Social Networks 34 (2012) 682–690

Contents lists available at SciVerse ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Network structure and team performance: The case of English Premier League soccer teams

Thomas U. Grund*

University of Oxford, Nuffield College, New Road, Oxford OX1 1NF, United Kingdom

ARTICLE INFO

Keywords:
Network structure
Centralization
Team performance
Soccer

ABSTRACT

A defining feature of a work group is how its individual members interact. Building on a dataset of 283,259 passes between professional soccer players, this study applies mixed-effects modeling to 76 repeated observations of the interaction networks and performance of 23 soccer teams. Controlling for unobserved characteristics, such as the quality of the teams, the study confirms previous findings with panel data: networks characterized by high intensity (controlling for interaction opportunities) and low centralization are indeed associated with better team performance.

1. Introduction

Although there is a consensus regarding the concept that a team is more than the sum of its parts, researchers focus on very different factors to explain why some teams are more successful than others. Some accounts stress the importance of the individual abilities and knowledge of group members, while others focus on group identification and consciousness or on leadership and the organization of work (see, e.g., Kozlowski and Bell, 2003; Sanna and Parks, 1997).

A growing body of evidence links the structural properties—e.g., network centrality—of interactions between group members to performance outcomes (see the reviews in Balkundi and Harrison, 2006; Katz et al., 2004; Flap et al., 1998; Borgatti and Foster, 2003). It has been argued that the orchestration of interactions and the relationships between team members are pivotal for team performance. The relevant unit of analysis is therefore the dyad between team members and not individual team members, per se. Such a social network approach suggests that interaction patterns matter for success. One team will be better than another because the individuals in that team interact in ways that members of the other team do not.

The rationale for this proposition is straightforward: some tasks require the involvement of different individuals or a combination of resources. Therefore, relationships between team members are important because they allow access to resources and facilitate the successful mobilization of these resources (Brass, 1984; Ibarra, 1993). Other researchers suggest that the structural properties of interaction and relationship patterns in teams are related to social expectations, identity, and support (Podolny and Baron, 1997). In addition, these researchers provide insight into unobserved team characteristics, such as group cohesion or the integration of individual members (see, e.g., Baldwin et al., 1997).

A meta-analysis by Balkundi and Harrison (2006: 59) summarizes previous findings on the relationship between within-team network structure and team performance in the following way: teams with denser networks tend to perform better and remain more viable. Additionally, centralized network structures are found to be negatively associated with team performance (Cummings and Cross, 2003).

Drawing on the innovative setting of team sports, this study overcomes some difficulties of previous research and investigates the interaction network and performance of professional soccer teams in the English Premier League (EPL) using panel data. A dataset of 283,259 passes between individual players in 760 soccer matches allows for the investigation of the network structure and team performance of 23 soccer teams in up to 76 repeated observations. The soccer context is ideal for the following reasons: the game is governed by clear rules; teams are more comparable in a soccer setting than in other settings; the boundaries of the teams are well defined; no players are missing; and the strength of interaction within teams and team performance can be assessed objectively.

After a review of the literature on group performance, the three main limitations of the previous research are identified. Then, the hypotheses that are tested in this study are presented. The next section describes the setting and the data. Then, the variables and measurements for team performance and network structure are introduced. The methods section describes the analytical strategy and the mixed-effects modeling approach that is used. The results section follows the methods section. The article then concludes with a discussion of the findings.

* Present address: ETH Zurich, Chair of Sociology (Modeling and Simulation), Clausiusstrasse 50, 8092 Zurich, Switzerland. Tel.: +41 44 632 07 42; fax: +41 44 632 17 67.

E-mail addresses: tgrund@ethz.ch, thomas.u.grund@gmail.com

0378-8733/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.socnet.2012.08.004

The main contribution of this article is to study the issue of within-team network structure and the performance of teams through an analysis of panel data. This study draws on an innovative setting, which allows for the analysis of 1520 different networks and performance outcomes. The article thereby contributes to existing debates about the role of embeddedness in the performance outcomes of teams and firms (see, e.g., Granovetter, 1985; Uzzi, 1996; Borgatti and Foster, 2003).

2. Literature

2.1. Previous research on network structure and team performance

One of the earliest empirical studies on interpersonal relations and team performance was conducted at the Hawthorne Works of Western Electric in the 1920s (see Roethlisberger and Dickson, 1939). While the Hawthorne studies were designed to find ways to increase workers' productivity, William Lloyd Warner and Elton Mayo probed interpersonal relations to describe group structures and used experiments to explore the impact of different work conditions on group productivity (Mayo, 1933).

A set of experiments conducted by Alex Bavelas in his Group Networks Laboratory aimed to investigate the role of communication structures in task performance. Participants were arranged in groups of five individuals, and each group had to solve a puzzle. The findings showed that "communication nets" with centralized structures (e.g., a wheel) improved the diffusion of information in simple tasks, whereas decentralized structures (e.g., a circle) delayed the diffusion of information (Bavelas, 1950: 730). Later research built on these experiments and demonstrated that decentralized communication structures are more efficient in solving complex tasks and lead to fewer errors (Leavitt, 1951; Guetzkow and Simon, 1955; Shaw, 1964).

Despite these early efforts and the increased use of work groups in organizations and firms (Guzzo and Salas, 1995; Hackman, 1990), research on network structures and team performance soon came to a halt. It was only recently that the topic of network structure and team performance resurfaced in academia (Katz et al., 2004). Less than ten years ago, Cummings and Cross (2003: 197) noted that "there has been relatively little social network research on the structural properties of natural work groups and their consequences for performance". Among the more recent research, Sparrowe et al. (2001) conducted a field study of 38 work groups in five organizations. The results are similar to those of Shaw (1964) and demonstrate that groups with decentralized communication patterns perform better than groups with centralized communication patterns. In another study, Cummings and Cross (2003) investigated 182 work groups performing complex tasks in a global organization and found that core–periphery and hierarchical group structures were negatively associated with performance. A study of 224 corporate R&D teams by Reagans and Zuckerman (2001) indicates that network density is positively related to productivity. Rulke and Galaskiewicz (2000) study the group network structure and performance of 39 teams of MBA students in management simulation games and find that decentralization is positively associated with stock price. Gloor et al. (2008) examined the number of e-mails sent between online team members and observed that balanced communication structures (e.g., an equal number of e-mails sent and received) are positively related to team performance. Studying 59 consulting teams, Carson et al. (2007) found that shared leadership predicts team performance. An additional overview on empirical studies that relate characteristics of team networks with team effectiveness can be found in Henttonen (2010). An associated set of studies relates the position of individuals in networks (e.g., node centrality) to individual performance outcomes (see, e.g., Baldwin et al., 1997; Sparrowe et al., 2001).

2.2. Limitations of previous research

The most important limitation of the previous research concerns the issue of causality. Do network structures drive team performance, or does performance promote certain network configurations in a team? All previous studies apply a cross-sectional design because of the difficulty of collecting longitudinal network and performance data. The absence of longitudinal analysis makes it problematic to say whether the network or the hypothesized effects of the network is causally antecedent (Lazer, 2001).[1]

Other limitations are related to general developments in social network research. First, social network scholars have long focused on friendships or advice relationships, rather than the interaction of individuals during the production process itself (for an exception, see Brass, 1981).[2] In many contexts, such a focus is appropriate (e.g., social-capital research), but in others it is less appropriate. While friendships and advice relationships are certainly pivotal for teams, they need to translate into the orchestration of a group production process to matter. Often, separating friendship, advice, and communication relationships is impossible because they are closely interwoven.

A second limitation is the focus on binary relationships. A large fraction of contemporary social network analyses treat network ties in a binary fashion—i.e., ties either exist, or they do not. In the context of network structure and team performance, for example, Cummings and Cross (2003) consider the strength of ties only together with a cutoff point to extract a binary projection of the "valued" relationship. Scholars have long questioned such a binary approach toward networks as it only encompasses qualitative relationships and neglects the strength of ties (see e.g., Festinger, 1949; Lévi-Strauss, 1963; Doreian, 1969; Peay, 1980). In the literature, there is consensus about the importance of variability in strength of interpersonal relationships (see Freeman et al., 1991). Consequently, several researches emerged which explicitly consider the values of social ties (e.g., MacKenzie, 1966; Peay, 1976; Freeman et al., 1991; Opsahl and Panzarasa, 2009). For example, a work relationship can be expected to vary depending on whether colleagues interact occasionally or all the time. Additionally, the intensity of interactions may be especially relevant in small teams, in which everybody is likely to be associated with everybody else.

Another intricate issue is the assessment and comparability of performance measures. Most commonly, team performance is assessed in a subjective manner. For example, Sparrowe et al. (2001) interviewed team leaders to rate the performance of their own team. Such an approach is often necessary because objective measures are unavailable. Furthermore, comparing the performance of different teams is not always a straightforward process, as teams are often involved in different types of work. As Sparrowe et al. (2001) have mentioned, it is also problematic that group leaders are commonly lenient and overrate group performance—especially because a group's effectiveness also reflects the performance of its leader.

[1] See also Aldrich (1991) and Reese and Aldrich (1995) concerning this issue in the context of entrepreneurial networks and business performance.

[2] Of course, many studies exist which focus on network relationships, which are not based on friendship. For example, scholars examined trust (Buskens, 2002) or gossip relationships (Ellwardt, 2011).

3. Hypotheses

3.1. Density-performance hypothesis

One of the most straightforward propositions in the literature is that network density and intense interactions between individual members increase team performance. Balkundi and Harrison (2006) label this the "density-performance hypothesis". For example, Sparrowe et al. (2001) suggest that when team members have strong relationships with many other team members, mutual interdependence increases. According to Molm (1994), such interdependence calls for cooperation and a coordination of efforts. A similar argument is made by Granovetter (1985), who showed that embeddedness in tightly linked and dense network structures (e.g., the network of New York diamond traders) increases trust and dependence in such communities. It has also been suggested that dense networks encourage information sharing and increase knowledge about other members of the network (Littlepage et al., 1997). Actors within these networks are more aware of other team members' potential and resources. Hence, dense network structures facilitate the mobilization of these resources. Similarly, intense interactions between team members increase visibility and accountability. It has been suggested that dense networks are effective in reducing social loafing—the tendency of individuals to put less effort into a job when they are in a group than when they are alone—because individuals can be held responsible more effectively (Wagner, 1995). Following this argument, I hypothesize the following:

Hypothesis 1. Controlling for interaction opportunities in teams, increased interaction intensity leads to increased team performance.

3.2. Centralization-performance hypothesis

Another hypothesis put forth in the literature is that centralization—the degree to which network positions are unequally distributed in a team—is related to performance. Group centralization is lowest when all members of a team are equally "central". Researchers do not always agree on how "group centrality" or "centralization" should be assessed. For example, Sparrowe et al. (2001) based their evaluation on degree distribution—the number of links the individual team members have. However, Cummings and Cross (2003) focused on core-periphery structures in networks.

Despite these differences, the theoretical rationale for a proposed negative relationship between team centralization and team performance remains the same. Sparrowe et al. (2001) and Molm (1994) claim that decentralized network structures foster interdependence, which ultimately encourages coordination and cooperation. Mutual relationships (not asymmetric ones) prohibit the exploitation of individuals. Furthermore, research suggests that "flat" hierarchies, or the ability of members to quickly reach others, affect crisis management (see Krackhardt and Stern, 1988; Cummings and Cross, 2003). Decentralized teams may provide more flexibility and timely information and are less dependent on specific "central" individuals. In line with this literature, I hypothesize the following:

Hypothesis 2. Increased centralization of interaction in teams leads to decreased team performance.

4. Setting and data

Despite sports teams having long been the subject of management, organization and group research (see, e.g., Grusky, 1963; Brown, 1982; Pfeffer and Davis-Blake, 1986), the network structure of such teams has hardly been examined. One exception is a study by Gould and Gatrell (1979) in which they investigate passes made between soccer players in the 1977 final between Liverpool and Manchester United. Their analysis, while innovative at the time, went largely unnoticed and has not been applied to the study of team performance. A more recent study by Duch et al. (2010) investigates passes made between players in matches of the European Cup 2008 soccer tournament to assess individual performance.

This article draws on a unique dataset of team structure and performance in the EPL, the top division of the English soccer system. The data were purchased from OPTA Sportsdata and contain detailed narratives of all 760 EPL matches in the 2006/07 and 2007/08 seasons. Based on video footage, the dataset includes 1,050,411 in-match events (including goals, passes, referee decisions and player behaviors).[3] Information on 283,259 passes between the players was used to construct networks for each team in each match. In total, 1520 networks were derived, including up to 76 matches for 23 teams.[4] The strength of the ties indicates the total number of passes between two players.

4.1. Performance variable

In soccer, the performance of a team is crucial for a team's standing at the end of the season. Each team accumulates points for winning or drawing a match. All the matches involved in this study were watched by thousands of spectators and broadcasted to millions of fans worldwide. Hence, there was no reporting bias. Although the points a team gains in a match ultimately determine its league table position, I focus on the number of goals a team scores in a match, an amount that indicates the offensive performance of the team. Limiting the analysis to offensive performance seems reasonable because the data only allows for the investigation of the orchestration of offensive team play (i.e., when a team possess the ball and players pass). Alternatively, one could also think of other performance measures, such as financial statements of teams. However, these measures would only be indirectly affected by how a team plays and cannot be assessed on a match-level basis.

4.2. Network structure and centralization

The setting of soccer allows for a direct assessment of interactions among team members. Furthermore, one of the advantages of investigating soccer is that the boundaries of the teams, and hence possible interactions of team members, are clearly defined. While only a player in possession of the ball can score a goal, players constantly pass the ball to each other, and team play is organized around such passes. Of course, team-relevant interactions can also take other forms—e.g., distracting the opponent team while not having the ball. While such other forms of team interactions are certainly important, direct passes between players of the same team are most likely the most consequential form of interaction in soccer matches and can be used to approximate the orchestration of group production.[5]

[3] Data has been retrieved using automated motion analysis (see Carling et al., 2008).

[4] For 17 teams 76 repeated observations are available and for 6 teams only 38 repeated observations.

[5] In the following text, I always refer to 'passes between players' with the term 'interaction' unless it is specified otherwise. Only successful passes are considered. It remains impossible to identify the intended destination of unsuccessful passes. Hence, network measures can only be calculated on the basis of successful passes. Such a limitation may be criticized for not accurately reflecting the interaction patterns in teams. At the same time, however, it seems plausible to argue that realized and not intended patterns of interaction are most crucial for team success.

Table 1
Passing pattern of Arsenal against Aston Villa; Saturday, August 19, 2006, Emirates Stadium.

	Fabregas	Silva	Hleb	Toure	Djourou	Henry	Eboue	Hoyte
Fabregas	–	9	24	5	2	12	10	3
Silva	17	–	15	11	5	8	3	11
Hleb	17	8	–	3	1	15	7	–
Toure	8	9	14	–	13	4	10	1
Djourou	5	13	2	17	–	1	–	6
Henry	4	5	10	3	2	–	3	1
Eboue	12	9	7	12	2	2	–	1
Hoyte	12	12	2	–	9	2	3	–

Note: Values indicate the number of passes from row to column player. Only information for the 8 most active players are shown. Ljungberg, Adebayor and Hoyte were substituted. Lehman was the goalkeeper.

During soccer matches, we can expect every player to make at least one pass to another team member. Interactions are frequent, but distinctive patterns emerge from the number of passes and the identities of the players involved.[6] An example of a 'pass network' is contained in Table 1. The table shows the number of direct passes between Arsenal players in the rows and Arsenal players in the columns in a match against Aston Villa on 19 August 2006. As suggested earlier, nearly all players make at least one pass to another team member (in Table 1, only four out of 56 directed ties are not realized). Therefore, a structural analysis based on unweighted ties would be of limited use, as distinctive patterns manifest themselves only through interaction frequencies—a common situation for small teams.

4.2.1. Network density and intensity

The most widely assessed feature of networks is density, which is traditionally calculated as the number of existing ties in a network divided by the number of potential ties. As suggested earlier, network density is less useful when ties are weighted and networks are almost complete (i.e., everybody is connected to everybody else). Already MacKenzie (1966) proposed to consider the weights of ties in networks explicitly, when structural features of a network are assessed. Various attempts have been made to generalize the assessment of network features to account for weighted ties (see, e.g., Barrat et al., 2004; Newman, 2004; Opsahl et al., 2010). In our context, the overall level of interaction can be assessed straightforwardly. Let us first define out-strength in professional soccer, $C_{OS}(i)$ of a node i as the sum of the values of outgoing ties (e.g. number of passes made by this player) and in-strength $C_{IS}(i)$ of a node i as the sum of the values attached to incoming ties (e.g. number of passes a player receives).

$$C_{OS}(i)\sum_{j=1}^{N} w_{ij} \tag{1}$$

$$C_{IS}(i) = \sum_{j=1}^{N} w_{ji} \tag{2}$$

with w_{ij} being the number of passes made from player i to player j in a team-match and N the number of nodes (here 8 players). As the opportunities for interaction heavily matter in sports—only the team with the ball can pass—I standardize the measure by the time T the team possesses the ball in the match.

One can then define *network intensity I* as the passing-rate for a team-match as:

$$I = \frac{1}{T}\sum_{i=1}^{N}\left(\frac{C_{OS}(i) + C_{IS}(i)}{2}\right) = \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} \tag{3}$$

4.2.2. Node and tie based network centralization

Group or network centralization is concerned with the distribution of individual network positions, e.g. node centrality. For example, a network is considered highly centralized when one actor is clearly more central than all other actors in the network. A network is decentralized when all actors have the same node centrality. Freeman (1978) proposed the following strategy to derive network centralization from node centrality scores: First, calculate the sum of the differences between the largest node centrality score and the scores of all other nodes in the network. Second, divide this sum by the maximum possible sum of differences (Wasserman and Faust, 1994). While such a method for calculating network centralization is purely node-based (as it refers to the potentially unequal distribution of node centrality scores), a tie-based counterpart to assess network centralization in weighted networks exists. In unweighted networks, ties are indistinguishable from each other. In weighted networks, however, ties have a value attached to them. We can therefore determine how unequally distributed these values are. From such a tie-based perspective, one would then measure network centralization not by examining the distribution of node characteristics (i.e. node centrality scores), but rather by investigating the distribution of tie characteristics (i.e. tie values). Applying Freeman's (1978) definition of centralization, a network would then be least centralized when all tie values are the same and most centralized when the sum of the differences between the highest tie value and all other tie values is at its maximum.

Of course, there are several plausible ways of assessing network positions, which are reflected in the variety of available indicators that aim to measure node centrality and network centralization (see e.g. Freeman, 1978; Borgatti and Everett, 2006). In the context of soccer, for example, Duch et al. (2010) put focus on betweenness and investigate the contribution of individual players in pass sequences. The current study follows the work of Sparrowe et al. (2001), who are concerned about the role of centralization for performance in teams and focus on degree distribution.[7]

4.2.3. Weight centralization

The most straightforward way to assess tie-based network centralization in weighted networks is to examine the distribution of tie values. In the case of passes between soccer players, the most decentralized interaction pattern is one where everybody interacts with everybody with the same intensity. In contrast, the most centralized network would be one in which most interactions involve the same two individuals. Formally, one can define weight centralization C_w as:

$$C_w = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(w^* - w)}{(N^2 - N - 1)IT} \tag{4}$$

[6] A complicating feature of soccer matches is substitution. While each team has 11 players (10 field players and 1 goalkeeper) up to 3 of these players can be substituted in the course of a match. Obviously, the opportunities for substitutes to participate in passes are limited. The strategy applied to overcome this difficulty is to reduce the analysis to those 8 players of a team who are most active in a match (in terms of passes). The network structure of the team interaction process is hence assessed on the basis of passes between these 8 most active players only. An alternative strategy would be to limit the analysis to those players who were not substituted and remained on the pitch for the whole duration of the match. The approach to focus on the 8 most active players (in terms of passes) adopted here, however, puts emphasis on the dominant pattern of interaction in the team. The substantive results remain the same for both strategies dealing with substitutes.

[7] Examining the number of passes between players seems most intuitive and resembles the measures used in Sparrowe et al. (2001) best, but I acknowledge that other ways of assessing network positions are also feasible. Borgatti (2005) provides a more detailed discussion about the interpretation of various centrality measures in different settings.

Table 2
Descriptive statistics.

	Mean	Std dev	Min	Max	Obs
(1) *IT*: total passes	186.36	59.68	59	461	1520
(2) *T*: ball possession (in minutes)	45	5.48	24.44	65.56	1520
Network intensity					
(3) *IT*: passing rate (passes per minute weighted by possession)	4.14	1.20	1.43	9.59	1520
Network centralization					
(4) C_w: weight centralization	0.05	0.02	0.02	0.14	1520
(5) C_I: in-strength centralization	0.08	0.03	0.01	0.25	1520
(6) C_O: out-strength centralization	0.08	0.03	0.02	0.24	1520
Team performance					
(7) Goals	1.27	1.21	0	8	1520

Table 3
Correlation matrix of variables.

	(1)	(2)	(3)	(4)	(5)	(6)
(1) *IT*: total passes	1					
(2) *T*: ball possession (in minutes)	0.38	1				
Network intensity						
(3) *IT*: passing rate (passes per minute weighted by possession)	0.92	0.00	1			
Network centralization						
(4) C_w: weight centralization	−0.34	−0.15	−0.32	1		
(5) C_I: in-strength centralization	−0.31	−0.15	−0.28	0.40	1	
(6) C_O: out-strength centralization	−0.158	−0.06	−0.15	0.31	0.20	1
Team performance						
(7) Goals	0.206	0.09	0.18	−0.10	−0.09	−0.06

whereas w^* is the maximum empirically observed tie value. The denominator in formula (4) simply reports the centralization when all interaction (i.e. number of passes) as defined as the total sum of weights (or passes) would be concentrated in a single and directed dyad—when one player only passes to one other player.

4.2.4. Strength centralization

The simplest *node-bas*ed network centralization—as used by Sparrowe et al. (2001)—is degree centralization. This measure is concerned about the distribution of the number of ties that network members have. In our setting, degree centralization is not a useful measure as (in most matches) every player is connected with all players. However, one can define centralization for the incoming and outgoing node strength:

$$C_I = \frac{\sum_{i=1}^{N}(C_{IS}^* - C_{IS}(i))}{(N-1)IT} \tag{5}$$

$$C_O = \frac{\sum_{i=1}^{N}(C_{OS}^* - C_{OS}(i))}{(N-1)IT} \tag{6}$$

With C_{IS}^* and C_{OS}^* being the largest empirically observed in-strength and out-strength of nodes. Then, for example, C_I (in-strength centralization) is highest when one player receives all passes and lowest when every member of the team receives an equal number of passes. Similarly, C_O (out-strength centralization) would be highest when one player makes all the passes and lowest when every member of the team makes the same number of passes.

4.3. Factor analysis

Some descriptive statistics of the variables used in the analysis are presented in Tables 2 and 3. On average, soccer teams make 186.36 passes in each match. Considering the ball possession, this translates into 4.14 passes per minute on average. This number is in line with our understanding of professional soccer. The mean scores for weight, in-strength, and out-strength centralization that are presented in Table 2 are rather low. However, notice that they are standardized to the rather extreme case, in which all passes are made from one player to another player. Finally, a team scores an average of 1.27 goals per match.

To derive a single dimension for network centralization, factor analysis was applied. The Kaiser–Meyer–Olkin measure of sampling adequacy is 0.60, and Bartlett's test of sphericity is significant ($\chi^2(3) = 427.79$, $p < 0.00$). Principal component factor analysis is used because the primary purpose is to identify and compute a composite score for the underlying notion of network centralization. Only one factor with an eigenvalue greater than one is retrieved. The factor loadings and communality scores are presented in Table 4. A composite score for the extracted factor based on all items is created for each team-match. A high score indicates a passing pattern characterized by centralized team play. A low score indicates a decentralized passing pattern. The skewness (0.96) and kurtosis (1.49) were within the tolerable range ($-2 < x < 2$) for assuming a normal distribution. An examination of the histograms also suggested that the distributions looked approximately normal.

5. Method

5.1. Analytical strategy and mixed-effects modeling

Longitudinal panel-data analysis is preferable to cross-sectional analysis, as it focuses on the dynamics of change and facilitates a discussion of causality by incorporating the time structure of

Table 4
Factor loadings and communalities based on a factor analysis on three network centralization indices for 1520 weighted team-match pass networks.

	Network centralization	Communality
C_w: weight centralization	0.80	0.65
C_I: in-strength centralization	0.65	0.42
C_O: out-strength centralization	0.74	0.54

Table 5
Crossed random effects Poisson regression results for scored goals.

	Expected effect	(1)	(2)	(3)
Fixed part				
Ball possession (in minutes)		1.014** (0.000)	1.006 (0.201)	0.995 (0.324)
Network intensity (passing rate)	>1 (Hypothesis 1)	1.127** (0.000)	1.062* (0.016)	1.045+ (0.088)
Network centralization (composite score)	<1 (Hypothesis 2)	0.950* (0.048)	0.946* (0.033)	0.951* (0.049)
Random part				
$\sqrt{\psi_1}$ [team]		–	0.209	0.227
$\sqrt{\psi_2}$ [opponent]		–	–	0.276
Random team effects		No	Yes	Yes
Random opponent effects		No	No	Yes
Observations		1520	1520	1520
Log likelihood		−2215.55	−2200.64	−2163.02

Exponentiated coefficients; p-values in parentheses.
\+ $p < 0.10$.
* $p < 0.05$.
** $p < 0.01$.

data. In the debate about network structure and team performance, several scholars have argued that longitudinal analysis would be necessary to disentangle the proposed relationship between the embeddedness of individuals and group performance outcomes (Sparrowe et al., 2001; Katz et al., 2004). So far, it has seemed impossible to acquire the necessary data to perform such analyses. Repeated observations for both the network and performance of different teams are needed. The data used in this study, however, meet these requirements.

Mixed-effects models provide a powerful and flexible tool for the analysis of grouped and longitudinal data. They allow one to take heterogeneity and dependence structure into account. It has been shown that mixed-effects models incorporate previous panel-data approaches with random effects (Rabe-Hesketh and Skrondal, 2008). Practically, mixed-effects models allow intercepts and/or slopes of regressions to vary across groups. In the case of longitudinal data, a repeatedly observed unit is treated as such a group. Furthermore, mixed-effects regression models have been extended to deal with multiple grouping layers (Bryk and Raudenbush, 1992; Snijders and Bosker, 1999; Hox, 2010). Models with multiple layers that are nested within each other are called multilevel (ML) or hierarchical linear models (HLM). However, data do not necessarily have to be hierarchically structured, and grouping layers do not need to be nested within each other. In these cases, data are considered to be cross-classified. The soccer data used in this study have such a cross-classified data structure. Each performance assessment made in a match is grouped not only according to the identity of the team whose performance is assessed but also in relation to the opposing team. While crossed random-effects models (also known as two-way error component models) extend classical hierarchical multilevel models, they can be fitted with procedures designed for purely hierarchical or multilevel structures (Goldstein, 1987). When applied to our setting, such a strategy addresses the following question: does a team perform better than usual when the passing network is denser or more centralized than it usually is for the team (controlling for opportunities to interact and unobserved team/opponent characteristics)?

5.2. Crossed random-effects Poisson regression

The number of goals scored by a team is a count variable and modeled with a Poisson regression (Rabe-Hesketh and Skrondal, 2008; Cameron and Trivedi, 1998). The value $E(y_{mkl})$ is the number of goals one can expect team k to score against team l in a particular match m. Each season, 20 teams play against each other. However, as three teams get relegated and replaced by three other ones after the first season, both k and l can range from 1 to 23. The variable T_{mkl} measures ball possession of team k in a particular match and, hence, controls for opportunities to pass. The variables I_{mkl} and C_{mkl} describe the passing network structure of team k; I_{mkl} refers to network intensity or the number of passes weighted by ball possession for team k in this match and C_{mkl} depicts the network centralization composite score for team k that comes out of the factor analysis. The mixed-effects model with crossed random effects suggests to split the error ξ_{mkl} of a simple Poisson regression into three components: (1) a component ζ_{1k}, which is specific to each team k whose performance is assessed; (2) another component ζ_{2l}, which is specific for team l being the opponent of the team whose performance is assessed and; (3) a residual error term ε_{mkl}. The components ζ_{1k} and ζ_{2l} are often called random effects or random intercepts; they have a population mean of zero, a variance of ψ_1 and ψ_2 and are assumed to be independent and normally distributed.

$$\zeta_{1k} \sim N(0, \psi_1) \tag{7}$$

$$\zeta_{2l} \sim N(0, \psi_2) \tag{8}$$

The full model can then be specified as:

$$\ln(E(y_{mkl})|I_{mkl}, C_{mkl}) = \beta_1 + \beta_2 T_{mkl} + \beta_3 I_{mkl} + \beta_4 C_{mkl} + \zeta_{1k} + \zeta_{2l} + \varepsilon_{mkl} \tag{9}$$

6. Results

A defining feature of a team is how the team members interact. This study investigates two hypotheses put forth in the literature about the relationship between within-team network structure and team performance. The first hypothesis suggests that network intensity or an increased overall level of interaction (controlling for opportunities for interaction) leads to better team performance. The second hypothesis suggests that network centralization decreases team performance. In other words, a team does not perform as well as it could when the team production process is centralized.

Table 5 shows the regression results. The regression coefficients are exponentiated. Thus, they indicate by how much the expected number of goals is multiplied by a unit increase of the independent variable. p-Values are reported in parentheses. Three model specifications are given. The first column presents results with no random team or opponent effects. This model mirrors the cross-sectional analyses conducted in previous studies.[8] A clear

[8] Practically, ζ_{1k} and ζ_{2l} are excluded from the model specification given in Eq. (9).

network intensity effect is found. Increases in the passing rate lead to increased team performance. As predicted, a clear network centralization effect is present. Increases in the centralization of team play lead to decreased team performance. Therefore, we can replicate previous findings about teams in the context of this study. The second column shows results with random effects for the team whose performance is assessed, and the last column shows the full crossed random-effects specification, which controls for both unobserved team and opponent effects. Columns (2) and (3) present the panel structure of the data. While causality can still be disputed, these two model specifications come much closer to making such claims than the cross-sectional designs that are traditionally used. The magnitude of both predicted effects declines, but nevertheless, passing rate remains significant at the 0.1 level, while network centralization is at the 0.05 level. Hence, this study replicates earlier findings about network structure and team performance in cross-sectional settings and provides supporting evidence for these findings to hold in a panel design.

7. Conclusion

What makes one team more successful than another? A promising body of research claims that the answer lies within the orchestration of the group production process. Adopting a network approach, this research tradition suggests that the pattern of interactions between individuals plays a pivotal role in team performance. It has been argued that a team relies on the successful mobilization of resources, especially during complex tasks. Team members need to draw on each other's abilities and knowledge. Hence, the ways in which team members interact is crucial for what a team can achieve.

Previous studies on group performance examined two aspects of network structure: (1) the overall level of interaction in teams measured as network density or intensity; and (2) team centralization or the distribution of network positions and roles. While network density is assumed to be positively related to team performance, a negative association is expected between network centralization and team performance.

The current study draws on an innovative setting and data, including repeated observations from the realm of team sports, to overcome problems of previous studies. Information on 283,259 passes between individual soccer players in 760 matches in the English Premier League allowed for the mapping of the group performance and group production process of 23 teams. In total, 1520 different networks were analyzed.

Controlling for the opportunities for interaction, our results confirm previous findings about the link between the level of interaction in teams and performance. High levels of interaction (i.e. passing rate) lead to increased team performance. In addition, our results support previous findings about the link between centralization and team performance. Centralized interaction patterns lead to decreased team performance.

An important contribution of this article lies in extending previous research on network structure and team performance in a way that incorporates repeated observations of teams. Previous results were limited due to the use of cross-sectional data and designs. Confirming some of these earlier results with panel data for both network structure and team performance allows for mixed-effects modeling and leads to more robust claims. Future work could exploit the time-series character of the data even more by investigating the dynamics of network structure and team performance. Such a perspective could be useful in examining changes in individual team members' interaction and performance trends.

However, this article also contributes to a more general debate about the role and importance of embeddedness in network structures. Researchers have been interested in the role of network structures in relation to reputation, conflict, attitudes and general behavior in groups. By examining the role of interpersonal relations in team performance, this study builds on this wider network paradigm.

Obviously, this study also has its limitations. While the analysis of sport teams is established in management research (see Grusky, 1963; Brown, 1982; Pfeffer and Davis-Blake, 1986), one may question the degree to which the results of the current study can be generalized to other sorts of teams. Brown (1988) emphasizes the nature of tasks in playing a significant role for group processes. Similarly, the work of Bavelas (1950), Leavitt (1951) and Shaw (1964) illustrates that decentralized network structures may be more efficient only in solving complex tasks. However, Elias and Dunning (1966) emphasize the advantages of investigating sports (particularly soccer), as configurational dynamics within social units are more visible in, for example, the activities of sport teams.

Another limitation involves the tactical setup of teams and the roles of players. For example, the goalkeeper in a team is allowed to use his hands, while all other players are not. Some players are mostly defenders and midfielders, while others are strikers. However, these different roles and even differences in the tactical setup between teams (e.g., the number of strikers a team nominates for a match) would not affect the results because of the mixed-effects modeling approach that is applied in this study. The question that the model addresses is the following: does one team perform better than it usually does when the network intensity or network centralization is higher than it usually is for this team? This problem would therefore only be of concern if teams changed their tactical setup significantly, depending on (previous) team performance.

The current analysis implicitly assumes that team performance and network structure are measured at exactly the same time. Admittedly, this is not the case with our data. It is possible for teams to play well and score goals in the first half of a match and then switch to a different mode of team play to defend their lead in the second half. The network data used for approximating the orchestration of team play build on passing events and not on stable relationships. For example, in the statistical modeling of network evolution, the difference between events and states is of crucial importance; however, the explicit focus on weighted networks, where tie strengths represent the sum of passes between players, makes the difference between events and states less of a concern. In the context of this study, it can be argued that the realized passing structure represents an underlying pattern of orchestrated team play. Although within-match dynamics are undoubtedly important and theoretically interesting, this study presents an original insight into the role of network centralization in sport teams. Additional research could apply a relational event perspective (e.g., Butts, 2008) and investigate within-match dynamics explicitly. Such an approach could include the investigation of relational sequences—e.g., closure of triads—and develop a multilevel perspective, wherein team performance is related to the dyadic level of passes.

From a modeling point of view, while the random effects for teams are constants, the players nominated by the teams in the matches are not. A team normally consists of up to 25 players, and the team manager selects 11 players for each match. Sometimes players are injured and cannot play anymore. Other times, the team manager decides to give other players a chance. However, the core of each team is rather constant. In nearly all teams, one finds players who play in almost all the matches. Additionally, all players of the team, including all substitutes, practice together on a daily basis. Therefore, one can think of the interaction pattern exhibited by the eleven players in a match as a sample of general team play.

This study also provides several opportunities for future work. It draws on a setting, which entails a special paired structure of the observations. In this analysis, the characteristics of the setting are considered in the modeling strategy using mixed-effects models with crossed random effects. Researchers should aim to replicate these results in settings without such an implicit dependency structure.

Additionally, this study examines only the role of network structure for team performance. A related body of research applies similar theoretical frameworks and techniques to examine the association between individual network positions and individual performance. This type of research also generally draws on cross-sectional designs and could benefit from panel data. The data of this study can be used for such a task. However, a difficulty lies in retrieving objective measures for individual performance. When considered in relation to existing strategies to measuring individual performance in soccer matches (e.g., the Actim-Index),[9] this data, which includes narratives of matches and a wide array of individual behaviors (e.g., goals, assists, tackles, fouls, challenges, off-sides and many more), allows for a possible and promising avenue for further research.

Last, this study stresses the importance of different types of networks. While a large amount of social network analyses focus on, for example, friendships, this study investigates the group production process in teams more closely. A direct observation of interactions in teams has been used to measure the orchestration of team play. However, most certainly, other types of relationships might be important and might even be antecedents of actual observed interactions. Another promising avenue for research would be to examine the relationships between different types of networks and team play. Using the current study as an example, we could ask the following: do soccer players make more passes to each other when they know each other well or when they have both played on the same team for a long time? Answering questions of this type could potentially be useful for engineering group production—i.e., it might offer opportunities for increasing team performance.

[9] URL: http://www.premierleague.com/page/ActimFAQs, last retrieved 20 November 2011.

References

Aldrich, H.E., 1991. Methods in our madness? Trends in entrepreneurship research. In: Sexton, D., Kasarda, J.D. (Eds.), The State of the Art of Entrepreneurship. PWS-Kent Publishing, Boston, pp. 191–213.

Baldwin, T.T., Bedell, M.D., Johnson, J.L., 1997. The social fabric of a team-based M.B.A. program: network effects on student satisfaction and performance. Academy of Management Journal 40 (6), 1369–1397.

Balkundi, P., Harrison, D., 2006. Ties, leaders, and time in teams: strong inference about network structure's effects on team viability and performance. Academy of Management Journal 49 (1), 49–68.

Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of complex weighted networks. Proceedings of the National Academy of Sciences of the United States of America 101 (11), 3747–3752.

Bavelas, A., 1950. Communication patterns in task-oriented groups. Journal of the Acoustical Society of America 22 (6), 725–730.

Borgatti, S.P., 2005. Centrality and network flow. Social Networks 27 (1), 55–71.

Borgatti, S.P., Foster, P.C., 2003. The network paradigm in organizational research: a review and typology. Journal of Management 29 (6), 991–1013.

Borgatti, S.P., Everett, M.G., 2006. A graph-theoretic perspective on centrality. Social Networks 28 (4), 466–484.

Brass, D., 1981. Structural relationships, job characteristics, and worker satisfaction and performance. Administrative Science Quarterly 26 (3), 331–348.

Brass, D., 1984. Being in the right place: a structural analysis of individual influence in an organization. Administrative Science Quarterly 29 (4), 518–539.

Brown, C., 1982. Administrative succession and organizational performance: the succession effect. Administrative Science Quarterly 27 (1), 1–16.

Brown, R., 1988. Group Processes. Dynamics Within and Between Groups. Blackwell, Oxford.

Bryk, A.S., Raudenbush, S.W., 1992. Hierarchical Linear Models: Applications and Data Analysis Methods. Sage, Newbury Park.

Buskens, V., 2002. Social Networks and Trust. Kluwer, Boston.

Butts, C., 2008. A relational event framework for social action. Sociological Methodology 38 (1), 155–200.

Cameron, A.C., Trivedi, P.K., 1998. Regression Analysis of Count Data. Cambridge University Press, Cambridge.

Carling, C., Bloomfield, J., Nelson, L., Reilly, T., 2008. The role of motion analysis in elite soccer: contemporary performance measurement techniques and work rate data. Sports Medicine 38 (10), 839–862.

Carson, J.B., Tesluk, P.E., Marrone, J.A., 2007. Shared leadership in teams: an investigation of antecedent conditions and performance. Academy of Management Journal 50 (5), 1217–1234.

Cummings, J.N., Cross, R., 2003. Structural properties of work groups and their consequences for performance. Social Networks 25, 197–210.

Doreian, P., 1969. A note on the detection of cliques in valued graphs. Sociometry 32, 237–242.

Duch, J., Waitzman, J.S., Amaral, L.A.N., 2010. Quantifying the performance of individual players in a team activity. PLoS One 5 (6), e10937.

Elias, N., Dunning, E., 1966. Dynamics of group sports with special reference to football. British Journal of Sociology 17 (4), 388–402.

Ellwardt, L., 2011. Gossip in organizations. ICS-dissertation. Groningen.

Flap, H., Bulder, B., Völker, B., 1998. Intra-organizational networks and performance: a review. Computational & Mathematical Organization Theory 4 (2), 109–147.

Festinger, L., 1949. The analysis of sociograms using matrix algebra. Human Relations 10, 153–158.

Freeman, L.C., 1978. Centrality in social networks: conceptual clarification. Social Networks 1, 215–239.

Freeman, L.C., Borgatti, S.P., White, D.R., 1991. Centrality in valued graphs: a measure of betweenness based on network flow. Social Networks 13 (2), 141–154.

Gloor, P.A., Paasivaara, M., Schoder, D., Willems, P., 2008. Finding collaborative innovation networks through correlating performance with social network structure. International Journal of Production Research 46 (5), 1357–1371.

Goldstein, H., 1987. Multilevel covariance components models. Biometrika 74 (2), 430–431.

Gould, P., Gatrell, A., 1979. A structural analysis of a game: the Liverpool v Manchester United Cup Final of 1977. Social Networks 2, 253–273.

Granovetter, M., 1985. Economic action and social structure: the problem of embeddedness. American Journal of Sociology 91 (3), 481–510.

Grusky, O., 1963. Managerial succession and organizational effectiveness. American Journal of Sociology 69 (1), 21–31.

Guetzkow, H., Simon, H., 1955. The impact of certain communication nets upon organization and performance in task-oriented groups. Management Science 1 (3/4), 233–250.

Guzzo, R., Salas, E.E. (Eds.), 1995. Team Effectiveness and Decision Making in Organizations. Jossey-Bass, San Francisco.

Hackman, R. (Ed.), 1990. Groups that Work (and those that do not): Creating Conditions for Effective Teamwork. Jossey-Bass, San Francisco.

Henttonen, K., 2010. Exploring social networks on the team level: a review of the empirical literature. Journal of Engineering and Technology Management 27, 74–109.

Hox, J.J., 2010. Multilevel Analysis. Techniques and Applications, 2nd edition. Routledge, New York.

Ibarra, H., 1993. Network centrality, power, and innovation involvement: determinants of technical and administrative roles. Academy of Management Journal 36 (3), 471–501.

Katz, N., Lazer, D., Arrow, H., Contractor, N., 2004. Network theory and small groups. Small Group Research 35 (3), 307–332.

Kozlowski, S.W.J., Bell, B.S., 2003. Work groups and teams in organizations. In: Borman, W.C., Ilgen, D.R., Klimoski, R. (Eds.), Handbook of Psychology: Industrial and Organizational Psychology, vol. 12. Wiley, New York, pp. 333–375.

Krackhardt, D., Stern, R., 1988. Informal networks and organizational crises: an experimental simulation. Social Psychology Quarterly 51 (2), 123–140.

Lazer, D., 2001. The co-evolution of individual and network. Journal of Mathematical Sociology 25 (1), 69–108.

Leavitt, H., 1951. Some effects of certain communication patterns on group performance. Journal of Abnormal and Social Psychology 46 (1), 38–50.

Lévi-Strauss, C., 1963. Structural Anthropology. Basic Books, New York.

Littlepage, G., Robison, W., Reddington, K., 1997. Effects of task experience and group experience on group performance, member ability, and recognition of expertise. Organizational Behavior and Human Decision Processes 69 (2), 133–147.

Mayo, E., 1933. The Human Problems of an Industrial Civilization. Macmillan Press, New York.

MacKenzie, K.D., 1966. Structural centrality in communications networks. Psychometrika 31 (1), 17–25.

Molm, L.D., 1994. Dependence and risk: transforming and structure of social exchange. Social Psychology Quarterly 57 (3), 163–176.

Newman, M.E.J., 2004. Analysis of weighted networks. Physical Review E 70, 056131.

Opsahl, T., Panzarasa, P., 2009. Clustering in weighted networks. Social Networks 31 (2), 155–163.

Opsahl, T., Agneessens, F., Skvoretz, J., 2010. Node centrality in weighted networks: generalizing degree and shortest paths. Social Networks 32 (3), 245–251.

Peay, E.R., 1976. A note concerning the connectivity of social networks. Jounral of Mathematical Sociology 4 (2), 319–321.

Peay, E.R., 1980. Connectedness in a general model for valued networks. Social Networks 2, 385–410.

Pfeffer, J., Davis-Blake., A., 1986. Administrative succession and organizational performance: how administrator experience mediates the succession effect. Academy of Management Journal 29 (1), 72–83.

Podolny, J.M., Baron, J.N., 1997. Resources and relationships: social networks and mobility in the workplace. American Sociological Review 62 (5), 673–693.
Rabe-Hesketh, S., Skrondal, A., 2008. Multilevel and Longitudinal Modeling Using Stata, 2nd edition. Stata Press, College Station, TX.
Reagans, R., Zuckerman, E., 2001. Networks, diversity, and productivity: the social capital of corporate R&D teams. Organization Science 12 (4), 502–517.
Reese, P.R., Aldrich, H.E., 1995. Entrepreneurial networks and business performance: a panel study of small and medium-sized firms in the research triangle. In: Birley, S., MacMillan, I.C. (Eds.), International Entrepreneurship. Routledge, London, pp. 124–144.
Roethlisberger, F.J., Dickson, W.J., 1939. Management and The Worker. Harvard University Press, Cambridge, MA.
Rulke, D.L., Galaskiewicz, J., 2000. Distribution of knowledge, group network structure, and group performance. Management Science 46 (5), 612–625.
Sanna, L., Parks, C., 1997. Group research trends in social and organizational psychology: whatever happened to intragroup research? Psychology Science 8 (4), 261–267.
Shaw, M.E., 1964. Communication networks. In: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology. Academic Press, New York, pp. 111–147.
Snijders, T.A.B., Bosker, R.J., 1999. Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. Sage, London.
Sparrowe, R., Liden, R., Wayne, S., Kraimer, M., 2001. Social networks and the performance of individuals and groups. Academy of Management Journal 44 (2), 316–325.
Uzzi, B., 1996. The sources and consequences of embeddedness for the economic performance of organization. The network effect. American Sociological Review 61 (4), 674–698.
Wagner, J.A., 1995. Studies of individualism-collectivism: effects on cooperation in groups. Academy of Management Journal 38 (1), 152–172.
Wasserman, S., Faust, K., 1994. Social Network Analysis. Cambridge University Press, Cambridge.