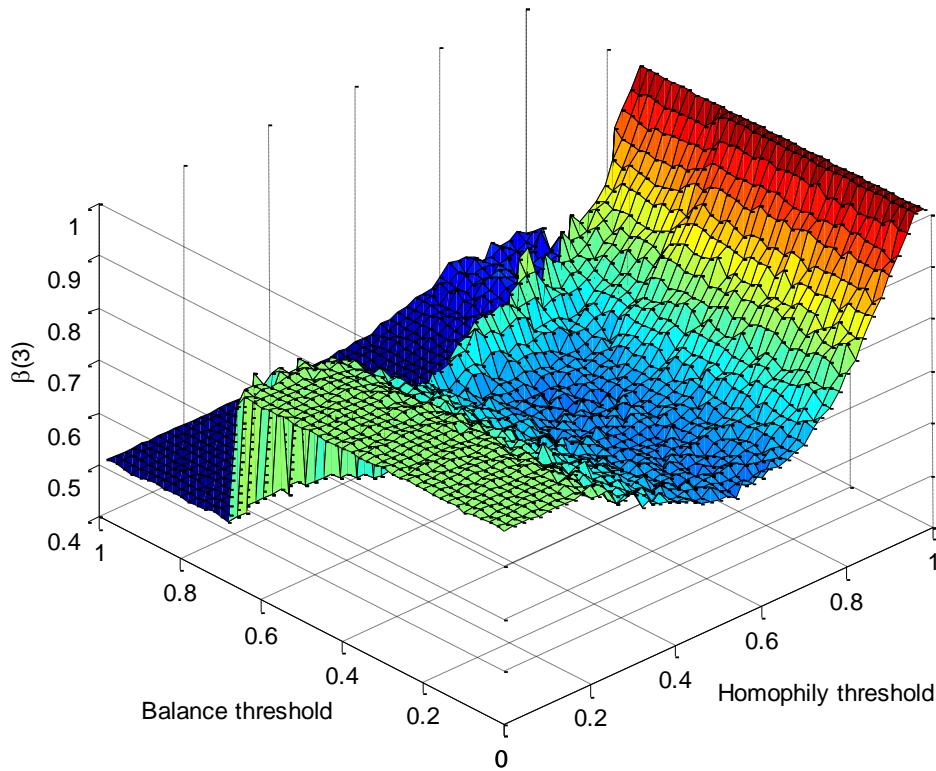
Figure 5. The stationary degree of balance in a complete network $(n = 50, \propto_0 = 0.9)$. **The sizes of the two subgroups are 10, 40.**

So, comparing the results in Figure 4 with those in Figure 5 suggests that:

1. When the value of the balance threshold is less than about 0.85 and the homophily threshold does not exceed about 0.24, the stationary degree of balance reaches about 0.75. Thus, the initial high proportion of positive links, favours a strong tendency towards groups.

2. If the balance threshold is over 0.85, whilst the value of homophily threshold is below about 0.6, the stationary value of $\beta(3)$ drops to 50% (c.f. Fig 4). When the value of

homophily threshold exceeds about 0.6, the stationary value of $\beta(3)$ increases along with the increasing value of the homophily threshold. Eventually the degree of balance will reach about 60% when the homophily threshold is 1.

3. When the balance threshold is smaller than 0.85, and the homophily threshold exceeds about 0.24, the stationary value of $\beta(3)$ first decreases along with the increases of the homophily threshold, then eventually reaches 100% when the homophily threshold is 1. Once again it appears that when α_0 is high, balance undermines homophily/heterophobia but at a higher threshold than when less positive links are initially introduced.

6 Adjustment under the joint impact of balance and homophily

The adjustment depicted in the previous section of the paper allows the operation of the balance and homophily rules to alternate with each other. In practice one might anticipate that the rules may well interact. Accordingly a joint balance and homphily rule, BH, was studied, again in randomly evolved complete signed structures.

1. All nodes are ascribed an identical balance threshold $B_{\text{threshold}}$.

2. All nodes are ascribed an identical homophily threshold $H_{\text{threshold}}$.

3. Select a node k randomly from the network.

4. If node k falls below the balance and homophily thresholds, change the sign of one edge of node k to enhance the balance and homophily ratio of the node. If the

adjustment does not enhance both the balance and homophily, select the other link incident at k and change its sign to enhance both balance and homophily. If this is not possible return to 3.

It is important to acknowledge that the balance adjustment rule is local in nature and therefore a link can adjust in a particular 3-cycle reducing balance in other cycles though homophilic adjustment will always improve the level of homophily/heterphobia. In effect balance can provide a break upon the rate at which groups emerge.

The simulated results with the total number of nodes set at 50, the initial proportion of positive links at $\alpha_0 = 0.1$, and the sizes of two potential subgroups at 10 and 40, are depicted in Figure 6.
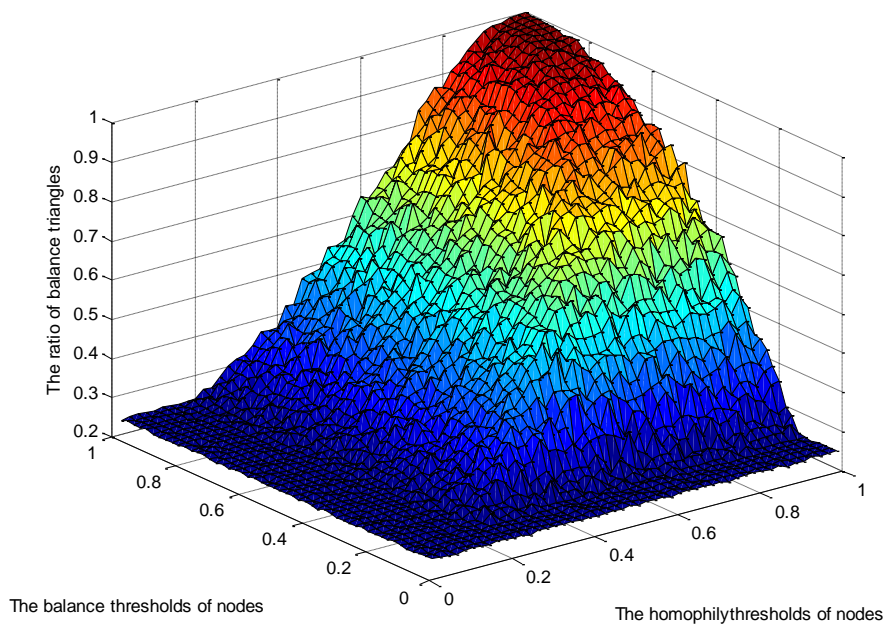
Figure 6. The stationary degree of balance under the joint impact of balance and homophily $(n = 50, \propto_0 = 0.1)$**. The sizes of the two subgroups are 10, 40.**

The results indicate that:

1. When the values of the balance and homophily thresholds are, respectively, less than about 0.2 and 0.35, the stationary value of $\beta(3)$ is about 0.25. The joint impact of the balance and homophily delivers few adjustments in the network.

2. When the values of the balance and homophily thresholds are respectively beyond about 0.2 and 0.35, $\beta(3)$ increases slowly with the increasing values of both the balance and homophily thresholds. The impact of balance has approximately the same effect as that of homophily. However, the stationary $\beta(3)$ curve appears to increase linearly with the homophily threshold though is convex with the balance threshold.

3. When the balance threshold and homophily thresholds both equal 1, $\beta(3) = 1$; the nodes divide into two homogeneously labelled groups.

Changing the initial ratio of positive links from 0.1 to 0.9 gives somewhat similar results, as portrayed in Figure 7, though the grouping is not perfect.
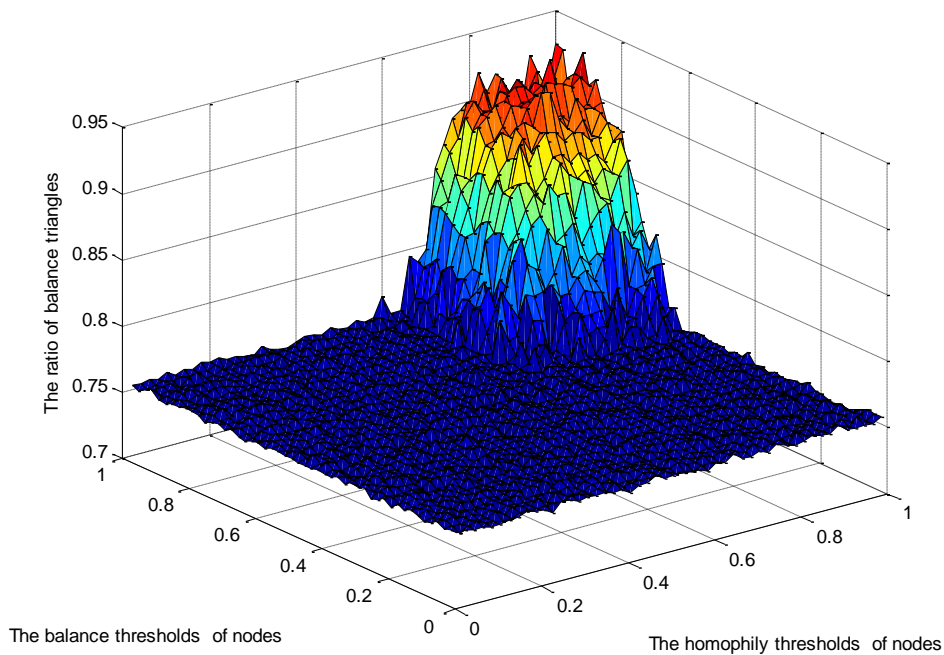
Figure 7. The stationary degree of balance in complete networks under the joint impact of balance and homophily $(n = 50, \propto_0 = 0.9)$. The sizes of the two subgroups are 10, 40.

Comparing the results in Figure 6 with those in Figure 7 demonstrates that:

1. When the values of the balance and homophily thresholds are both less than about 0.6, $\beta(3)$ becomes stationary at about 0.75. The initial structure of the randomly generated structure is largely undisturbed.

2. When the values of the balance and homophily thresholds are both beyond about 0.6, the stationary value of $\beta(3)$ increases quickly to 1. The curves are very similar to those in Figure 6.

7 The impact of sub-group size under the impact of the joint balance and homophily rules

It is clearly of interest as to whether the size of the exogenously labelled potential groups will alter the stationary degree of balance and, thus, the tendency to group formation. Accordingly simulations were conducted with many different initial group sizes. The results with equal sized groups are reported here (i.e. n = 50 with 25 As and 25 Bs) in Figs 8 and 9.
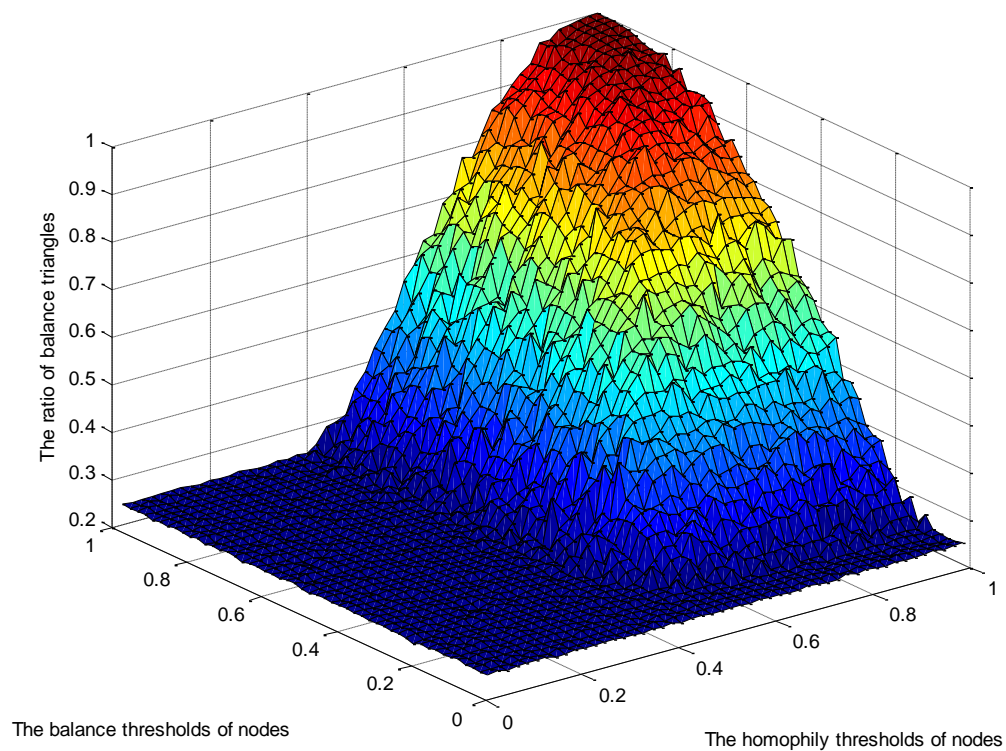

Figure 8. The stationary degree of balance in complete networks under the joint impact of balance and homophily $(n = 50, \propto_0 = 0.1)$. The sizes of the two subgroups are 25 and 25.
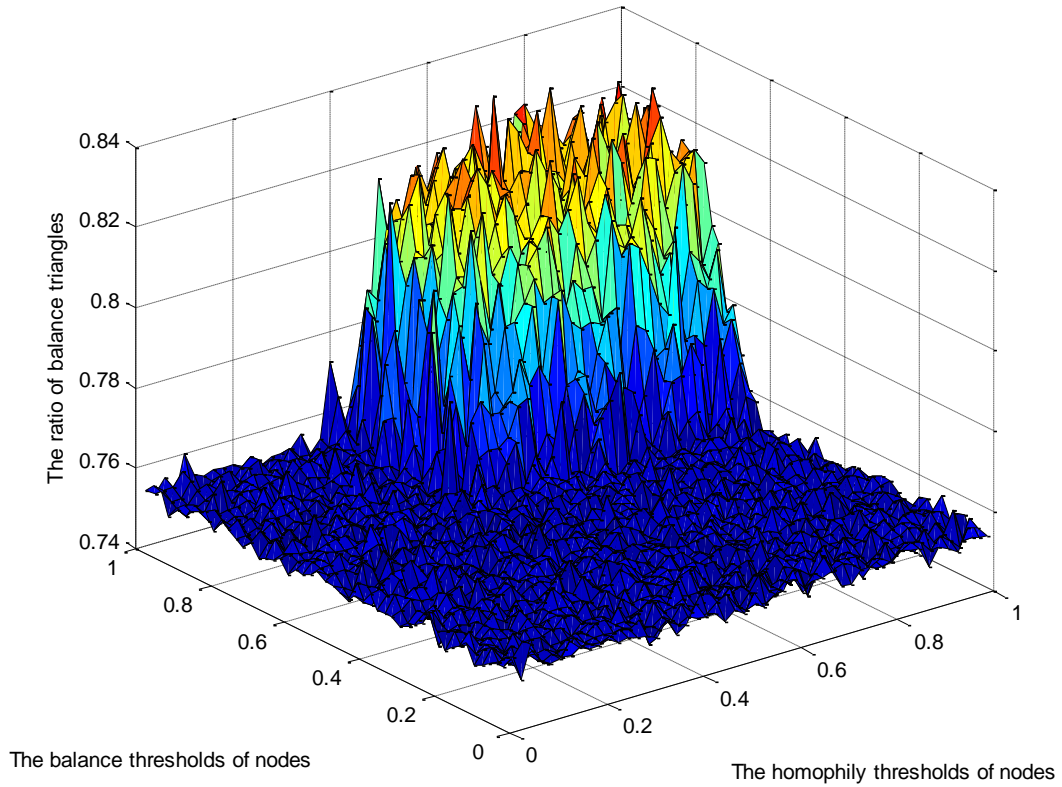
Figure 9. The stationary degree of balance in complete networks under the joint impact of balance and homophily $(n = 50, \propto_0 = 0.9)$. The sizes of the two subgroups are 25,25.

Comparing the results in Figures 6, 7, 8 and 9 demonstrate that:

1. The stationary degree of balance does not change very significantly with the size of the exogenously labelled subgroups. The shapes in the figures, controlling for α_0, are rather similar.

2. The small difference between Figures 8 with 6 (i.e. when $\alpha_0 = 0.1$) emerges when the homophily threshold is between 0.2 and about 0.5. The reason is that, with the same initial ratio of positive links, equal sized potential subgroups will have a higher initial balance

ratio than subgroups of size 10:40.

3. We find similar patterns when comparing the results in Figure 9 with Figure 7 (i.e. when α_0 is high.

8 Distributed balance and homophily thresholds

So far all nodes have been assigned, in each simulation, identical values across all nodes for both their balance and homophily thresholds. This assumption appreciably simplified the simulations but is probably an over simplification, though cultural norms may conceivably have this sort of effect. However, one might expect both types of thresholds to be distributed in value. Consequently simulations were run with uniformly distributed balance thresholds (i.e. a maximum entropy distribution) as follows:

1. If the average balance threshold, $B_{threshold}$, is less than 0.5, all the nodes' balance tolerances follow a uniform distribution in the zone $[1, 2*B_{threshold}]$. For example, if $B_{threshold} = 0.3$, the distribution zone is [0. 0.6].

2. If the average balance threshold is above 0.5, all the nodes' balance tolerances follow a uniform distribution in the zone $[(2*B_{threshold}-1), 1]$. For example, if $B_{threshold} = 0.7$, the distribution zone is [0.4. 1].

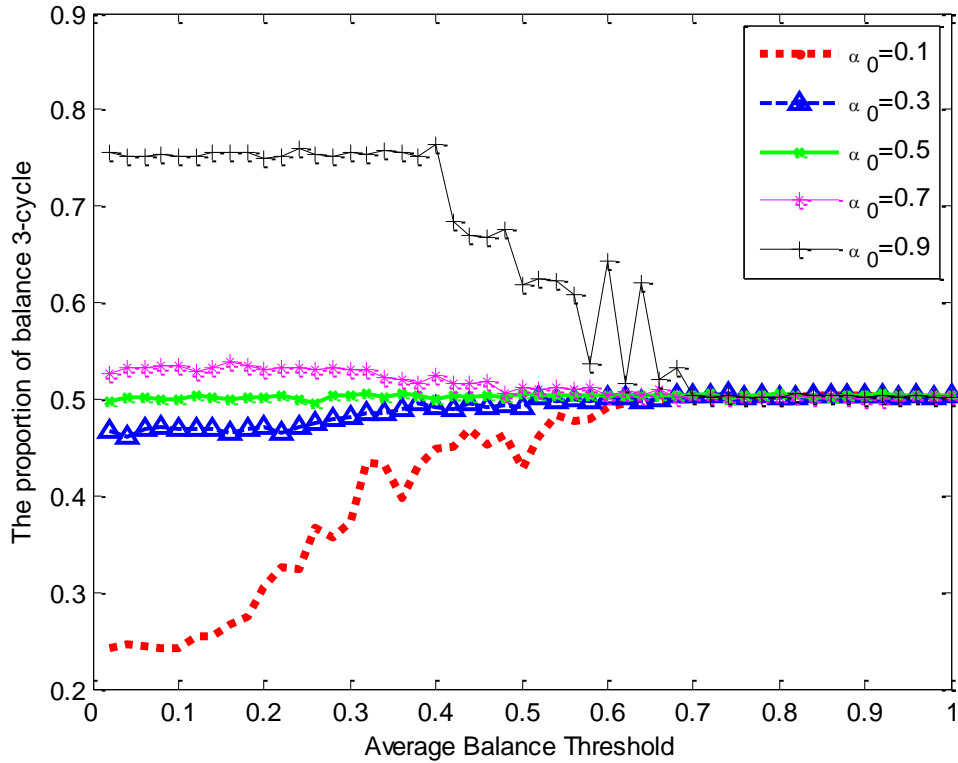With these assumptions in place the results in Fig 10 are obtained.

Figure 10. The distribution of balanced 3-cycles after reaching stationary state, complete networks $(n = 50)$ with uniformly distributed balance thresholds.

Comparing the results in Fig. 10 with those in Fig. 1 demonstrates that they are very similar indeed. The non-monotonic fluctuations in the curves (Fig.10) where, $\alpha_0 = 0.1$ and 0.9 are attributable to the relatively small n. We draw the general conclusion from this comparison and others not reported, namely that the local balance adjustment does not lead to a stationary state that depends upon the distribution of the threshold in the population.

9 The impact of multiple groups on network evolution

To determine whether a network with more than two subgroups will reach a stationary state, five equal sized exogenous labelled potential groups were studied under the joint rule and

with identical thresholds across all the nodes. The simulation results, averaged over numerous
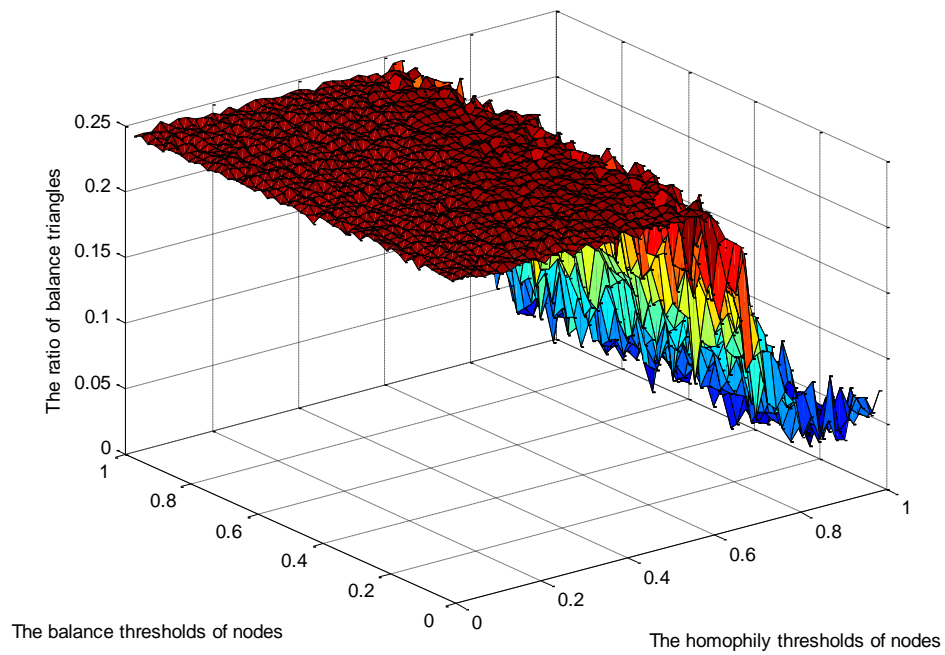
runs, are depicted in Figures 11 and 12.

Figure 11. The stationary degree of balance in complete networks $(n = 50, \propto_0 = 0.1)$ divided into five subgroups, each with identical node balance and homophily thresholds (compare with Figure 6)
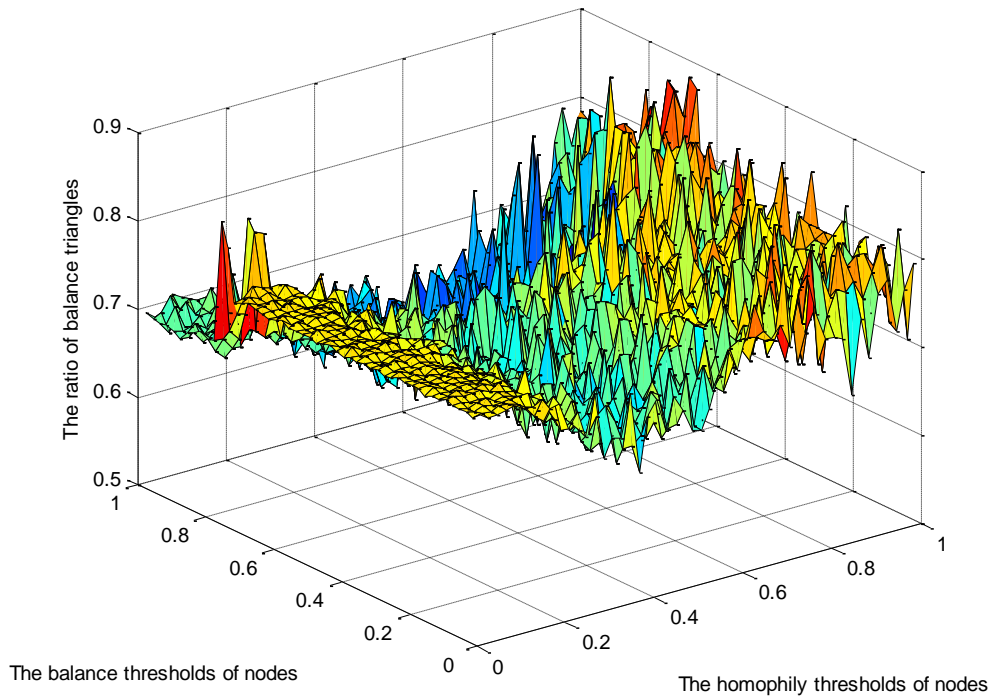
Figure 12. The stationary degree of balance in complete networks $(n = 50, \propto_0 = 0.9)$ divided into five subgroups, each with identical node balance and homophily thresholds (compare with Figure 7)

The results in Fig 11 reveal that, where the initial proportion of positive links is low and the balance threshold is lower than about 0.5, adjustments are infrequent and the stationary balance ratio of whole network is at about 25% (the initial expected value from equation (1)). Increasing the homophily threshold has no impact upon the stationary state as long as the balance threshold stays below about 0.5. In effect the tolerance to local imbalance undermines the impact homophily even when the threshold goes to unity. When the balance threshold moves above 0.5, balance declines precipitously whatever the value of the homophily threshold. Homophily cannot drive the structure into homogenous groups in the presence of significant levels of a tendency towards balance.

Comparing the results in Figure 11, with those in Figure 12, shows that the degree of balance is profoundly influenced by the initial proportion of positive links; a higher initial proportion enables the structure to form clearer group boundaries. The fluctuations are, however, quite notable in Fig 12. These results are quite intuitive. Networks with a great deal of initial hostility find it almost impossible to break into many groups (but relatively easy into two groups as thresholds approach unity (Figs 6 and 8)) whilst, if the initial proportion of positive links increases multi-grouping becomes easier (Fig12) and bipolarisation more difficult (Figs 7 and 9). It is worth recalling (footnote 6) that the results here pertain to situations where all negative 3-cycles retain their unbalanced status.

10 Conclusions

We have studied the balancing process, tracked in terms of $\beta(3)$, in randomly generated complete signed networks, containing symmetric positive and negative relations/links, under the impact of varying (though initially identical across nodes) local balance and homophily/heterophobia thresholds. The nodes were exogenously labelled as either As or Bs. The separate effects of local balance and homophily/heterophobia thresholds upon the adjustment in proportions of balanced 3-cycles ($\beta(3)$) have been studied, using simulations. The justification for addressing these issues through simulation rather than empirical study stems from their inherent complexity which, in our view, rules out premature empirical inquiry. Confirming earlier results with a slightly different local link deletion adjustment rule (Abell and Ludwig 2008), the movement to balance under local adjustment rules is markedly

non–monotonic. Thus, the impact upon group formation (courtesy of the First Structural Theorem) is correspondingly uneven and in so far as groups are formed they may, of course, contain heterogeneously labelled nodes. Homophily/heterophobia, by contrast, necessarily procures groups which contain homogeneously labelled nodes.

Consequently the joint effect of balance and homophily/heterophobia upon the proportions of balanced 3-cycles has been studied. The initial adjustment rule adopted is rather simple allowing for the sequential alternation of (local) balance and homophily/heterophobia effects, each with varying thresholds. The rules were then allowed to interact. The results are, we believe notable in the context of group formation as they point to local balance as a randomising mechanism in the presence of homophily/heterophobia. If the threshold of intolerance to imbalance is low, we expect homophily/heterophobia to dominate, leading to the formation of homogeneous groups. If, however, the threshold increases then, abruptly, the dynamics do not lead to groups but to a near random distribution of positive and negative links.

The results reported in this and in our earlier papers confirm that local balance adjustment does not provide an unequivocal theory of group formation. Furthermore, the adjustment dynamics are characterised by abrupt changes of various sorts. Our initial objective in mounting these theoretical enquiries was to ascertain what role balance theory might be allowed play, alongside other mechanisms, in a general theory of group formation in signed structures. Our conclusion in this respect is, at this juncture, one of some considerable

caution. Empirical studies which observe structures at one or a few points in time may well prove misleading, as results will depend upon where in the complex dynamical process measurements are made. Clearly the timing of adjustment becomes crucial –something we hope to take up at a later date. The interaction of balance and homophily is (Wimmer and Lewis, 2010) accordingly complex. If homophily were to be complete in a structure then balance is guaranteed. If, however, homophily (inbreeding) is only a tendency then local balance effects can undermine homophily.

11 Footnotes

[1] By local adjustment we imply that links change sign (or, in the case of deletion dynamics, are severed) in particular imbalanced 3-cycles without taking account of the role of the link in any of the, up to $n - 2$, other 3-cycles in a complete structure. We will argue that local adjustment is probably more realistic than global adjustment. We also concentrate upon 3-cycles as our evidence suggests that larger cycles are not unstable. However, computing balance in terms of 3-cycles only and cycles of all lengths are identical procedures in complete structures.

[2] The degree of balance is given by the ratio balanced 3-cycles (triangles) to total 3-cycles in a structure. Balanced 3-cycles contain either one or three positive links.

[3] The Dunbar number is conjectured as the upper limit of the number of relations human beings can maintain. This probably suggests that incomplete structures become more pertinent

above this number. In selecting complete structure with n set at 50, each node will need to maintain 49 relations.

[4] The first structural theorem states that a balanced structure consists of either one group in which all connected nodes are positively linked, or two such groups (plus-sets) with negative links running between them. The second structural theorem assumes that all negative linked 3-cycles are balanced (stable) and balancing structures (containing negative links) consist of more than two positive linked groups with negative links between them.

[5] The above described adjustment rule is entirely consistent with the results for similar local rules previously reported (Deng et al 2010). When the transition probabilities between the 3-cycle types constitute an irreducible Markov process then the stationary distribution is independent of the starting point; It could be assumed that the individual (node) would adjust the link which most improves its balance. But, once again, in anything other than very small structures this would imply a calculative ability which is incredible.

[6] A certain caution is necessary in interpreting this conclusion. With a different local sign change adjustment rule in place (Deng et al, 2012), when $\alpha_0 > 80\%$ and the intolerance to imbalance is total ($\beta_{threshold} = 1$) the structure can reach global balance $\beta(3)=1$.

[7] With multiple groups it is often assumed that all negative 3-cycles are balanced/stable when the second structural theorem applies. In a subsequent paper we shall explore this possibility but here we stay with the standard interpretation of balance. Clearly, the instability of all negative 3-cycles will allow for formation of more than two groups.

12 References

Abell, P. (2014). "Recent Developments in Structural Balance Theory" *International Encyclopedia of the Social and Behavoiural Sciences.* Elsevier, Forthcoming.

Abell, P. and M. Ludwig (2008). "Structural Balance: A Dynamic Perspective", *Journal of Mathematical Sociology,* 33, 1-27.

Anatal, T., Krapivsky, P.L. and S. Redner (2005). "Dynamics of Social Balance on Networks", *Physical Review, E,* 72,036121.

Anatal, T., Krapivsky, P. L., & Redner, S. (2006). "Social balance on networks: The dynamics of friendship and enmity". *Physica D: Nonlinear Phenomena*, 224(1), 130-136.

Cartwright, D., C., and F. Harary (1956). "Structural Balance: A Generalisation of Heider's Theory" *Phsychological Review,* 63, 277-292.

Davis, J., A. (1967)."Clustering and Structural Balance in Graphs", *Human Relations*,20, 181-187.

Deng, H. and P. Abell (2010). "A Study of Local Sign Change Adjustment in Complete Balancing Structures", *Journal of Mathematical Sociology,*34, 253-282.

Deng, H., Abell, P., Li, J., & Wu, J. (2012). "A study of sign adjustment in weighted signed networks". *Social Networks*, 34(2), 253-263.

Doreian, P., and D. Krackhardt (2001). "Pre-transitive balance mechanisms for signed networks", *Journal of Mathematical Sociology*, 25(1), 43-67.

Doreian, P. and A. Mrvar (1996). "A Partitioning Approach to Structural balance", *Social Networks,* 18,149-168.

Doreian, P. and A. Mrvar (2009). "Partitioning Signed Social Networks", *Social Networks*, 31, 1-11.

Doreian, P., and A. Mrvar (2014). "Testing Two Theories for Generating Signed Networks Using Real Data", *Advances in Methodology & Statistics/Metodoloski zvezki*, 11(1).

Easley, D. and Kleinberg, J. (2010). *Networks, Crowds and Markets,* Cambridge, Cambridge University Press.

Kulakowski, K., P. Gawronski and P. Gronek (2008). "The Heider Balance: a Continuous Approach", *International Journal of Modern Physics,* C,16, 707 -716.

Ludwig M. and P. Abell (2007). "An Evolutionary Model of Social Networks", *European Physical Review B,* 58, 97-105.

Marvel, S.A., J. Kleinberg, R.D., Kleinberg and S.H. Strogatz (2011)."Continuous -time Model of Structural Balance .*Proceedings of the National Academy of Sciences,*108, 1771-1776.

McPherson, M., L. Smith–Lovin and J. M Cook, (2001). "Birds of a Feather: Homophily in Social networks", *Annual Review of Sociology,* 27, 415-444.

Newcomb, T. M. (1961). *The acquaintance process,* (Holt, Rinehart and Winston, New York)

Sampson, S. F. (1968). *A novitiate in a period of change: An experimental and case study of social relationships* (Doctoral dissertation, Cornell University, September).

Radicchi, F., Vilone, D., Yoon, S., & Meyer-Ortmanns, H. (2007). "Social balance as a satisfiability problem of computer science". *Physical Review* E, 75(2)

Snijders, T.A.B. and R.J. Bosker, (2012) *Multilevel Analysis,* Sage. London.

Wimmer, A. and K. Lewis, (2010). "Beyond and Below Racial Homophily: ERG Models of a Friendship Network Documented on Facebook", *American Journal of Sociology*,116, 583-642.

Acknowledgement