CONNECTIONS 22(2): 12-21

International Management Research and Social Networks

Nicholas Athanassiou[1]
College of Business Administration, Northeastern University

INTRODUCTION

Social networks concepts and methodology have been used increasingly in empirical research published in leading U.S.-based management journals over the past ten years. Yet, use of similar social networks concepts for the study of international management phenomena has lagged this trend, although these concepts offer an untapped opportunity for research that would add considerable insights to our understanding of how multinational corporations are managed.

The purpose of this article is threefold. First, it surveys the management literature to identify and discuss the trends in empirical studies that use social networks concepts and that have been published in leading management journals from 1989 to 1999. Then, within this broader set of articles, the subset that focuses on MNCs is examined. Second, this article briefly discusses how certain key concepts about multinational corporation (MNC) structure and its management can be viewed from a social networks perspective. Third, it summarizes the constructs examined by recent (1982-1996) empirical studies of international strategic management literature. The article then concludes that such constructs can be studied through a social networks lens.

MANAGEMENT JOURNALS AND SOCIAL NETWORK METHODS

One objective of this review is to examine studies in academic journals that are perceived as accepting papers that are rigorous in their methodology and influential in the field of management and international business management. Several recent studies identify the Academy of Management Journal (AMJ), the Administrative Science Quarterly (ASQ), the Strategic Management Journal (SMJ), Organization Science (OS), and the Journal of International Business Studies (JIBS) as leading management journals (Johnson & Podsakoff, 1994; MacMillan, 1991; Tahai & Meyer, 1999). Journals such as the Academy of Management Review or the Harvard Business Review were not included because they do not publish empirical research and as such do not explain methods and construct/variable operationalization. Also, non-U.S. based journals were not included, in order to control for editorial preferences and review practices. U.S. based journals that focus mostly on organizational behavior were not reviewed because the focus of the study was on the strategy formulation and implementation level of management research.

A total of 60 articles identified as using social network methodology were published in these journals between 1989 and mid-1999 (listed in Appendix I). A close examination of these publication data reveal that the number of articles with a social networks methodology published annually increased considerably during the most recent half of the period examined (Table 1). Further, one can see that AMJ and SMJ were the journals that published more than half of such articles while ASQ recorded the third largest volume over this period. Significantly, JIBS, the journal dedicated to international business research, published only three articles based on social network concepts during this period. Thus, one can see that social networks concepts and methods are becoming more known and more acceptable among management researchers and journal editors.

Table 1
US management journal articles with a social networks perspective

Journal	'89-'90	'91-'92	'93-'94	'95-'96	'97-'98	'99*	Total
AMJ	1	0	2	10	10	0	23
ASQ	3	2	1	2	1	1	10
SMJ	3	1	5	3	4	3	19
OS	0	0	1	3	1	0	5
JIBS	0	0	0	0	2	1	3
Total	7	3	9	18	18	5	60

*(6 months)

When the 60 articles were reviewed to identify articles with international business/management focus, only 8 were found (Appendix II). Of these, 3 were those published by JIBS, 4 by SMJ and 1 by AMJ. ASQ and OS published no such articles (Table 2).

Table 2
U.S. Management Journal Articles
With a Social Networks Perspective *and* an International Business/Management Focus

Journal	'89-'90	'91-'92	'93-'94	'95-'96	'97-'98	'99*	Total
AMJ	0	0	0	1	0	0	1
ASQ	0	0	0	0	0	0	0
SMJ	0	1	1	1	0	1	4
OS	0	0	0	0	0	0	0
JIBS	0	0	0	0	2	1	3
Total	0	1	1	2	3	2	8

*(6 months)

Of the 8 articles with an international focus, 6 have a level of analysis at the country or the interorganizational level. The two remaining articles have a level of analysis internal to the organization, the individual manager level (Table 3). It is noteworthy that more than half of the 53 non-international articles examine phenomena internal to the organization.

Table 3
Level of Analysis for Social Network-Based International Management/Business Articles

	AMJ	ASQ	SMJ	OS	JIBS	Total	Int'l.
Country	0	0	0	0	1	1	1
Interorganizational	10	3	14	1	1	29	5
Intraorganizational	0	1	4	1	1	7	0
CEO	1	0	0	0	0	1	0
Groups/Teams	3	1	1	2	0	7	0
Managers	9	5	0	1	0	16	2
Total	23	10	19	5	3	60	>8

From the above analysis we can conclude that international management articles using social networks methods are beginning to be published in U.S. management journals but at a slow pace. There is no way to know why the pace is so slow. As the following section points out, the multinational corporation is a networked organization. Therefore, a social network approach to its analysis should be appropriate.

THE MNC'S ORGANIZATION AS A NETWORK

The multinational corporation (MNC) is a dynamic and multifaceted organizational phenomenon. It is a corporation that has a physical presence through direct investment in multiple national markets over which it exerts managerial control. Such an investment had traditionally been made to secure resources overseas and/or to manufacture and sell products in the MNC's overseas location. Today, in addition to tangible products, an MNC's investment overseas may also generate intangible products such as services and other knowledge-based products that may be destined for any market around the world.

Because the MNC has a physical presence in multiple national markets, its value chain activities (Kogut, 1985a; Kogut, 1985b; Kogut, 1993; Porter, 1986) resemble a multinational network whose nodes are located in the firm's operating entities that are dispersed across countries around the world. These dispersed operating entities are interdependent to different degrees depending on the MNC and its strategic needs. Also, these operating entities may have a range of different forms. At one extreme, the MNC may operate overseas through wholly owned subsidiaries, i.e., through cross-border transactions that take place within its boundaries and can be thought of as interdependent value chain activity flows across the MNC's wholly owned entities around the world. At the other extreme, the MNC may operate overseas through direct export sales originating in its home market or through licensing, i.e., through arms-length external transactions between the MNC and independent business entities located in the foreign markets. These external transactions can be conceptualized as linkages of the MNC's value chain activities with those of the independent businesses with which the MNC is interacting. Between the two extremes lie hybrid operating arrangements such as international joint ventures, strategic alliances and franchising which add further complexity to the value chain activity interdependencies that may exist within and across the MNC's boundaries in different foreign markets. Furthermore, an MNC may employ several of these overseas operating forms at any one time.

When an MNC operates overseas through wholly owned subsidiaries, it assumes capital investment risk exposure and has managerial control over all of its activities throughout its international operations (Kogut, 1985a; Kogut, 1985b; Porter, 1986). In this case, its structure and processes must be designed to control a complex network of value chain activities that are embedded in diverse cultural environments across many national borders and cultures around the world. Furthermore, an MNC that operates overseas through direct exports has virtually no overseas investment or investment risk exposure and needs a structure and processes to manage a very limited number of value chain activities in each overseas market at the cost of no managerial control.

The dispersion and complexity of its world-wide value chain activity network causes the MNC its major organizational challenge: how can it integrate these dispersed and diverse activities to achieve maximum efficiencies while it retains the necessary levels of local market responsiveness in order to sustain its competitive advantage (Bartlett & Ghoshal, 1998; Prahalad & Doz, 1987). Thus, decisions on the initial distribution of the MNC's discrete activities and the configuration of its assets are only one factor that influences the MNC's long term performance. Additionally, long term performance is determined by the MNC's ability to leverage the flexibility of the sequential advantages of its multinational network and coordinate its activities to face the contingencies and constraints imposed by the firm's evolving complex environment at any particular time (Kobrin, 1994; Kogut, 1993). This flexibility allows the MNC to shift activities to the most advantageous of its international locations in order to maintain its competitive edge as conditions shift over time.

Much empirical work on MNC strategy finds that normative integration is a primary means to the challenge of coordinating dispersed activities (Boyacigiller, 1990; Edström & Galbraith, 1977; Ghoshal & Nohria, 1989; Nohria & Ghoshal, 1997; Roth & Nigh, 1992; Roth, Schweiger, & Morrison, 1991). An MNC strives for global strategic advantage through normative integration, which is a horizontal process of informal task forces, committees, and other information and knowledge sharing and enhancing activities. These activities depend on nurturing the learning process in a direction such that the desired coordination linkages are developed. Thus, normative integration is created through personal relations among managers who develop shared beliefs and increasingly harmonize their managerial behavior around the world.

Thus, the MNC's configuration of operating linkages around the world and the nature of the mechanisms it uses to manage these linkages in order to achieve organizational integration are two broad sets of managerial challenges that can be studied using social networks methods. A further set of relationships that can be studied with social networks methods concerns the MNC's web of relationships within its external environment of multiple and differentiated overseas markets. To date, international management researchers have studied MNC organizations internally and their external embeddedness through the examination of dyadic relationships. As the following section shows, these relationships can be studied more comprehensively through the social network perspective.

INTERNATIONAL STRATEGIC MANAGEMENT CONSTRUCTS AND SOCIAL NETWORKS

A recent study of international strategic management literature (Samiee & Athanassiou, 1998) identified 42 articles published by leading journals between 1982 and 1993. A review of these articles identifies 16 constructs that have been studied to better understand management of the MNC. None of these studies used social network methodologies, though with the exception of one, R&D intensity, they describe relationships among actors in a network (Table 4).

Table 4
International Business Strategy Research Relational Constructs

Behavior control

Centralization of decisions

Competitive intensity

Conflict

Control

Coordination

Flexibility

Foreign commitment

Headquarters-subsidiary integration

Headquarters-subsidiary relations

Information processing capacity

Interdependence--Activity

Interdependence--complexity of
Normative integration

Social harmony (procedural justice)

R&D Intensity

As discussed earlier in this article, the present review of studies on management research identified eight studies with an international business focus that were published since 1993. Of these, only one addresses MNC management issues (Athanassiou & Nigh, 1999). This article treats the MNC's top-management-team as a network and examines how internationalization of the MNC changes the way these managers interact to exchange advice on international business issues.

There could be a number of reasons for this minimal influence of social network concepts on international research. One reason may be the difficulty of collecting social network data across multiple organizations that span multiple national borders. Another reason may be a lack of awareness of social network methods, of the concepts that can be examined through these methods, and of their advantages over more commonly used methodologies by the more general management research community. The lack of awareness is clearly changing, as shown by the increasing number of social networks based articles that have been published more recently (Table 1). Difficulties of collecting social network data across multiple organizations that span multiple national borders are likely to be overcome as more international management researchers become increasingly familiar with the nature of social network research and its natural fit with the nature of the multinational corporation's organization and its managerial challenges.

Clearly, the networked nature of the MNC organization and the prominence of relational tools such as normative integration as means to manage the MNC would indicate that social networks methods should offer an appropriate approach with which to study these phenomena. The MNC remains an untapped opportunity for future research based on social networks concepts and methodology.

REFERENCES

Bartlett, C. A., & Ghoshal, S. (1998). *Managing across borders: The transnational solution*. (2nd ed.). Boston, MA: Harvard Business School Press.

Boyacigiller, N. (1990). The role of expatriates in the management of interdependence, complexity and risk in multinational corporations. *Journal of International Business Studies, 21*(3), 357-381.

Edström, A., & Galbraith, J. R. (1977). Transfer of managers as a coordination and control strategy in multinational organizations. *Administrative Science Quarterly, 22*, 248-263.

Ghoshal, S., & Nohria, N. (1989). Internal differentiation within multinational corporations. *Strategic Management Journal, 10*, 323-337.

Johnson, J. L., & Podsakoff, P. M. (1994). Journal influence in the field of management: An analysis using Salancik's index in a dependency network. *Academy of Management Journal, 37*(5), 1392-1407.

Kobrin, S. J. (1994). Is there a relationship between a geocentric mind-set and multinational strategy? *Journal of International Business Studies, 25*(4), 493-511.

Kogut, B. (1985a). Designing global strategies: Comparative and competitive value-added chains. *Sloan Management Review, 26*(4), 15-28.

Kogut, B. (1985b). Designing global strategies: Profiting from operational flexibility. *Sloan Management Review, 27*(1), 27-38.

Kogut, B. (1993). Foreign direct investment as a sequential process. In P. J. Buckley & P. N. Ghauri (Eds.), *The internationalization of the firm: A reader* (pp. 217-230). New York, NY: Harcourt Brace Jovanich.

MacMillan, I. C. (1991). The emerging forum for business policy scholars. *Strategic Management Journal, 12*(2), 161-165.

Nohria, N., & Ghoshal, S. (1997). *The differentiated network: Organizing MNCs for value creation*. San Francisco: CA: Jossey-Bass.

Porter, M. E. (1986). Competition in global industries: A conceptual framework. In M. E. Porter (Ed.), *Competition in global industries* (pp. 15-60). Boston, MA: Harvard Business School Press.

Prahalad, C. K., & Doz, Y. (1987). *The multinational mission: Balancing local demands and global vision*. New York: Free Press.

Roth, K., & Nigh, D. (1992). The effectiveness of headquarters-subsidiary relations. *Journal of Business Research, 23*, 1-25.

Roth, K., Schweiger, D., & Morrison, A. (1991). Global strategy implementation at the business unit level: Operational capabilities and administrative mechanisms. *Journal of International Business Studies, 22*(3), 369-402.

Samiee, S., & Athanassiou, N. (1998). International strategy research: Cross-cultural methodology implications. *Journal of Business Research, 43*(2), 79-96.

Tahai, A., & Meyer, M. J. (1999). A revealed preference study of management journals' direct influences. *Strategic Management Journal, 20*(3), 279-296.

Appendix I

Empirical Research Articles, Published in Management Journals, that Draw on Social Network Concepts

Academy of Management Journal

Baum, J. A. C., & Korn, H. J. (1996). Competitive dynamics of inter-firm rivalry. *Academy of Management Journal, 39*(2), 255-291.

Baum, J. A. C., & Oliver, C. (1996). Toward an institutional ecology of organizational founding. *Academy of Management Journal, 39*(5), 1378-1427.

Belliveau, M. A., O'Reilly III, C. A., & Wade, J. B. (1996). Social capital at the top: Effects of social similarity and status on CEO compensation. *Academy of Management Journal, 39*(6), 1568-1593.

Brass, D. J., & Burkhardt, M. E. (1993). Potential power and power use: An investigation of structure and behavior. *Academy of Management Journal, 36*(3), 441-470.

Carroll, G., & Teo, A. C. (1996). On the social networks of managers. *Academy of Management Journal, 39*(2), 421-440.

Deephouse, D. L. (1996). Does isomorphism legitimate? *Academy of Management Journal, 39*(4), 1024-1039.

Gulati, R. (1995). Does familiarity build trust? The implications of repeated ties for contractual choice in alliances. *Academy of Management Journal, 38*(6), 85-112.

Human, S. E., & Provan, K. G. (1997). An emergent theory of structure and outcomes in small-firm strategic manufacturing networks. *Academy of Management Journal, 40*(2), 368-403.

Hundley, G., Jacobson, C. K., & Park, S. H. (1996). Effects of profitability and liquidity on R&D intensity: Japanese and U.S. companies compared. *Academy of Management Journal, 39*(6), 1659-1674.

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

Ibarra, H. (1995). Race, opportunity, and diversity of social circles in managerial networks. *Academy of Management Journal, 38*(3), 673-703.

Kraatz, M. S. (1998). Learning by association? Interorganizational networks and adaptation to environmental change. *Academy of Management Journal, 41*(2), 621-643.

Labianca, G., Brass, D. J., & Gray, B. (1998). Social networks and perceptions of intergroup conflict: The role of negative relationships and third parties. *Academy of Management Journal, 41*(1), 55-67.

Leicht, K. T., & Marx, J. (1997). The consequences of informal job finding for men and women. *Academy of Management Journal, 40*(4), 967-987.

Mehra, A., Kilduff, M., & Brass, D. J. (1998). At the margings: A distinctiveness approach to the social identity and social networks of underrepresented groups. *Academy of Management Journal, 41*(4), 441-452.

Nelson, R. E. (1989). The strength of strong ties: Social networks and group conflict in organizations. *Academy of Management Journal, 32*(2), 377-401.

Osborn, R. N., & Hagedoorn, J. (1997). The institutionalization and evolutionary dynamics of interorganizational alliances and networks. *Academy of Management Journal, 40*(2), 261-278.

Provan, K. G., & Sebastian, J. G. (1998). Networks within networks: Service link overlap, organizational cliques, and network effectiveness. *Academy of Management Journal, 41*(4), 453-463.

Saxton, T. (1997). The effects of partner and relationship characteristics on alliance outcomes. *Academy of Management Journal, 40*(2), 443-462.

Shah, P. (1998). Who are employees' social referents? Using a network perspective to determine referent others. *Academy of Management Journal, 41*(3), 249-268.

Tsai, W., & Ghoshal, S. (1998). Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal, 41*(4), 464-476.

Wade, J. (1996). Community-level analysis of sources and rates of technological variation in the microprocessor market. *Academy of Management Journal, 39*(5), 1218-1244.

Xin, K. R., & Pearce, J. L. (1996). Cuanxi: Connections and substitutes for formal institutional support. *Academy of Management Journal, 39*(6), 1641-1658.

Administrative Science Quarterly

Ancona, D. G., & Caldwell, D. F. (1992). Bridging the boundary: External activity and performance in organizational teams. *Administrative Science Quarterly, 37*(4), 634-.

Barley, S. R. (1990). The alignment of technology and structure through roles and networks. *Administrative Science Quarterly, 35*, 61-103.

Burkhardt, M. E., & Brass, D. J. (1990). Changing patterns or patterns of change: The effects of a change in technology on social network structure and power. *Administrative Science Quarterly, 35*(March), 104-127.

Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organizational subunits. *Administrative Science Quarterly, 44*(1), 82-111.

Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly, 37*, 422-447.

Ibarra, H., & Andrews, S. B. (1993). Power, social influence and sense-making: Effects of network centrality and proximity on employee perceptions. *Administrative Science Quarterly, 38*(277-303).

Krackhardt, D. (1990). Assessing the political landscape: Structure, cognition, and power in organizations. *Administrative Science Quarterly, 35*, 342-369.

Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly, 41*(1), 116-145.

Provan, K. G., & Milward, H. B. (1995). A preliminary theory of interorganizational network effectiveness: A comparative study of four community mental health systems. *Administrative Science Quarterly, 40*(1), 1-33.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly, 42*(1), 35-67.

Journal of International Business Studies

Chen, H., & Chen, T.-J. (1998). Network linkages and location choice in foreign direct investment. *Journal of International Business Studies, 29*(3), 445-468.

Koka, B., Prescott, J. E., & Madhavan, R. (1999). Contagion influence on trade and investment policy: A network perspective. *Journal of International Business Studies, 30*(1), 127-148.

Money, R. B. (1998). International multilateral negotiations and social networks. *Journal of International Business Studies, 29*(4), 695-710.

Organization Science

Abrahamson, E., & Rosenkopf, L. (1997). Social network effects on the extent of innovation diffusion: A computer simulation. *Organization Science, 8*(3), 289-309.

Constant, D., Sproull, L., & Kiesler, S. (1996). The kindness of strangers: The usefulness of electronic weak ties for technical advice. *Organization Science, 7*(2), 119-135.

D'Aveni, R. A., & Kesner, I. F. (1993). Top managerial prestige, power and tender offer response: A study of elite social networks and target firm cooperation during takeovers. *Organization Science, 4*(2), 123-151.

Porter Liebeskind, J., Lumerman Oliver, A., Zucker, L., & Brewer, M. (1996). Social networks, learning, and flexibility: Sourcing scientific knowledge in biotechnology firms. *Organization Science, 7*(4), 428-443.

Zack, M. H., & McKenney, J. L. (1995). Social context and interaction in ongoing computer-supported management groups. *Organization Science, 6*(4), 394-422.

Strategic Management Journal

Athanassiou, N., & Nigh, D. (1999). The impact of company internationalization on top management team advice networks: A tacit knowledge perspective. *Strategic Management Journal, 19*(1), 83-92.

Black, J. A., & Boal, K. B. (1994). Strategic resources: Traits, configurations and paths to sustainable competitive advantage. *Strategic Management Journal, 15*(Summer), 131-148.

Blankenburg Holm, D., & Ericksson, K. (1999). Creating value through mutual commitment to business network relationships. *Strategic Management Journal, 20*(467-486).

Blois, K. J. (1990). Transaction costs and networks. *Strategic Management Journal, 11*(6), 493-496a.

Boyd, B. (1990). Corporate linkages and organizational environment: A test of the resource dependence model. *Strategic Management Journal, 11*(6), 419-430.

Burgers, W. P., Hill, C. W. L., & Chan Kim, W. (1993). A theory of global strategic alliances: The case of the global auto industry. *Strategic Management Journal, 14*(6), 419-432.

Dyer, J. H. (1996). Specialized supplier networks as a source of competitive advantage: Evidence from the auto industry. *Strategic Management Journal, 17*(4), 271-291.

Garud, R., & Kumaraswamy, A. (1993). Changing competitive dynamics in network industries: An exploration of Sun Microsystems' open systems strategy. *Strategic Management Journal, 14*(5), 351-369.

Gulati, R. (1998). Alliances and networks. *Strategic Management Journal, 19*(4), 293-317.

Gulati, R. (1999). Network location and learning: The influence of network resources and firm capabilities on alliance formation. *Strategic Management Journal, 20*(5), 397-420.

Hagedoorn, J. (1995). A note on international market leaders and networks of strategic technology partnering. *Strategic Management Journal, 16*(3), 241-250.

Jarillo, C. J. (1990). Comments on *Transaction Costs and Networks*. *Strategic Management Journal, 11*(6), 497-499.

Madhavan, R., Koka, B. R., & Prescott, J. E. (1998). Networks in transition: How industry events (re)shape interfirm relationships. *Strategic Management Journal, 19*(5), 439-459.

Majumdar, S. K., & Venkatraman, S. (1998). Network effects and the adoption of new technology: Evidence from the U.S. telecommunication industry. *Strategic Management Journal, 19*(11), 1045-1062.

Nohria, N., & Garcia-Pont, C. (1991). Global strategic linkages and industry structure. *Strategic Management Journal, 12*(Summer), 105-124.

Robbins, J. A. (1993). Organization as strategy: Restructuring production in the film industry. *Strategic Management Journal, 14*(Summer), 103-118.

Shan, W., Walker, G., & Kogut, B. (1994). Interfirm cooperation and startup innovation in the biotechnology industry. *Strategic Management Journal, 15*(5), 387-394.

Stabell, C. B., & Fjeldstad, Ø. D. (1998). Configuring value for competitive advantage: On chains, shops, and networks. *Strategic Management Journal, 19*(5), 413-437.

Wade, J. (1995). Dynamics of organizational communities and technological bandwagons: An empirical investigation of community evolution in the microprocessor market. *Strategic Management Journal, 16*(Summer), 111-133.

APPENDIX II

Articles Published in Management Journals with Both a Social Network and an International Focus

Athanassiou, N., & Nigh, D. (1999). The impact of company internationalization on top management team advice networks: A tacit knowledge perspective. *Strategic Management Journal, 19*(1), 83-92.

Burgers, W. P., Hill, C. W. L., & Chan Kim, W. (1993). A theory of global strategic alliances: The case of the global auto industry. *Strategic Management Journal, 14*(6), 419-432.

Chen, H., & Chen, T.-J. (1998). Network linkages and location choice in foreign direct investment. *Journal of International Business Studies, 29*(3), 445-468.

Hagedoorn, J. (1995). A note on international market leaders and networks of strategic technology partnering. *Strategic Management Journal, 16*(3), 241-250.

Hundley, G., Jacobson, C. K., & Park, S. H. (1996). Effects of profitability and liquidity on R&D intensity: Japanese and U.S. companies compared. *Academy of Management Journal, 39*(6), 1659-1674.

Koka, B., Prescott, J. E., & Madhavan, R. (1999). Contagion influence on trade and investment policy: A network perspective. *Journal of International Business Studies, 30*(1), 127-148.

Money, R. B. (1998). International multilateral negotiations and social networks. *Journal of International Business Studies, 29*(4), 695-710.

Nohria, N., & Garcia-Pont, C. (1991). Global strategic linkages and industry structure. *Strategic Management Journal, 12*(Summer), 105-124.

CONNECTIONS 22(2): 22-29

A Core/Periphery Structure in a Corporate Budgeting Process

Noah P. Barsky [1]
Dept. of Accountancy, Villanova University

Many business organizations have adopted empowerment programs as a means of adapting to rapidly changing competitive environments. This paper examines the nature of ties among managers who participate in the budgeting process at a major apparel manufacturer. This case offers two primary contributions to social network analysis research. First, these data provide a concrete example of a context in which a core-periphery network structure can exist. Second, integrating the network data with ethnographic data to examine the structure of ties among managers is useful to help researchers to better understand the impact of corporate empowerment programs.

INTRODUCTION

In recent years, the academic and practitioner business literatures have documented and heralded the rise of the team-based, network-centered adaptive organization as a response to dynamic competitive market forces and emerging technology. Yet, limited research exists about the nature of managerial networks in firms in the current business environment. This paper examines the structure of ties among managers who participate in the budgeting process at a major apparel manufacturer.

Corporate budgeting presents a salient context to study managerial networks because it is a routine, widely-used, high-profile process that incorporates and impacts all organizational functions. Participatory or "bottom-up" budgeting is an example of a mechanism adopted by firms to promote employee empowerment and cross-functional interaction. Firms frequently adopt participatory budgeting programs to "empower" employees by allowing the workforce to set performance targets and allocate resources.

Consistent with these prevalent business practices, executives at our sample firm, to be referred to as Apparel Company (APCO), adopted a participatory budgeting process for fiscal year 1996-97. According to company executives, the primary goal of the process was to eliminate hierarchy and better integrate the network of managers. Unlike traditional "top-down" (hierarchical) budgeting, in a participatory budgeting process, the general managers should be most central and influential in the budgeting choices.

One common implementation difficulty that companies encounter is the ease in which the language of empowerment may be adopted, without substantively transferring control and substantive influence to front-line management. In such cases, employees may be involved in the mechanical aspects of the process, but often feel as thought their input was not substantively considered or incorporated into the final budget. Social network analysis enables us to examine the nature of ties amongst managers who participate in corporate budgeting. The structural data, when coupled with ethnographic data, can help to develop insight and understanding about corporate budgeting processes.

In this paper, social network analysis is used to examine two primary questions related to corporate budgeting at APCO. First, did the budgeting process work out as planned? Second, what type of structure emerged? The results are useful in helping researchers and practicing managers to understand the implications of adopting employee empowerment programs. The research site and data collection are discussed next.

THE RESEARCH SITE AND DATA COLLECTION

The Research Site

Data were collected from all 53 managers who participate in the budgeting process at APCO. APCO is a leading, brand-name designer and manufacturer of athletic performance and fashion clothing with annual sales of roughly $500 million. The U.S.-based company out-sources production to the Far East and Latin America. Hence, APCO's budgeting process de-emphasizes capital budgeting (asset acquisition) and focuses on developing an operating expense budget to meet sales goals. The centralization of services and employees at one North American location allows for a one-period, parsimonious research design that incorporates all of the managers who participate in the annual budgeting process (for more detail, see Barsky, 1999).

In a historical context, this study was conducted in an important year for APCO's budgeting process. For fashion and apparel designers, success and market growth can occur rapidly and can quickly transform the organization from a "niche" manufacturer to a major, large-scale international enterprise. Such events unfolded at APCO in the late 1980s and early 1990s. For nearly twenty years, the company had existed as a limited scale entrepreneurial venture. After experiencing rapid growth in the late 1980s, the company issued an initial public stock offering in the early 1990s. These events radically changed the size and scope of APCO's operations. In an effort to match company capabilities with global market demands, the firm underwent a rapid period of "professionalization" of its workforce. These changes primarily consisted of the hiring of professional accountants, production planners and a corporate marketing team.

In an effort to stabilize and formalize processes, such as budgeting, the CEO (and company founder) invested in the financial education of functional managers. The goal was to develop an integrated workforce, where employees would readily share bases of functional knowledge to achieve corporate performance goals. Budgeting was seen as a "flagship" process which could be used to impart strategic and financial goals, facilitate cross-functional communication, and enable the appropriate allocation of resources.

The finance group was directed to "financially educate" functional managers over a period of years through an extensive year-round use of profitability reports. During the budgeting process (in the years prior to this study), financial managers aided the operational managers in the basic techniques of how to develop a budget. Company management believed that, by the year of this study, the functional managers had gained a satisfactory level of financial understanding and experience to independently develop operating budgets based on their business knowledge. The financial managers were expected to act as coordinators and serve as a support function, as needed.

APCO's budgeting memorandum describes the process as "bottom-up" and one that should last about 2 ½ months. The memo encourages employees to, "Take Ownership. Be Proactive. This is your budget. You will develop it and ultimately be measured by it." The memo also specifically delineates each manager's responsibilities in the budgeting process. Clearly, the language in the memorandum and discussions with managers indicate that the budgeting process at APCO is intended to be bottom-up and participatory in nature. The design calls for a budget driven by the operating managers with executive oversight and final approval. In terms of the managerial network, one would expect general and front-line operating managers to be most central.

Data Collection

Data were collected in two primary phases. First, top managers were interviewed and documentation about the budgeting process was collected. These data were used to identify the managers who participate in the budgeting process and to refine the research questionnaire. Second, managers completed a research questionnaire that asked about demographics, rank, perceptions about budgeting process, and workplace contacts.

Network data were collected through "name-generator" type questions. Participants were asked to list the names of persons who (1) provide inputs or receive outputs of their routine work; (2) with whom they regularly talk about work-related activities; and (3) co-workers whom they consider to be friends or social associates. Wording of these questions in the questionnaire was intended to be as non-restrictive as possible, and consistent with instruments by Ibarra (1995).

Table 1
Work and Friendship Ties Among Members of APCO

```
                                  1111111111222222222233333333334444444444
                             5555
                             1234567890123456789012345678901234567890123456789
                             0123

             TECH_SERV       0000000000000000000000001000010110000000000001000000
 1          ACCTG_MGR1       0011
            ACCTG_MGR2       0011100000011111011100101000111100110001001111001
 2        BUDG_COORD1        1011
               PUB_REL       0101000000001001000101000001000100010000000110010
 3            DISTRIB        0000
   SOURCING_OUTERWEA         0110011110100111101000001010100100011101000110101
 4                   R       0000
       VPFIELDSALEOPS        0100000101000000010010101000110110110001011001001
 5  SR_SOURCING_PERFO        0011
                    RM       0001000000101000000000000100000011000000000000000101
 6  PROC_IMPR_ORDER_F        0000
                    UL       0001000100000001000000010100001010001000100001100 1
 7                MIS1       1010
              CUSTSERV       0001101000111000011010101000110110111011000011001
 8              CREDIT       0111
            ACCTG_MGR3       0001000000000001000000000010000000001000000000001
 9        SOURCING_DEV       0000
```

```
1             ACCTG4  0000110000010000010000000011001100001000000011101
0         PURCHASING  0000
1           SRVPSALES 0001000100010000010110000010000100001010000111000
1     GMSPECEVENTSBU  0001
1               MIS2  0100010101101000010010001011110100010001000000010
2          VPINTLSBU  1000
1             ACCTG5  0110000100010100010000000000100100000000000111001
3        VPMARKETING  1010
1       SRMGRTEAMSERV 0101000000001001100000000010100000000100000110000
4   GMACTIVEOUTERSBU  0000
1       PRODMGR_PERF  0101001010000001010000001000001100000100000111001
5               MIS3  0010
1          DIRSBUOPS  0111000000000110000001001000000100010101000010001
6       BUDGETCOORD2  0000
1            GMYOUTH  0001000000000100010000001000001100010011000011001
7            SRVPOPS  0011
1                CFO  0100100101111010101000100000100000001000000001001
8    DIRRETAILMARKET  1100
1 SRSOURCINGMGRACT-   0101000100000000010010000000100100000001100001100 1
9                 WE  1000
2       MARKETMGRADS  0110000000100000000000000010000000001000000000110
0          GMPERFSBU  0000
2               MIS4  0000100100110000001000101000111100010001010011000
1             ACCTG6  1000
2     DIRSPECMARKETS  0010000000000001000000000000000000000100000000000
2           GMHEADSBU 0000
2    CREATIVESERVMGR  0100101100000000010010011000110110110010101111001
3        NATSALESMGR  1101
2      DIRMARKETSERV  1000010000000000000000101000000100110011000010000
4       BUDGETCOORD3  0001
2         CONTROLLER  0101101100010011100010110011000001110001000010001
5      SRVPMARKETING  1100
2            PROJMGR  0000000010000000000000000000000000010000000000000
6    DIRRETAILSTORES  0000
2          TREASURER  0001000001110100000100001000011100011001000010100
7            VPSALES  0011
2              VPHRM  0010000001010000000000001000000100010001000010001
8                COO  0011
2                CEO  1101100100011100011010100000000100110111001111001
9                     0111
3                     0100101100010000000010100010000111000011000100001
0                     0110
3                     1100010001000010100010000010000101001000000011001
1                     0111
3                     1111111011110111010101100111100001101110011111 1
2                     1111
3                     0000100100000000000000100000010000100000011001000
3                     1000
3                     0000000000000000000000010000110000000000000000000
4                     0000
3                     0100100100000000000000111000100010010000011011000
5                     1000
3                     0111101110010001100010111111100100100001010111001
6                     1001
3                     0001000101100000001010000000100011000000000011100 1
7                     0000
```

```
3                            0001000000000111000000100000010000000000000000010101
8                            0000
3                            0000000100100000101000110000110100000000000110000
9                            1010
4                            0101101100010001101010011011110100010000001111001
0                            1111
4                            0000000000000000000001000000000100000000001001000
1                            0110
4                            0000100000000000000001000000000001011000000000000
2                            1000
4                            0100100000000000000001000001000101000011000010 10
3                            0000
4                            1111000000101110000000010000011010001101100001 1001
4                            0000
4                            0111001101101111101010111011101100111111000101001
5                            1011
4                            0100101101101010111010100000101110111001101110001
6                            1111
4                            0001010001000000000100000010000100000100000000001
7                            0100
4                            0010000000010000000100000000000100000000001000000
8                            0000
4                            0101111111001011111000101001111100011101000111100
9                            1111
5                            0100001000011000011010101000000110110011010011001
0                            0000
5                            0000000100000000010000101000111100000000110000 1101
1                            0010
5                            1100101100001010100000000001111110000000111000 11001
2                            0101
5                            1100100100100000010000001100111011000100010000 11001
3                            0010
```

Each participant's responses were coded as binary variables (i.e. presence of relationship) and entered into a name-by-name adjacency matrix . Since the research questions address group issues (rather than relationships between specific pairs), the matrix was symmetrized via the *maximum* method. This means that if either manager *i* mentioned manager *j* or the other way around, the tie between *i* and *j* was considered to be present. That is, x_{ij} and x_{ji} are both set to 1 if either $x_{ij} = 1$ or $x_{ji} = 1$. Symmetrization can be performed by UCINET 5 software (Borgatti *et al*., 1999). The symmetrized dataset is given in Table 1.

Participants also provided perceptions of other managers' relative influence, using a seven-point Likert scale. Participants consolidated responses to the name-generator questions into one non-redundant list. The participants then rated the relative budgetary influence of each member of their personal workplace network. Also, after listing workplace contacts, participants then listed the most influential managers and rated those persons' influence.

As a validation check, participants also rated their own budgetary influence relative to others. The mean self-ratings of influence and influence ratings assigned by others were moderately correlated (.3869, p<.01). This result suggests that participants were in relatively fair agreement with others' overall assessment of their influence.

ANALYSIS

Structural Analysis

A starting point in structural analysis is to examine a multidimensional scaling (MDS) of the graph-theoretic distances among all actors in APCO's managerial network. Figure 1 presents the resulting map.

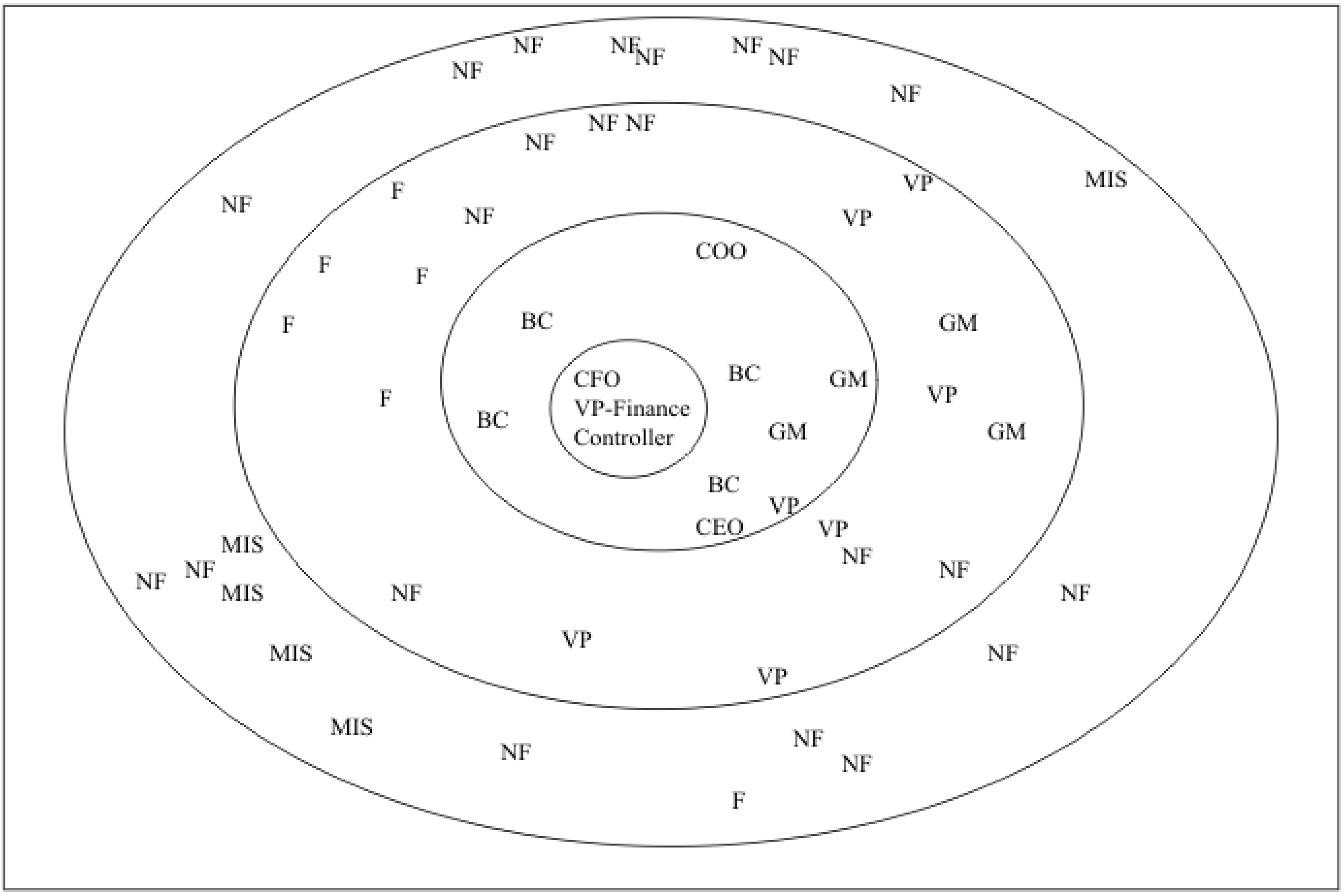


Figure 1. Geodesic Map of APCO's Managerial Network. Legend: BC=Budget Coordinator; F=Financial Manager; CEO=Chief Executive Officer; NF=Non-Financial Manager; CFO=Chief Financial Officer; VP=Vice-President; COO=Chief Operating Officer; GM=General Manager; MIS=Information Technologist.

The MDS of geodesic distances suggests a core/periphery (C/P) structure. In the context of a Euclidean representation, the defining feature of a C/P structure is a group of concentric circles where the average distance between points within the circles increases monotonically with distance from the center (Borgatti and Everett, 1997).

A C/P structure can also be characterized (Borgatti and Everett, 1997) by the absence of relatively exclusive cohesive subgroups (e.g., cliques). In other words, the network possesses just one tight subgroup – the core. An analysis of the clique structure of the network reveals few mutually exclusive groups, and instead hundreds of closely related cliques differing from each other by just one or two members. Similarly, hierarchical clustering analysis of the clique overlap matrix reveals – for the most part – a single core group, as shown in Figure 2.

```
             3   2   2 4 3 3 4 4 2 1 1   1 3   1 2 4 4 5   2 2 3 3 1 3 4 1 1 1 1 1 3   2 2 5 5   2
    Level    4 9 6 3 0 8 3 5 2 3 2 4 5 4 6 8 6 0 7 7 1 1 1 4 8 0 9 1 7 4 7 9 2 3 8 1 7 1 5 0 2 5 3 8
  -------    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  83.0000    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
  63.6667    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
  54.0667    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
  45.0000    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
  39.2800    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
  32.5633    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
  30.4512    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
  28.2278    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . XX
  20.6770    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . XXXX
  14.5600    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . XXXXXX
  13.6044    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . XXXXXX
  11.8981    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . XXXXXXXX
   9.0000    . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . XXX XXXXXXXX
   8.0000    . . . . . . . . . . . . . XXX . . . . . . . . . . . . . . . . . . . . . . XXXXX XXXXXXXX
   7.0000    . . . . . . . . . . . . . XXX . . . . . . . . . . . . XXX . . . . . . . . XXXXX XXXXXXXX
   6.0000    . . . . . . . . . . . . . XXX . . . . . . . . . . . . XXX . . . XXX . . . XXXXX XXXXXXXX
   5.6667    . . . . . . . . . . . . . XXX . . . . . . . . . . . . XXX . . . XXXXX . . XXXXX XXXXXXXX
   5.0000    . . . . . . XXX . . . . . XXX . . . . . . . . . . . . XXX . . . XXXXX . . XXXXX XXXXXXXX
   4.6419    . . . . . . XXX . . . . . XXX . . . . . . . . . . . . XXX . . . XXXXX . . XXXXXXXXXXXXXX
   4.3333    . . . . . . XXX . . . . XXXXX . . . . . . . . . . . . XXX . . . XXXXX . . XXXXXXXXXXXXXX
   3.6667    . . . . . . XXX . . . . XXXXX . . . . . . . . . . . . XXXXX . . XXXXX . . XXXXXXXXXXXXXX
   3.5195    . . . . . . XXX . . . . XXXXX . . . . . . . . . . . . XXXXX . . XXXXX . XXXXXXXXXXXXXXXX
   3.3333    . . . . . . XXXXX . . . XXXXX . . . . . . . . . . . . XXXXX . . XXXXX . XXXXXXXXXXXXXXXX
   3.0043    . . . . . . XXXXX . . . XXXXX . . . . . . . . . . . . XXXXX . . XXXXX XXXXXXXXXXXXXXXXXX
   3.0000    . . . . . . XXXXX . . . XXXXXXX XXX XXX . . . . . XXX XXXXX . . XXXXX XXXXXXXXXXXXXXXXXX
   2.5026    . . . . . . XXXXX . . . XXXXXXX XXX XXX . . . . . XXX XXXXX . . XXXXXXXXXXXXXXXXXXXXXXXX
   2.2382    . . . . . . XXXXX . . . XXXXXXX XXX XXX . . . . . XXX XXXXX . XXXXXXXXXXXXXXXXXXXXXXXXXX
   2.2143    . . . . . . XXXXX . . XXXXXXXXX XXX XXX . . . . . XXX XXXXX . XXXXXXXXXXXXXXXXXXXXXXXXXX
   2.0460    . . . . . . XXXXX . . XXXXXXXXX XXX XXX . . . . . XXX XXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXX
   2.0000    . . . XXX . XXXXX . . XXXXXXXXX XXXXXXX XXX . . . XXX XXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXX
   1.9778    . . . XXX . XXXXX . . XXXXXXXXX XXXXXXX XXX . . . XXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXX
   1.7448    . . . XXX . XXXXX . . XXXXXXXXX XXXXXXX XXX . . . XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   1.3333    . . . XXXXX XXXXX . . XXXXXXXXX XXXXXXX XXX . . . XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   1.1478    . . . XXXXX XXXXX . . XXXXXXXXX XXXXXXX XXX . . XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   1.0895    . . . XXXXX XXXXX . . XXXXXXXXX XXXXXXX XXX . XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   1.0000    . XXX XXXXX XXXXX . . XXXXXXXXX XXXXXXX XXX . XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.9675    . XXX XXXXX XXXXX . . XXXXXXXXX XXXXXXX XXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.8346    . XXX XXXXX XXXXX . . XXXXXXXXX XXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.8333    . XXX XXXXX XXXXX . XXXXXXXXXXX XXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.7337    . XXX XXXXX XXXXX . XXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.6667    . XXX XXXXX XXXXXXX XXXXXXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.5684    . XXX XXXXX XXXXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.3082    . XXX XXXXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.0873    . XXX XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.0356    . XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
   0.0000    XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

Figure 2. Hierarchical clustering (average method) of clique overlaps.

Borgatti and Everett (2000) suggest two additional ways to identify a C/P structure:

Two-class partition of nodes. By examining the group's adjacency matrix, researchers may simultaneously identify the core and periphery. The core is a dense, well-connected subgraph. Conversely, the peripheral nodes are loosely connected to each other, but connected to some members of the core. Essentially, such a model extends the well-known "Freeman star" to include a tight cluster of multiple group members at the center. Running the 2-class model in UCINET 5 for Windows (Borgatti *et al.*, 1999) yielded a fit (correlation criterion) of 0.390, which is moderately high.

Eigenvectors. This approach yields a continuous model in which each node is assigned some value of "coreness". Running the continuous C/P model in UCINET 5 yielded a fit (correlation criterion) of 0.533, which is quite high.

Members of the Core

Given the presence of the C/P structure, the next logical and interesting question is to consider why this type of structure exists at APCO. First, a review of data shows that rank was significantly correlated with coreness. A close examination of the network shows that the most "core" group consists of the three top financial managers – the Vice-president of finance (treasurer), the Chief Financial Officer (CFO), and the corporate controller. The next ring features the budget coordinators, the CFO and the COO. The non-financial managers, general managers, and a few vice-presidents are not in the inner rings.

The network structure appears to be quite similar to a topographical view of a traditional corporate hierarchy. In terms of the budgeting process, this network is quite revealing about the actual interactions among the players. However, the participative budgeting process was designed to invert the hierarchical pyramid. By definition, the espoused purpose of empowerment programs is to place front-line managers at the center of processes.

In this case, the finance function plays a highly central role in the network. The quantitative data indicated that coreness was highly correlated with rank and perceived budgetary influence. The budgeting process is designed to have operating managers play a significant role. The network graph suggests that coreness in the budgeting process is reserved for a set of individuals other than the operating managers.

DISCUSSION

This discussion will focus on why a core-periphery structure exists among the network of managers at APCO and how this structure can be used to help explain the outcomes of the budgeting process. Core-periphery structures can exist for two reasons. First, a core-periphery structure can arise simply as a function of indiscriminate ties and variation in degree. For example, in a pure commodity exchange market, a study of trader behavior is likely to reflect a core-periphery structure. That is, the variation in the degree of trading would cluster high volume traders and place relatively low volume traders along the periphery. Second, a core-periphery structure may reflect deliberate inclusion/exclusion of members based on a single attribute. The data suggest that this second case applies to the APCO managerial network.

Examples of attributes that may account for membership in the core include importance to the budgetary process, rank, (un)willingness to share (restrict access) to information and desire to maintain the corporate hierarchy. Integrating the structural data, interview data and responses to research questionnaires provides insights to explain the presence of this core-periphery structure.

This close clustering of the managers is consistent with interview responses which indicate that politics (i.e. isolated factions) are relatively non-existent at APCO. These managers also commented that high levels of communication across functions are common at APCO. Nonetheless, the interviewees suggested that senior management and top financial managers played the most significant role in the budgeting process.

Interviews with top managers at APCO indicate that asking employees to spend time on the budgeting process and not incorporating their input may have a detrimental effect on senior management's credibility. Many managers questioned whether the budgeting process at APCO was simply form over substance. This "pseudo-participation" undermined the credibility of the budget process and impaired employees' perceptions of top management. Many employees commented that a top-down process would actually be preferred over a process that asks employees to spend several working weeks on an activity that has little impact on outcomes.

For example, one of APCO's general managers commented:

> "[APCO's budgeting process is] bottom-up until it reaches senior management. You work to achieve an end…[then, budgets] get handed in and get kicked right back to you saying [here's the number and the allocation]."

Another top manager remarked that managers believed in the process:

> "We felt very good about the process from the bottom-up point of view. [Finance] gave people historical information as best [Finance] could … we felt pretty good that [managers] spent the time pulling together their budget…
>
> Instead, senior management...didn't get together…they didn't meet the deadlines. They were the ones who weren't getting together with their direct reports (SBU and cost center managers)…
>
> And before you knew it, the budget that was put together on a detailed basis was gone….It was all thrown to the wayside….it fell apart…all of a sudden the budget just became some top-down high level number…the only thing that we had was some massaged number that [went] into the cost center reports.
>
> [The] goal of the budget process was to make the budget be legit…how 'bout that? Actually, come to an agreement. That's where it fell apart when it went to a higher [management] level"

The manager commented that the events damaged the credibility of the process:

> "So, months and months of work and respect for the budget process was lost. Until mid-October when it was down at the first line management level….It was a process we could all be proud of….and by the time it was finished, we didn't even want to talk about it.

> We didn't care at that point. You should have never had us go through y'know 2-1/2 months of serious work putting together a budget if it didn't really mean anything to [top management] in the first place."

When asked about networks and communication, one top manager commented that the budgeting process initially helped functional managers better understand finance, and helped the financial managers better understand the business. However, the employees quickly realized the outcomes were directed and determined by senior management.

One manager summarized the events as follows:

> "We all still talk bottom-up...you want bottom-up...y'know, but the reality in a lot of large companies, and being here at [APCO] is that senior management…create(s) an overall top number and allocate that to this [strategic business unit]."

These comments shed light on why the central core is so influential at APCO. In essence, a "rhetoric-reality" gap was present at APCO. That is, while management captured the language of empowerment, the reality (as reflected in the core-periphery structure) is an environment of a strong command and control based hierarchy.

These results are supported by the data from a secondary validation instrument – the Management Communications Diagnostic Questionnaire (see Jablonsky, et al., 1993) – that indicates that the financial function maintains a "command and control" orientation. That is, financial managers focus on supporting top management and reinforcing the corporate hierarchy. As reflected in the budgeting process, top management and the financial team were reluctant to give up control. Hence, the process that was designed to promote employee empowerment served to reinforce the corporate hierarchy.

CONCLUSIONS

This paper presents research that examines the structure of ties among managers at a major apparel manufacturer. The results have implications for the social networks and management literatures. First, these data provide an example of a core-periphery structure in a defined social context. Applying this model to other data sets may be useful in developing richer insights about why C/P structures can develop in social settings.

These results have several important implications for management practices and research. First, many companies use participatory budgeting to "empower" employees. The structural data provide a basis for managers to compare designs to outcomes. In this case, APCO spoke of having budgeting in the hands of operating managers. However, the results show that while many managers participated in the budgeting process, a few top ranking finance officers constituted the core.

More generally, the results are indicative of a "rhetoric-reality gap" that is prevalent in the financial management of many organizations (Jablonsky and Barsky, 1999). For example, while the budget memos tell employees that "this is your budget; take ownership," control at APCO still rests with top management. The employees with the highest influence ratings and coreness values were senior executives and financial managers in the core. Social network analysis was instrumental in identifying this result.

Superficially, APCO appears to have strong participatory orientation. Our examination shows that the realities of the budget process network are actually quite different. Specifically, the results show that "capturing the language" of participation while maintaining a traditional "command and control" emphasis can impair management credibility and undermine the effectiveness of processes such as participatory budgeting.

While these results identified a "rhetoric-reality" gap with respect to budgeting at APCO, it should be noted that executives at APCO commented that the results of this study provided rich insights about employee perceptions. Company executives agreed that future budgeting efforts would attempt to more closely align budget goals, processes and outcomes. Participating managers felt the study offered an avenue to express these concerns and to develop an integrated perspective for future dialogue.

In conclusion, this study of participatory budgeting at a large, publicly-held U.S. company provides a concrete example of a C/P structure. The integration of structural and qualitative data provides a foundation for enriching academic and practitioner understanding of corporate programs. Management communication and control systems are playing an increasingly important role in corporate strategy and accountability. Simply adopting programs and practices is not a guarantee that a firm will be able to achieve its communication and performance objectives. Our data suggest that while participation is commonly considered to be beneficial for organizations, a "rhetoric-reality" gap may undermine management credibility and impair process effectiveness.

Last, this study demonstrates that senior management must be aware of the firm's underlying social structure. Implementing programs solely related to formal hierarchy may be inadequate to achieve desired goals. This research identifies important considerations regarding the tension inherent in balancing empowerment and control in the modern organization that attempts to adapt to increasingly competitive markets.

REFERENCES

Barsky, N.P. 1999. *Organizational Determinants of Budgetary Influence and Involvement*. New York: Garland Publishing.

Borgatti, S.P., M.G. Everett and L.C. Freeman 1999. *UCINET 5 for Windows*. Boston, MA: Analytic Technologies.

Borgatti, S.P. and M.G. Everett. 1997. "Models of Core/Periphery Structures." Sunbelt International Social Networks Conference, 2/97, San Diego, CA.

Borgatti, S.P. and M.G. Everett. 2000. Models of Core/Periphery Structures. *Social Networks* (forthcoming).

Ibarra, H. 1995. Race, Opportunity, and Diversity of Social Circles in Managerial Networks. *Academy of Management Journal, Vol. 38*, No. 3, 673-703.

Jablonsky, S.F. and N.P. Barsky. 1999. Integrating Strategy, Control and Accountability. *Corporate Finance Review, Vol. 4*, No. 2, 27-35.

Jablonsky, S.F., P.J. Keating, and J.B. Heian. 1993. *Business Advocate or Corporate Policeman? Assessing Your Role as a Financial Executive*. Morristown, NJ: Financial Executives Research Foundation.

CONNECTIONS 22(2): 30-36

Boundedness and Connectivity of Contemporary Families: A Case Study[1]

Eric D. Widmer and Linda-Ann La Farga
University of California, Irvine

We present a rationale for considering significant family units as open low-density ego networks rather than as small close-knit groups. Using a case study approach, we show that individuals are likely to have significant family members who are not strongly connected to each other, and whose own significant family members differ. However, based on relationships among parents and children, family networks follow well- known sociometric tendencies such as reciprocity and transitivity of choices. We further underline some implications of our findings for family research.

INTRODUCTION

For a long time, research has considered the question of family boundaries as settled, using either a predefined set of family roles or a common residence as valid criteria for defining what is the significant family unit. More recently, however, researchers interested in recomposed families have underlined that those boundaries are not obvious, because divorce and remarriage have created ties among different households and have extended the set of family roles. Another trend of research has emphasized that the connection between adults and their family of orientation are well developed and functionally important, and independence between family of orientation and family of reproduction is no longer taken for granted. Concepts such as “the modified extended family” (Litwak, 1960), “the new extended family” (Furstenberg, 1987), or the “remarriages’ chain” (Cherlin & Furstenberg, 1994) suggest that many contemporary families are not nuclear in nature. However, researchers have not yet drawn the logical conclusion from those observations. If it is true that one can think of contemporary families as chains of relations, one can, and indeed should, use a network approach to study them. Up until now, such attempts have been extremely limited in number (for instance, Jedlicka, 1977).

In previous research (Widmer, 1997; Widmer, 1998), the senior author asked 25 female college students to define their significant family members. It was found that lists of significant family members not only included cohabiting parents and children, but also non-cohabiting siblings, grandparents, other kins, and friends considered as family. The variety of roles encountered was great, leading to the conclusion that it was difficult to define boundaries of significant family units in terms of residence or with a limited set of roles. Interestingly, a large number of family members were reported by interviewees not to be strongly connected to each other. However, because this previous study was based on a single interview per family, it could not properly illustrate the fact that families are likely to be unbounded, low-density ego-networks, rather than the small close knit groups to which research often refers.

Expanding on our previous work, we undertook a case study, interviewing all the significant family members of a twice-divorced female. A case study approach is appropriate for exhibiting the operation of some general theoretical principle (Mitchell, 1983). In this case, we want to show that, by their very nature, contemporary families tend to be low density, unbounded networks rather than small, bounded groups with a high density of interactions. Thus, we are mainly interested in two questions. First, to what extent do people cited as significant family members by someone cite the same persons as significant family members? Second, to what extent do these persons cite each other as significant family members? These questions further assess the degrees of boundedness and connectivity of contemporary families. A limited overlap between definitions of significant family members and a small number of intercitations are expected to exist among family members, especially if some of them are divorced or remarried.

DATA

Ego (Betty) is a 54 year old, Hispanic female working as a social worker in a state-run facility for abused and neglected children. She lives in a middle-sized American town located on the West Coast and has been married twice, with both marriages ending in divorce. Betty has two children from her first marriage, as well as a fairly large family of orientation consisting of three sisters, two brothers, and one mother. Ego was chosen primarily because she has experienced both divorce and remarriage. The fact that Ego feels she lives in a cohesive family and has not reported any major family conflicts also confirms that our results are not caused by some family dysfunction.

In order to know whom the significant family members of Ego are, we draw from research which was mostly concerned with the definition of stepfamily boundaries (Furstenberg, 1987). We created a name generator, which reads:

> ‘*Give me the first name of the people in your family who are significant for you at this time*’.

The question then specifies:

> ‘*By significant, I mean those people in your family who have played a role, either positive or negative, in your life during the past year. As I just said, I am not only interested in the people that are significant to you because you love them or respect them, but I am also interested in those who have upset you or made you angry during the last year*’ (Widmer, 1998).

Ego cited nine significant family members who also had to report on who were their significant family members. Then, using a snowball sampling technique, we interviewed any person who was cited by at least two persons cited by Ego, on the assumption that they are actors in the group dynamics.

Table 1
Ego's Network

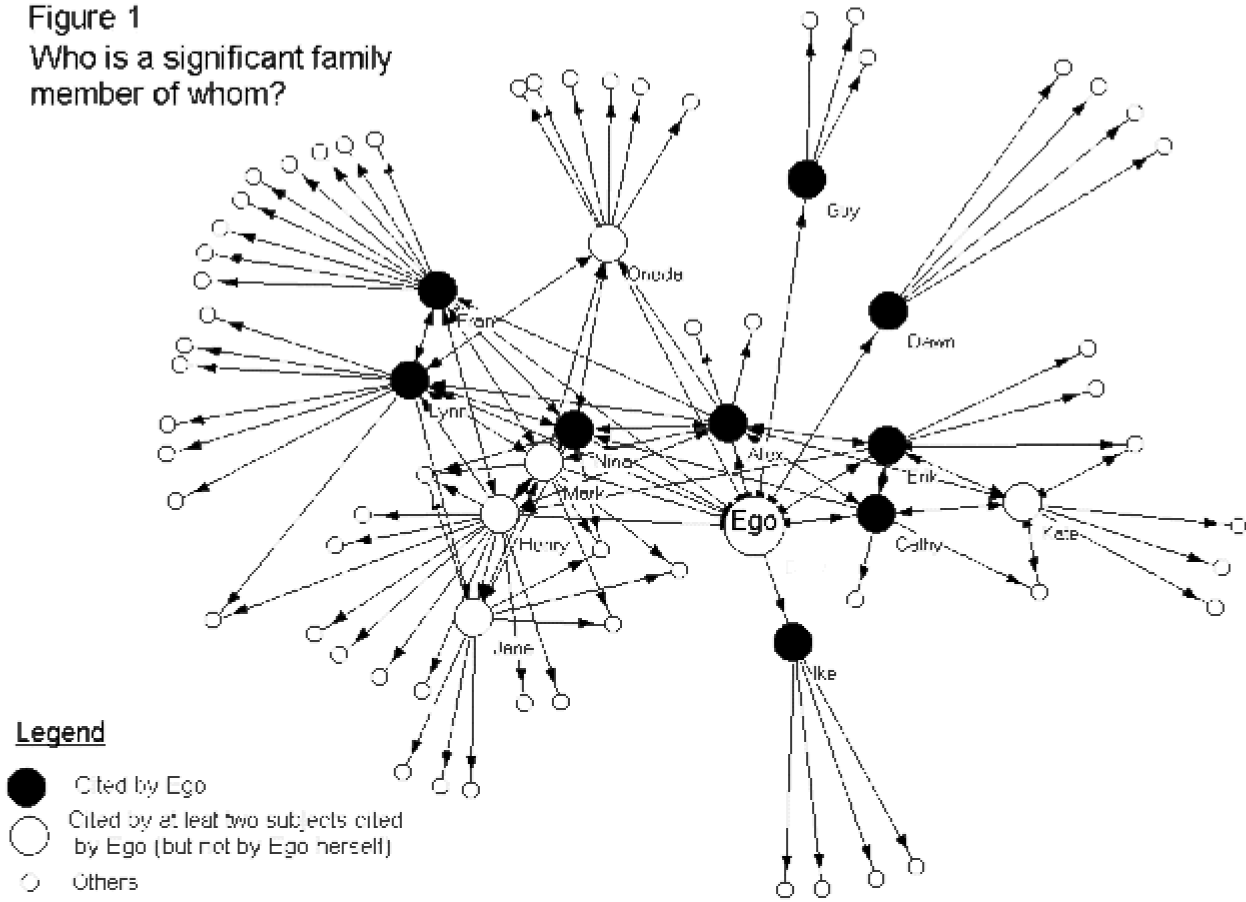
Name	*Description (relation, age, sex, education, occupation)*
Betty	Ego, 53, female, high school, social worker
Alex	Son, 27, male, high school, dock worker
Cathy	Daughter, 24, female, in college
Dawn	Friend/co-worker, 24, female, high school
Erik	1^{st} ex-spouse, 52, male some college
Fran	Middle sister, 42, female, vocational training, office manager
Guy	2^{nd} ex-spouse, 51, male, college, disabled veteran
Henry	Youngest brother, 49, male, college, construction worker
Ike	Friend/co-worker, 32, male, college, program director
Jane	Eldest brother's 2^{nd} wife, 36, female, high school, bank teller
Kate	Ex-mother-in-law, 72, female, high school, retired homemaker
Lynn	Eldest sister, 46, female, law school, civil attorney
Mark	Eldest brother, 50, male, college, computer technician
Nina	Mother, 74, female, high school, retired homemaker
Oneda	Youngest sister, 42, female, high school, insurance claim authorizer

Ego cited an extended range of role-persons as significant family members: her two children, her two ex-husbands, two close friends, her mother, and two sisters. After interviewing these individuals, we added five subjects, who were cited by at least two persons cited by Ego. They were three other siblings, a sister-in-law, and Ego's former mother-in-law from her first marriage. Information on these fifteen core family members of Ego is presented in Table 1.

RESULTS

Figure 1 depicts the family network up to a distance of three from Ego. The larger nodes represent the fifteen core members of Ego's family. Those nodes in black indicate the persons that have been cited directly by Ego as significant. This graph shows a total of 79 persons cited with 145 arcs linking them together.

Figure 1
Who is a significant family member of whom?



This network is 46% male. Eleven percent of subjects are younger than 20 years old, 34% between 20 and 40, 41% between 41 and 60, and 14% are older than 60. Eighteen percent of subjects live in the same county as Ego, 60% in the same state but not in the same county, and 22% in another state. It is striking to note that 42 of the 79 persons mentioned (53%) have no recognized tie with Ego.

Boundedness

From Figure 1, one can see that Ego's significant family is not a bounded closed group, but a rather widespread unbounded network. A first confirmation stems from the fact that among the 79 persons cited, 59 were cited by only one subject. To further assess the boundedness of this network, we computed a matrix of citation matches among the 15 core members.[2] The overall match is .14, meaning that on average core members share 14% of their own network with others. When only the nine persons cited directly by Ego, plus Ego, are considered, this percentage remains about the same (15%), as is the case when Ego is taken out of her network (13%).

Let us also point out that 27 of the 105 dyads included in the core family network do not share any significant family members at all, and 25 share only one person, Ego in 21 cases. Two thirds of the pairs (70 pairs) share less than four persons as significant family members, and only 10% report sharing more than 6 members. On average, each pair shares only 2.57 members. Those different estimates suggest that the degree to which each person's family network overlaps with others is weak, thus lending support to the conclusion that Ego's family network is not clearly bounded.

Connectivity within the core

We now want to address the issue of intercitations among the 15 core family members of Ego. Density of citations within the core is .34. This means that many members do not consider each other as significant. If we measure the density only among Ego and subjects who were cited as significant by Ego (for a total of 10 subjects) the density is .38. If we get rid of Ego, who obviously cites every other person and is over-cited because of the study design, the density of citations becomes only .24, that is, only about a quarter of the possible links among the 9 persons directly cited by Ego exist.

To measure the degree of significance of other core family members in a more precise way, we asked each subject to report the degree of closeness they feel toward each other member.[3]

Results for the core network, as well as for only those subjects cited directly by Ego, are reported in Table 2.

Table 2
Closeness in the Core Family Network of Ego

Closeness among family members of Ego	Codes	Fifteen core members (n=210)	Nine subjects directly cited by Ego and Ego (n=90)	Nine subjects directly cited without Ego (n=82)
Very close	1	15%	18%	8%
Close	2	17%	17%	15%
Somewhat close	3	16%	16%	14%
Acquaintance	4	12%	18%	22%
Barely known	5	11%	10%	13%
Not known at all	6	29%	22%	28%
Total		100 %	100%	100%
Average		3.7	3.52	3.98

Again, we face a similar result as with the citations of significance. Only a minority of relationships are among 'very close' or 'close' persons. Two thirds of the relationships among core family members of Ego are considered 'somewhat close' or less. Let us stress that this result holds true even when one considers only subjects cited by Ego. It is especially noteworthy that 28% of the subjects do not know each other at all.

Influence of Family Roles

We now take a closer look at the correlation existing between intercitations, overlap and family roles. Table 3 reports measures of boundedness and connectivity according to family roles.

Table 3
Family Ties, Overlap and Intercitations

Family ties	Number of relations	Average overlap	Proportion of mutual citations	Average closeness
Parent-child	11	.40	100%	1.54
Spouses/ex-spouses	3	.24	100%	1.5
Siblings	16	.27	56%	2.09
Kins	14	.23	29%	2.34
In-laws	8	.08	0%	3.94
Former steps & inlaws	19	.07	11%	4.05
Others	36	.02	3%	5.53

In this family, biological parents and biological children cite each other systematically as significant others; they also report a high level of closeness and their definition of significant family members overlaps much more than for other categories (even though it is still far from 1.0). Spouses and ex-spouses cite each other as well, although this result should be considered cautiously since it is based only on three relations. Results for other categories are much more nuanced. Even though siblings constitute a large number of the existing ties, the level of reciprocity in citations is only 56%. In-laws, former steps and former in-laws have a very low overlap, do not have many mutual citations, and report an average closeness at the 'acquaintance' level. Other relations (such as former in-laws or friends of siblings, etc.) are almost completely disconnected and show almost no overlap.

In order to know when transitivity[4] holds and when it fails to hold, we further investigated the roles of persons shared by any pair of the core network, as well as the roles of persons that are specific to just one member of the pair. It was striking that parents and their child always shared the other parent, even in cases of divorce. Sharing of another child (sibling of the child) was also frequent although less likely. However, parents' parents and parents' siblings, parents' friends and parents' new spouses, were much less likely to be shared as significant by the child.[5]

Spouses and ex-spouses were most likely to share their common children. However, they usually did not share their own children, their siblings, or parents. These crude tendencies point to the fact that marriage and remarriage do not create much transitivity. Furthermore, it suggests that the parent-child link is the backbone relationship on which family networks are built. Relations among siblings, kin, and divorced parents are likely to be explained by a tendency of the parent-child relationship to create transitivity.

CONCLUSION

Our study points to the fact that neither boundedness nor connectivity can be taken for granted in contemporary families. However, this does not mean that we face unstructured families. Transitivity of parent-child relationships, along with other structural properties still to be discovered, probably account for much of the dynamics by which each family network is shaped.

These results may modify the current understanding of family dynamics. If significant family units are chains of relations, then the type of social integration one can expect from them is different from that which is provided by smaller and more connected groups. Social control, social support, conflicts, material exchanges, and power structures of families are likely to have features discovered in other types of networks (for instance, structural holes, differential centrality of actors, cliques, etc.). Influence of those features on individual adjustment, especially for children with divorced or remarried parents, is worth studying.

Measurement issues are also raised by our findings. To measure properties of family systems, researchers have extensively used scales based on the assumption that well-defined family units exist. Scale items such as *"Family members consult other family members on their decision*", or "*we like to do things with just our immediate family*" (Farrel & Barnes, 1993; Olson, McCubbin et al., 1983), which are supposed to capture family cohesion and adaptability, have little relevance if families are low density unbounded networks. As a matter of fact, had we asked any pair of the 15 core members to estimate those items, we would have gotten reports concerning different sets of people, thus questioning the validity of such scales. Of course, it is always possible to define beforehand who is going to be included as part of "*the family*" and make it clear to interviewees (in considering, for instance, the household as the relevant unit). However, we are again faced with the problem of having some or most of them with significant family members not included in the *a priori* definition.

Although our results are based on a case study, we believe the conclusions we draw to be true for a large number of families. Divorce and remarriage obviously affect the transitive closure of families. Inclusion of kin and pseudo-kinship ties has the same effect. The lack of transitivity of relations with in-laws and steps confirms the hypothesis that families are not close-knit bounded groups, but rather chains of relations. Further studies should be conducted to confirm, extend and refine, what is presented here. Using a network perspective to study family dynamics is necessary, since a majority of persons, due to the increase of divorce and remarriage (Glick, 1989), are likely to deal with those unbounded low-density family networks in their everyday life.

REFERENCES

Borgatti, S.P., M.G. Everett, and L.C. Freeman. (1991). *UCINET*, version IV. Columbia, SC: Analytic Technology.

Cherlin, A.J., and F.F. Furstenberg, (1994). Stepfamilies in the US: a reconsideration. *Annual Review of Sociology*, 359-381.

Farrell M. P. and Barnes (1993). Family systems and social support: A test of the effects of cohesion and adaptability on the functioning of parents and adolescents. *Journal of Marriage and the Family*, 55, 119-132.

Furstenberg, F.F. (1987). The new extended family: The experience of parents and children after remarriage. In: Pasley K., Ihinger-Tallman, M. *Remarriage and Stepparenting: Current Research and Theory*, 42-61. (New York: Guilford).

Glick, P. (1989). Remarried families, stepfamilies and stepchildren: A brief demographic analysis. *Family relations*, 38: 24-27.

Jedlicka D. (1977). *Serial marriage networks: A study of change in the American Family.* University of Georgia, unpublished Manuscript.

Litwak, E. (1960). Geographical mobility and extended family cohesion. *American Sociological Review*, 25, 385-394.

Mitchell, J.C. (1983). Case and situation analysis. *The Sociological Review*, 31(2), 187-211.

Olson, D.H., H.L. McCubbin, et al. (1983). *Families: What makes them work.* (Beverly Hills, CA: Sage).

Widmer E.D. (1997). Families as cognitive networks: Theoretical and methodological propositions. Paper presented at *the National Council on Family Relations, 59 Th annual conference program*. November 7-10, 1997, Arlington, VA.

Widmer E.D. (1998). Families as networks: A structural approach of family relationships (Under review for publication).

CONNECTIONS 22(2): 37-49

Normative Versus Instrumental Functions: Evidence of Social Network Differentiation Among Rural Kenyan Men [1]

Steven R. Green
Policy Research Division, The Population Council, New York2

INTRODUCTION

The study of communications networks is marked by a relative paucity of descriptions of network characteristics of rural African men. It might also be observed that the significance of these network attributes for interpreting processes of social change is also understudied in the developing countries. This paper seeks to address both of these shortcomings in the literature by providing a detailed description of the communication networks of one group of rural African men, members of the Luo ethnic group of Kenya's Nyanza Province, and by analyzing the possible implications of these network attributes for social change.

Specifically, the network attributes of individual Luo men will be compared for two topics related to fertility change, the process by which a society moves from a relatively high average number of births per woman to a substantially lower average number of births per woman. The first network centers on the people with whom the respondent discusses the issue of children supporting their parents in old age. The second network focuses on family planning and the people with whom the respondent discusses the desire for fewer children, the use of contraceptive methods, the side effects of these methods, and so forth. This paper seeks to answer two questions: 1) Can the two networks be differentiated on the basis of their normative and instrumental functions?; and 2) What does this distinction in network characteristics suggest about the future of fertility change among the Luo?

Both semi-structured interview data and survey data confirm that there is considerable differentiation in the composition of the two networks for individual Luo men, and many of these discernible differences suggest that Luo men discuss these two issues with somewhat different types of partners. The evidence indicates that the parental support network serves more of a normative function, while the family planning network plays more of an instrumental role. Both normative and instrumental changes are integral components of any process of social change; each of these elements is, therefore, essential for fertility decline. The fact that Luo men are discussing both the normative context and the instrumental factors associated with the decision to have fewer children augers well for the pace of future contraceptive adoption and fertility decline among the Luo.

THE LUO OF KENYA

The Luo migrated from the Sudan and settled near the eastern shore of Lake Victoria during the last 500 years (Shipton, 1989, p.16). They now live primarily in three districts of Kenya's Nyanza Province, and number over 2.5 million (Shipton, 1992, p.357). The Kenyan census of 1979 measured Luo population growth at approximately 4 per cent per year (Shipton, 1992, p.361), and Shipton (1995) describes the Luo as having "one of Africa's denser rural populations (now averaging over 200 per sq. km.)" (p.166). Rapid population growth among the Luo has resulted in high population densities and a strong contraction in the size of individual landholdings (Hoddinott, 1992, p.562). In addition, the Luo have stood in opposition to Kenya's national government since independence in 1963 and have suffered relative neglect in terms of development and governmental allocation of resources (Shipton, 1995, p.166; Francis and Hoddinott, 1993, p.116). This excerpt, from a 47 year old stonecutter and cotton farmer with 4 years of primary education, demonstrates what many Luo men perceive as the economic difficulties they face:

> "The goodness of having many children? Yes, we have talked about it. If you have 4 or 5 children, not all of them will go astray, you might find one who is willing to support you, if you educate all of them. Even if you don't educate them, you will find at least one who can really do the farm work. You can find people saying that the son of so-and-so is really a hard worker. They all want to be lazy. The badness is that you might have only a few acres of land, and each son will claim land. You might not have wealth and as such will leave your children in poverty. All your sons want to get married, and you have to give each one of them a cow [for their brideprice]. All of them want to go to school, and yet you can't afford to pay fees." [Kawadhgone 5]

The Luo are characterized by patrilineal descent (Shipton, 1992, p.362), virilocal residence (Shipton, 1992, p.362; Hoddinott, 1992, p.548), and a relatively high level of polygyny. They combine the growing of staple crops, primarily maize, millet, sorghum, pulses, cassava, green vegetables, sweet potatoes, and bananas, and the production of cashcrops, such as groundnuts, sugarcane, tobacco, and cotton, with the herding of cattle, sheep, and goats (Francis and Hoddinott, 1993, p.120). The main sources of income in the four research locations, aside from selling agricultural products, include fishing, stone cutting, and skilled and unskilled labor (e.g., as carpenters, masons, thatchers, agricultural workers, etc.).

NORMATIVE VERSUS INSTRUMENTAL FUNCTIONS OF COMMUNICATIONS NETWORKS

Mitchell (1969) distinguishes between two functions of an individual's communication networks. The first of these functions is to help the individual define norms and maintain a set of attitudes consistent with other network members (Mitchell, 1969, pp.36-37). These normative networks tend to contain more kin members and intimate friends. A communication network may also serve a second function as an instrumental means for achieving desired ends (Ibid., p.38). Mitchell asserts that normative communication networks tend to be more stable, while instrumental communication networks tend to be mobilized specifically to address a particular need or objective. He posits that instrumental networks "may be looked upon as an aspect of a personal network isolated in terms of a specific short-term instrumentally-defined interactional content: the personal network itself is more extensive and more durable" (Ibid., pp.39-40).

Moore (1990) asserts that, "A broad range of ties, with many strong or weak connections to diverse others, is often seen as a valuable instrumental resource, while network density is more closely associated with social support" (Moore, 1990, p.728). Marsden (1987) also reports that, "[I]nfluence processes and normative pressures operate through intimate, comparatively strong ties" (p.123).

The hypothesis guiding this analysis was that the attributes of parental support network partners, taken in the aggregate, would provide evidence that these conversations serve more of a normative function, while the characteristics of the family planning network partners, taken in the aggregate, would suggest that family planning discussions play more of an instrumental role. It is important to recognize, however, that the parental support network is likely to contain a considerable instrumental component, while the family planning network is sure to feature a strong normative content. The question, therefore, is not whether or not these two networks are examples of "pure" normative and instrumental networks; rather, the question is whether or not significant differences along these lines can be discerned between them.

EXPECTATIONS OF LUO NETWORK CHARACTERISTICS

As little literature exists describing network attributes of rural, sub-Saharan African men, we can speculate about what they might be, based on studies conducted in developed countries. Marsden (1987) finds that network range (i.e., diversity and heterogeneity) is greater in urban, as opposed to rural, settings (p.129), so we might expect rural, Luo men to have fairly homogeneous network partners. Fischer (1982) also reports differences in kin ties between rural and urban Americans (p.83), as well as significant differences in network densities, with rural respondents reporting more kin ties and much denser networks than urban respondents (p.156).

Network density, which ranges from 0.0 to 1.0, is a measure of the degree to which the network partners named by a respondent know one another. The degree of density is important, because it indicates the likelihood that new information and influences will, or will not, permeate a network. In a highly dense network, all or most of the people named by an individual know one another, while in a less dense network the network partners are less likely to know one another.[2]

Network partners in dense networks tend to share background characteristics (e.g., age, education, socioeconomic status), spend more time with one another, talk to the same people, read the same newspapers, listen to the same radio programs, etc. Network partners in less dense networks, conversely, tend to have more diverse background characteristics. This means that a respondent who is at the center of a less dense network is more likely to be exposed to a greater variety of information than a respondent at the center of a highly dense network.

It might be taken for granted, based on research findings from developed countries, that the characteristics of the communication networks of rural, Kenyan men would tend to be fairly predictable and homogeneous, with high numbers of kin ties and high densities. As the Luo live in villages that are composed largely of clan members (i.e., relatives or kin), one might expect the communications networks of these men to be composed almost exclusively of kin members who know one another well (i.e., high density). Furthermore, two of the four field sites included in this research are in remote areas of Nyanza Province, accessible only by poor roads and footpaths or boat; a third field site is accessible by better roads, but travel within the area is generally restricted to poorly maintained trails. For the residents of these three sites, opportunities to develop extensive networks outside their vicinities appear to be limited. We might expect that the characteristics of the two networks for individuals would be nearly identical irrespective of the subject matter, as we could anticipate a great deal of overlap in the network partners who comprise the two networks.

On the other hand, we might expect that differences in the subject matter of the two networks would result in discernible differences in the characteristics of the two networks, despite the impediments to travel to and within these areas. It is plausible that discussions of parental support would necessarily involve network partners with age and sex characteristics significantly different from the family planning network partners. If this is the case, there should be less overlap in the conversational partners of the two networks, and there should be discernible differences in their network attributes. If so, one must assume that the two topics of discussion are viewed as somewhat distinct by the respondents; to some extent the respondents are engaging different types of conversational partners on the two issues. If so, this could have important implications for predicting the pace of fertility decline among the Luo.

The Luo are characterized by a traditional, patriarchal social system, in which older men hold a disproportionate share of control over the distribution of resources. Shipton (1989) calls Luo society a gerontocracy, in which elder males control bridewealth, land, and, to a lesser degree, labor, and he asserts that in Luo ideology, age, wealth, and respect are intertwined (p.20). Parkin (1978) describes the normative control of older men in this passage, "When young Luo monogamists equivocate about the advantages of taking a second wife, elders place repeated stresses on the word for polygynist, which has the additional connotation of arbiter and man of eminence and authority" (p.25).

Parkin asserts that, "This opposition between, from a male viewpoint, threats of disorder and reassertions of order is most vividly expressed in Luo public life and speeches, which are almost exclusively the domain of men" (Ibid., p.29). Elsewhere, Parkin argues that older Luo men act forcefully to oppose practices they see as violating 'proper' Luo conduct (Ibid., pp.215-216; p.287). Arguing about the foreign notion of private landholdings, Shipton (1992) asserts that the British found that implementing a plan to register and title landholdings among the Luo in the 1930s was met by opposition from "*ad hoc* assemblies of Luo spokespeople (nearly all men...)" and that even today, "Elders insist that lending at interest...is a foreign idea that came only during the colonial period" (pp.363-364).

NORMATIVE DISCUSSIONS AND FERTILITY BEHAVIOR: THEORETICAL CONSIDERATIONS

Parkin and Shipton agree that in Luo culture older men are considered the proper individuals to evaluate and pass judgment on normative matters. Thus, the normative issue of parental support should fall within their purview. There is a strong theoretical argument in the demographic literature that men's expectation of old age support is one of the cornerstones of high fertility in traditional, patriarchal societies (see, for example, Caldwell,1976, pp.343-345; Mason and Taj, 1987, pp.614-615). Caldwell and Caldwell (1987) contend that the expectation of future support from children "is emerging as the strongest support of all for undiminished fertility" (p.421). Throughout Luoland, however, the socially-sanctioned practice of children supporting their parents in old age is weakening, reshaping the normative landscape.

If the parental support network more closely resembles a communication network with a normative function, it should be composed of more older and male network partners who know one another well (i.e., higher density) and more kin members. As the issue of parental support is inextricably tied to the issue of the value of children and, therefore, the normatively correct number of children one should have, the parental support discussions between older and younger men have implications for desired family sizes, contraceptive use, and fertility decline among the Luo. If, on the other hand, the family planning network is correctly conceived of as having an instrumental function, it should contain more heterogeneous network partners, including more women, more network partners who are agemates of or younger than the respondent, more nonkin, and fewer network partners who know one another well (i.e., lower density).

This analysis uses data from both a survey and semi-structured interviews. The survey data are communication network data collected among Luo men of Kenya's Nyanza Province from December 1994 through January 1995. The survey data represent two separate, but related, communication networks (i.e., parental support and family planning). These Luo network data are ego-centered, meaning that the information about conversational partners is collected for each respondent; there was a total of 709 respondents. These survey data provide a snapshot of individuals' conversational partners, rather than a more detailed, but limited, view of the network partners of all members of a bounded geographical area.

Such whole-network data provide extensive insight into the way that all individuals interact within a bounded setting (see Montgomery and Casterline, 1993, for instance), but must be artificially truncated at some geographical marker, such as the village limits. The ego-centered approach adopted here, in contrast, asks about all conversational partners with whom the respondent has discussed the relevant issue. As a result, conversational partners can include people who live great distances from the respondent, those who are seen infrequently, and others who would be less likely to be captured if data collection were limited to a geographic setting that utilized saturation sampling (i.e., asking about all of the potential discussion partners in the bounded area).

The network data are supplemented by qualitative data collected in Nyanza Province, Kenya in June and July, 1994. A total of forty semi-structured interviews were conducted, ten in each of four locations. While the data were collected using a sampling strategy that attempted to gather the data from a representative sample, the realities of conducting research in the field necessitated that the data be collected from respondents who could be located in a short time frame. The sample of men selected for qualitative interviews is biased, therefore, against men who were more likely to be away from their homes for extended periods of time. As a result, the qualitative data do not come from a representative sample of Luo men, and are presented for illustrative purposes only.

The inclusion of qualitative data addresses one of the limitations of network analysis, as elaborated by Barnett, Danowski, and Richards (1993). These authors lament that traditional communication network research has treated "message content as transparent or unimportant" (p.10). They assert that the majority of network researchers are not trained in communication science, so they are likely to ignore "the concept of message transmission or information exchange [that] is perhaps the defining concept of the discipline" (Ibid.). Network research has generally concentrated on the characteristics of network partners and the overall frequency of communication, at the expense of its content. The addition of qualitative data, although utilizing small samples that are not generalizable to a larger population, permits us some insight into what information is exchanged with whom, and the influence of these conversations on attitudes and behavior concerning parental support and family planning.

EVIDENCE OF DIFFERENCES IN NETWORK CHARACTERISTICS IN THE QUALITATIVE DATA

All of the parental support and family planning discussions described by the semi-structured interview respondents were tallied, in order to get a quantitative sense of the content of the discussions and the attributes of the respondents' discussion partners. Respondents were asked to elaborate upon the specific content of the conversation and to describe the characteristics of the discussion partner(s). The qualitative data are particularly helpful in exploring the content of discussions that are glimpsed only indirectly through the survey data.

Table 1.
Comparison of Selected Measures of Parental Support and Family Planning Networks of Luo Men, from Semi-structured Interviews, Nyanza Province, Kenya, 1994

Measure	Categories	Parental Support Network (%)	Family Planning Network (%)
Discussion Type	Normative	59.6	24.9
	Instrumental	22.9	44.4
	Normative-Instrumental	17.4	30.7
Sex of Discussion Partner	Male	90.3	65.1
	Female	9.7	34.9
Age of Discussion Partner	Older than respondent	43.3	27.5
	Agemate of respondent	42.3	45
	Younger than respondent	14.4	27.5
Relationship with Discussion Partner	Relative	46.5	43.5
	Friend	44.7	46.3
	Acquaintance	8.8	10.2

Comparisons of the measures for the two networks from the qualitative data are presented in Table 1. Beginning with the measure of discussion type, the two networks can be distinguished from one another by the content of their discussions. The parental support discussions are more than twice as likely as the family planning discussions to have a normative content, while the family planning conversations are nearly twice as likely to have an instrumental content. This result is consistent with the expectation that parental support discussions would be more likely to focus on normative issues, while the family planning discussions were expected to focus on instrumental matters.

An even more telling difference between the two networks involves their sex compositions. According to the qualitative data, women constitute more than one-third of the family planning discussion partners, while comprising less than one-tenth of the parental support discussion partners. Women are more than three times as likely to be named as family planning network partners than as parental support network partners. For Luo men, the two subject matters are clearly distinct in terms of the appropriateness of discussing them with women.

There is a similar distinction between the two networks regarding the ages of discussion partners. For the parental support network, only a small minority of the discussion partners are younger than the respondents, and the largest proportion of conversational partners is older than the respondent. For the family planning network, conversely, the greatest proportion of discussion partners are among the respondents' agemates, while the remainder of the family planning conversational partners are equally divided among those older and those younger than the respondent. The qualitative evidence suggests that the topic of parental support is viewed by Luo men as more of a normative issue, more appropriately discussed with older men; family planning, on the other hand, is seen as a more instrumental matter, about which discussions with women and men of various ages is acceptable. For the other network attributes, the two networks are virtually indistinguishable.

EVIDENCE OF DIFFERENCES IN NETWORK CHARACTERISTICS IN THE SURVEY DATA

How do the parental support and family planning networks compare in the survey data? Perhaps the most surprising finding is the limited extent to which the conversational partners for the two networks named by individual respondents overlap. Only 418 of 1555 (26.9%) conversational partners named in the family planning network were also named in the parental support network. The two sets of network partners tend to be distinct, despite the limited opportunities for travel to and within the research sites and the clan-based organization of the villages.

Comparisons of specific measures for the two networks are presented in Table 2. The respondents, on average, named more than one additional network partner for the parental support network than for the family planning network. The difference between the two networks is highly statistically discernible. The survey respondents were nearly twice as likely to report having discussed family planning with no one as they were to respond that they discussed parental support with no one (29.8% vs. 15.4%), and they were somewhat more likely to report discussing parental support with six or more network partners than talking about family planning with that number of network partners (29.9% vs. 22.4%). By all measures, parental support is discussed with more partners than is family planning.

Another network variable was created to measure the respondent’s assessment of his emotional closeness to the network partners he named. The “closeness” measure used here is constructed from the responses to the question of how the respondent categorizes his relationship with the network partners he names: Are they "confidants", "just friends", or "acquaintances". The closeness variable was created for each respondent for each of the two networks, by giving a value of 1.0 for each network partner named as a confidant, a value of 0.5 for each partner named as just a friend, and a value of 0.0 for each network partner named as an acquaintance. The scores for all of a respondent’s parental support or family planning network partners were summed and then divided by the number of network partners named. The result is a measure of closeness between the respondent and all of his network partners named, given in a single variable with a value between 0.0 and 1.0.

Table 2
Univariate Distributions of Network Attributes of
All Luo Men, from Survey Data, Nyanza Province, Kenya, 1994-95

	Rural Luo Men Parental Support Network	Rural Luo Men Family Planning Network
Overall network size mean / sd (N)	5.10 / 5.02 (709)	3.83** / 4.81 (709)
0	15.4	29.8
1	2.5	7.2
2	9.3	11.4
3	14	11.3
4	19.2	12.8
5	9.7	5.1
6+	29.9	22.4
Kin network size	2.01 / 1.40 (600)	1.86 / 1.42 (488)
0	19	23.4
1	19.7	20.5
2	22.8	19.9
3	18.3	18.9
4	20.2	17.4
5	n.a.	n.a.
Nonkin network size	1.47 / 1.32 (600)	1.28* / 1.21 (488)
0	32.2	32.8
1	22	30.1
2	21.7	19.3
3	15.2	11.5
4	9	6.4
5	n.a.	n.a.
Proportion kin	0.57 / 0.37 (600)	0.57 / 0.39 (488)
0	19	23.2
0.01-0.33	13.8	10.2

0.34-0.66	17.8	16.2
0.67-0.99	17.2	17.6
1	32.2	32.8
Closeness to partners	0.77 / 0.22(600)	0.73* / 0.27 (488)
0.00-0.34	4.2	9.6
0.35-0.67	32.2	32.4
0.68-1.00	63.7	58
Density	0.62 / 0.36(576)	0.53** / 0.41 (346)
<0.25	19.8	31.8
0.25-0.49	12.9	11.3
0.50-0.74	22.1	18.8
>0.74	45.3	38.2
Sex heterogeneity (IQV)	0.11 / 0.29 (600)	0.24** / 0.39 (488)
0	86.7	72.1
0.01-0.90	8.8	17.2
>0.90	4.5	10.7
Proportion male	0.92 / 0.21 (600)	0.81** / 0.31 (488)
(avg. prop. female)	7.9	19.0
Age	1.28 / 0.59 (600)	1.12** / 0.60 (488)
Avg. Proportion Older (1)	54	43.5
Avg. Proportion Same (2)	21.1	24.8
Avg. Proportion Younger (3)	24.9	31.7

*p<0.02, **p<0.001

There is a modest, yet statistically discernible, difference in the closeness measures for the two networks, with respondents reporting that they feel somewhat closer to their parental support network partners than to their family planning network partners. While the difference is slight, it is statistically significant and in the predicted direction. The respondents are naming parental support network partners to whom they feel somewhat closer than their family planning network partners. This is what we would expect if the parental support network is playing more of a normative function, while the family planning network more of an instrumental role.

Table 3
Univariate Distributions of Network Attributes of
Luo Men Naming 3 or More Contacts, from Survey Data, Nyanza Province, Kenya, 1994-95

	Rural Luo Men Parental Support Network	Rural Luo Men Family Planning Network
Overall network size mean / sd (N)	6.72 / 4.97 (516)	6.83 / 5.05 (366)
3	19.2	21.9
4	26.4	24.9
5	13.4	9.8
6+	41.1	43.4
Kin network size	3.76 / 0.49 (516)	3.67* / 0.59 (366)
0	0.0	0.0
1	0.2	1.4
2	2.1	2.2
3	19.6	24.6
4	78.1	71.9
5	n.a.	n.a.
Nonkin network size	1.56 / 1.36 (516)	1.44 / 1.30 (366)
0	31.0	31.4
1	20.5	25.7
2	20.4	19.1
3	17.6	15.3
4	10.5	8.5
5	n.a.	n.a.
Proportion kin	0.58 / 0.36 (516)	0.60 / 0.36 (366)
0	15.9	16.7
0.01-0.33	16.1	13.6
0.34-0.66	17.1	14.8
0.67-0.99	20.0	23.5
1	31.0	31.4
Closeness to partners	0.76 / 0.21(516)	0.73* / 0.25 (366)
0.00-0.34	3.7	7.7
0.35-0.67	34.3	37.2

0.68-1.00	62.0	55.2
Density	0.62 / 0.34 (514)	0.52** / 0.39 (298)
<0.25	17.9	30.9
0.25-0.49	14.4	13.1
0.50-0.74	24.7	21.8
>0.74	43.0	34.2
Proportion male	0.93 / 0.20 (516)	0.82** / 0.28 (366)
(Avg. Proportion female)	7.4	18.0
Age	1.29 / 0.57 (516)	1.12** / 0.60 (488)
Avg. proportion Older (1)	54.7	43.5
Avg. proportion Agemate (2)	20.5	24.5
Avg. proportion Younger (3)	24.7	32.1

*p<0.02, **p<0.001

The mean density for the parental support network, similarly, is greater than and statistically discernible from the mean density for the family planning network (i.e., 0.62 vs. 0.53). This is additional evidence that the family planning network serves more of an instrumental function than the parental support network. As Mitchell argued, the normative network acts to define and maintain attitudes and norms among network members. This is more easily achieved if network members know one another well (i.e., if network density is higher). The instrumental network, on the other hand, is mobilized in order to achieve an end. In this instance, the family planning network is used by respondents to garner information about the advantages and disadvantages of particular contraceptive methods, their side effects, and where they can be obtained. For the Luo respondents, discussions with people beyond an immediate circle of intimates who all know one another well will be more effective for acquiring new family planning information (Granovetter, 1973, pp.1370-1371). This is reflected, to some degree, in the lower mean density for the family planning network.

Speaking broadly, Luo men are engaging in conversations about changes in the normative system of parental support and their influence on the demand for children with one group of network partners – to some extent, older men – while participating in instrumental discussions about family planning with a somewhat different and more heterogeneous set of network partners. This statistically discernible difference in network densities constitutes modest evidence that the processes of ideational and behavioral change in fertility are occurring among the Luo. This is of particular importance given that there are serious limitations to the number and heterogeneity of potential conversational partners in a partially kin-based village setting with poor transportation.

The statistically discernible differences in the two networks extend to the sex and age characteristics of the network partners. As predicted, the parental support network partners are more likely to be male and older than the family planning network partners. Both differences in the means for these variables are statistically significant. This confirms, indirectly, that rural Luo respondents tend to seek older men -- those who speak with authority on normative matters -- as discussion partners about the erosion of the traditional practice of children supporting their parents in old age. It might also be argued that older men are seeking the younger male respondents as discussion partners on the issue of parental support, in order to remind them of their normative responsibilities. Whatever the case, the family planning network shows much greater heterogeneity in the age and sex compositions of the discussion partners.

The findings reported in Table 2 are for all network partners (up to four) named by respondents. There might be reason to think, however, that there are differences in the network attributes of all respondents and respondents who name three or more partners. Perhaps a more reticent Luo would be more likely to speak only to one older man about parental support or family planning, while a more garrulous Luo might speak to a more heterogeneous set of network partners. Table 3 presents network attributes for respondents naming three or more discussion partners. The findings for these respondents closely parallel those for all respondents, with the two networks having statistically discernible differences in closeness between respondents and their network partners, network densities, and age and sex compositions. Evidence that the two networks can be distinguished along normative and instrumental lines, therefore, is also found in the sample of respondents with greater numbers of network partners.

IMPLICATIONS FOR THE PACE OF FERTILITY DECLINE

Caldwell and Caldwell (1987) assert that the expectation of future support from children is one of the main incentives for sustained high fertility in sub-Saharan Africa (p.422). They argue that if parents are secure in the belief that their children will support them in old age, then having many children is an economically rational strategy for parents to follow. If, however, parents' certainty of future support from their children deteriorates, then the economic rationality of continued high fertility is less clear. The norm of children supporting their parents in old age is beginning to erode among the Luo. Some Luo respondents described how modern economic realities clash with normative responsibilities, preventing them from fulfilling their obligations to their parents. An excerpt from an interview with a 38 year old farmer with seven years of primary education demonstrates this point:

> Interviewer: Do you expect your children to support you in old age?
>
> Respondent:I think they will help me, but again it is difficult with the way the world is today. If I am the one finding it difficult to support my parents, how can I expect help? The prices of things are increasing steadily. [Oyugis 6]

About one-quarter of the semi-structured interview respondents indicated they had serious doubts about their children's ability or willingness to assist them in old age, and nearly every respondent could relate a story of someone who was not able or willing to support his parents. Some men even believe that circumstances have changed so dramatically that parents can now expect to support their children indefinitely, rather than receiving assistance from their children. The following excerpt, from a 24 year old with a secondary education who earns his income from odd jobs and repairing electronic equipment, demonstrates this point:

> Interviewer:If you grow old, do you expect your children to help you?
>
> Respondent: Me, if I grow old the way I'm seeing the world of today, you can educate your child and then later, he has bad luck and doesn't get a job. So you sympathize with

> him and it forces you to help him. You cannot rely on him for future help. So it's better to help him so that he can help himself. [Oyugis 1]

We can speculate that the frequency of normative discussions of parental support indicates that this erosion in the certainty of old age assistance is a contentious issue between older and younger men. Younger men are searching for a normatively acceptable way to balance their responsibilities to their parents and their children. The option of having fewer children offers Luo men a better chance to fulfill all of their obligations. Findings from both the qualitative data and the survey data suggest that the normative discussions essential to a change in desired family size, most likely as a result of the erosion in the norm of parental support, are occurring among the Luo.

The data also demonstrate that Luo men are engaging in the instrumental discussions that will provide the information they will need should they choose to limit their fertility. The frequency of instrumental discussions about contraception and fertility limitation suggests that, as the demand for children declines, either as a result of the growing sense that children are becoming less reliable as sources of old age support or as a result of the increasing cost of raising and educating children, Luo men will have access to the instrumental information they need to limit their family size effectively.

DISCUSSION

At this early point in the fertility decline of the Luo, the family planning networks are as capable of carrying negative information about contraceptive methods as they are of carrying positive information. Many contraceptive adopters, for example, experience unpleasant side effects. These dissatisfied users certainly relate their negative experiences with contraceptive methods to their network partners. An example from one respondent, a 34 year old stonecutter with four years of primary education, demonstrates how negative information about contraceptive methods passes through communication networks:

> Interviewer: You also said there were some people you saw swallowing [oral contraceptives]. How did you know?
>
> Respondent: You can ask. I remember one time I asked another woman why she is constantly sick and she told me her stomach has been affected by "family" tablets. [Gwassi 5]

Some male respondents described conversations they had with dissatisfied users or other discussion partners who had heard about gynecological problems associated with intrauterine devices, birth defects they attribute to the use of hormonal methods, etc. Luo men's descriptions of incredible side effects, particularly those associated with oral contraceptive use, show that a modest amount of inaccurate information continues to flow through these networks. Most of the arguments against family planning voiced by the Luo men in the qualitative interviews centered on side effects and fear of the long term health and fertility implications of contraceptive use, rather than strong normative or religious opposition to fertility limitation. This parallels network research findings for women in voluntary associations in Cameroon (Valente *et al.*, 1997). Such rumors and inaccuracies, as well as stories about genuinely dissatisfied users, circulating through the family planning networks impede the spread of contraceptive use among the Luo.

While the Luo men's family planning networks are undoubtedly carrying some negative information about contraceptive methods, evidence from other sources suggests that the majority of family planning discussions probably convey positive messages about contraceptive methods and fertility limitation, and that Luo men are getting the information they need to make informed decisions about contraceptive adoption. In the following excerpt, the same respondent quoted immediately above describes another family planning conversation he remembered:

> Interviewer: You have said you always see somebody swallow a tablet. Have you talked to any of these people?
>
> Respondent: Yes, I had talked to one woman from our village.
>
> Interviewer: What did she say?
>
> Respondent: She said that she doesn't have any problems when swallowing these tablets. She can take up to three or four years without having a child, and when she decides to have a child, she just does it without any problem. She is my sister-in-law, my wife's sister. [Gwassi 5]

In addition, data from the *Kenya Demographic and Health Survey 1993* show that well over 90% of the Luo men sampled know a modern contraceptive method and a source for obtaining it (Table 11.5, p.145). More than one-in-five of the Luo men (or their partners) surveyed was currently using a modern method of contraception (Ibid., Table11.8, p.148) and slightly more than 90% of these men reported approving of family planning (Ibid., Table 11.15, p.153). According to these data, Luo communication networks are carrying generally positive information about contraception and family planning. Given a decline in the desire for large families, broad approval of family planning, the existence of active family planning networks, and a significant number of current contraceptive users, there are reasons to expect that contraceptive use among the Luo could increase rapidly, with an accompanying fall in fertility rates.

As the value of children to their parents declines, either due to the increasing uncertainty that children will provide support to their parents in old age or due to the increasing expense of raising children, Luo fertility will continue to fall. Limited declines in Luo fertility have already been observed. The *Kenya Demographic and Health Survey 1993* reports, for instance, that between 1984 and 1992 total fertility rates in Nyanza Province fell by 18% (Table 3.4, p.25). Further declines are likely, given the perceived erosion of the normative understanding between parents and children that made high fertility economically rational in the past.

CONCLUSION

Communication networks are essential components of social change; fertility decline is a case in point. Social change requires transformations in both normative understandings and the behaviors that translate those understandings into socially sanctioned outcomes. The Luo men of Kenya's Nyanza Province currently are engaged in discussions of a number of issues regarding fertility change, including the value of children and the perceived breakdown in the norm of children supporting their parents in old age, ideal family size, and the best means for achieving the aim of having fewer children. Their communication networks are central to the processes of normative and behavioral change that will result in substantial fertility decline. While these processes have begun only recently, contraceptive prevalence is increasing and fertility decline has begun among the Luo. The communication networks of the Luo men of Kenya's Nyanza Province are likely to facilitate and accelerate these changes.

This search for a new normative agreement suggests that the confidence parents once had in their children's ability and willingness to care for them in old age -- a major incentive for continued childbearing -- is eroding among the Luo. In its place, uncertainty and a new calculus of the demand for children is emerging. The qualitative evidence indicates that Luo men clearly perceive greater benefits from having fewer children than their fathers had, as a result of both the greater expense of raising children and the reduced support parents can expect from their children. The normative underpinnings for fertility decline, therefore, are being established now, expressed primarily in a reduced demand for children. The strong normative content of the parental support discussions suggests that the process of ideational change, the first step in the broader process of fertility decline, is occurring among the Luo.

Family planning, conversely, proves to be more of an instrumental topic, and when Luo men seek information about limiting family size or contraceptive methods and their side effects, a more heterogeneous set of discussion partners is culturally appropriate. These diverse others include more women, agemates, and people younger than the respondents, as well as more individuals who do not know the respondent's other network partners well. This heterogeneity and differentiation might surprise those who would expect the networks of rural men living in villages composed largely of clan members, and who have limited means of transportation beyond their immediate locations, to be fairly predictable and homogeneous. The existence of diverse, instrumental networks, used for gathering information about the means for limiting fertility, demonstrates that the process of behavioral change, the transformation from ideational change to actual fertility decline, is also under way among the Luo.

Further research on the role that communication networks play in fertility decline and other processes of social change seems warranted. Measures of the attributes of respondents' networks, such as their densities and age and sex compositions, provide important evidence of the way in which discussions influence normative and behavioral aspects of fertility decisions. While the network data provide important evidence about the attributes of the partners the respondents are selecting for these discussions, qualitative data provide useful insights into the content of those discussions. Both longitudinal network data and comparative studies of the network characteristics of other rural African men would add greatly to our understanding of the links between parental support and family planning discussions and the process of fertility decline.

REFERENCES

Barnett, G. A., J. A. Danowski, and W. D. Richards, Jr. (1993). “Communication Networks and Network Analysis: A Current Assessment” in William D. Richards, Jr. And George A. Barnett (eds.), *Progress in Communication Sciences, volume XII.* Norwood, New Jersey: Ablex Publishing Corporation, pp. 1-19.

Caldwell, J. and Caldwell, P. (1987). "The Cultural Context of High Fertility in sub-Saharan Africa." *Population and Development Review 13*:409-437.

Caldwell, J. C. (1976). “Toward a Restatement of Demographic Transition Theory.” *Population and Development Review 2*(3-4):321-366.

Fischer, C. S. (1982). *To Dwell among Friends: Personal Networks in Town and City.* Chicago: The University of Chicago Press.

Francis, E. and J. Hoddinott. (1993). “Migration and Differentiation in Western Kenya: A Tale of Two Sub-Locations.*” Journal of Development Studies 30*(1):115-145.

Granovetter, M. (1973). "The Strength of Weak Ties." *American Journal of Sociology 78*:1360-1380.

Hoddinott, J. (1992). “Rotten Kids or Manipulative Parents: Are Children Old Age Security in Western Kenya?” *Economic Development and Cultural Change 40*(3):545-565.

Kenya Demographic and Health Survey, 1993 (1994). Calverton, Maryland: National Council for Population and Development, Central Bureau of Statistics, Office of the Vice President and Ministry of Planning and National Development and Macro International Inc.

Marsden, P. (1987). "Core Discussion Networks of Americans." *American Sociological Review 52*:122-131.

Mason, K. and Taj, A. (1987). "Differences between Women's and Men's Reproductive Goals in Developing Countries." *Population and Development Review 13*:611-638.

Mitchell, J. C. (1969). “The Concept and Use of Social Networks” in J. Clyde Mitchell (ed.), *Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns*. Manchester: Manchester University Press, pp. 1-50.

Montgomery, M. R. and J. B. Casterline. (1993). “The Diffusion of Fertility Control in Taiwan: Evidence from Pooled Cross-Section Time-Series Models.” *Population Studies* (47):457-479.

Moore, G. (1990). "Structural Determinants of Men's and Women's Personal Networks." *American Sociological Review 55*:726-735.

Parkin, D. (1978). *The Cultural Definition of Political Response: Lineal Destiny Among the Luo*. London: Academic Press.

Shipton, P. (1989). *Bitter Money: Cultural Economy and Some African Meanings of Forbidden Commodities*. Washington, D.C.: American Anthropological Association, American Ethnological Society monograph series, no.1.

Shipton, P. (1992). “Debts and Trespasses: Land, Mortgages, and the Ancestors in Western Kenya.” *Africa 62*(3):357-388.

Shipton, P. (1995). “Luo Entrustment: Foreign Finance and the Soil of the Spirits in Kenya.” *Africa 65*(2):165-196.

Valente, T. W., S. C. Watkins, M. M. Jato, A. van der Straten, and L. P. M. Tsitsol. (1997). “Social Network Associations with Contraceptive Use among Cameroonian Women in Voluntary Associations.” *Social Science and Medicine 45*(5):677-688.

CONNECTIONS 22(2): 50-60

A General Permutation-Based QAP Analysis Approach for Dyadic Data from Multiple Groups[1]

John Levi Martin
Rutgers University

The QAP approach has been extended from the analysis of bivariate correlations to multiple regressions, and is assuming the "workhorse" role for social network analysis analogous to that played by the OLS linear regression for non-network analysis. But there are severe limitations to current implementations, namely the restriction to one group at a time, and to linear regressions. Contemporary computing power makes feasible an implementation that has neither of these restrictions. This paper describes the various alternatives to such an implementation, makes public a user-friendly program for linear and non-linear QAP regressions for data that may come from more than one group, and illustrates its use with an example that sheds some light on the dynamics of influence processes in naturally occurring groups.

THE PROBLEM OF CORRELATED ERROR STRUCTURES AND LIMITATIONS TO EXISTING ROUTINES

As network researchers are painfully aware, the standard errors of model estimates that are normally produced by statistical packages to conduct tests of statistical significance are usually invalid for dyadic data. This is because most derivations of the standard errors require the assumption that all error terms are independent of one another, so that the covariation of any two errors is zero (Hanushek and Jackson 1977: 53). But given dyadic data, each person will contribute to (N-1) dyads, and hence it is quite likely to be the case that the "error" that characterizes one dyad involving ego is similar to the error characterizing another dyad involving ego, or that the errors are "autocorrelated." In this case, the formula giving the variances of the estimates is incorrect, even if the estimates themselves are correct. Therefore, statistical inferences based on network data are likely to be wrong. It is this problem that the Quadratic Assignment Procedure (QAP) attempts to solve.

While the Quadratic Assignment Procedure (Hubert and Schultz 1976; Baker and Hubert 1981; Hubert 1985) is a general--and complex--analytic strategy, it is most commonly implemented for network data in a form analogous to an exact test; we use the observed data to generate a distribution of possible alternative outcomes; we then compute the statistical significance of our observation against this distribution. While certain usages of QAP may in fact be biased (Krackhardt 1992), Krackhardt's (1987, 1988) simulations show that QAP analyses are more likely to discriminate between significant and insignificant correlations than are conventional OLS (Ordinary Least Squares) statistics.

QAP multiple regressions have been incorporated in UCINET (Borgatti, Everett and Freeman 1992), the most popular general network program. But the existing implementations have two limitations. First, one cannot simultaneously fit models to more than one group at a time. (If we were to use a standard QAP package, it would attempt to permute members into groups in which they could not possibly belong, leading to an overwhelming amount of missing data.) Second, models for data with dependent variables that are dichotomous (e.g. logistic or probit regression), and models for data with dependent variables that take on only a few small non-negative integer values (e.g. Poisson or negative binomial regression) cannot be fit. These limitations close off many avenues of investigation to network researchers. There are a number of existing techniques that can be used to handle one or the other of these limitations; in the next section, I briefly review these, and then go on to present a general QAP framework for the analysis of multiple networks, and compare the results obtained via this program to those obtained via other techniques.

OTHER OPTIONS

Multiple Networks

The first limitation discussed was that existing QAP implementations are unable to analyze data from more than one group at a time. But it is possible to do separate analyses for each network, and then summarize the results in a meta-analytic procedure (for example, see Krackhardt and Porter 1986). In addition to doing a meta-analytic test of the significance of dyadic-level covariates, one may conduct a second model at the group level, by regressing the constants on group-level variables (for a two-stage analysis of network data using somewhat similar reasoning, see Blau and Alba 1982). Such a procedure has certain advantages; for example, by separating the groups, the biases that are associated with mixed level models are avoided. However, this procedure has drawbacks. Most importantly, we are unable to constrain the coefficients to be equal across groups; while the meta-analysis may still allow us to pass judgement as to the significance of some coefficient in the sample as a whole, we are not able to use the multiple groups to attain any precision in the estimation of this coefficient.

Correlated Error Structures

The other limitation of the standard QAP implementation has to do with the restriction to linear regression, as opposed to other models specifically designed for dichotomous dependent variables or counts with low means. Since dichotomous data appear frequently in network research (often, relations are either present or absent), and because the problems of using a linear model for dichotomous dependent variables (namely that predicted probabilities can be greater than one or less than zero) are more severe than those that come from fitting a linear model to data that may be Poisson distributed, I focus on the former case for the rest of this section, and discuss existing ways of fitting logistic (or probit) regression models to network data. There are many approaches to correlated error structures for logistic models, but these approaches are in general not applicable. That is because they generally assume that the correlations are across a set units, all of which are hierarchically nested within a some categorical schema (for example, given nonlinear models such as logistic models, see Goldstein 1991, Longford 1994). But there are recent models for network data that in some cases can be used to produce logistic regressions.

The *p** Model

First, there has been an explosion of work extending the Holland-Leinhardt (1981) p_1 model to handle situations of where the dyads are not independent after conditioning on node characteristics. While this approach begins from the "one tie" case, it can be extended to cover the relation between two dyadic variables. For a simple example, given two dichotomous relations **X** and **Y**, (=0,1), a graph statistic $G=\sum x_{ij}y_{ij}$ may be used as the basis for an estimate of a parameterization of the association between the two relations (Wasserman and Pattison 1996: 421). Thus the simplest "positive association" model can easily be incorporated as a form of dyadic dependence of **X** on **Y** restricted to those dyads $x_{ij,}$ y_{kl} in which i=k and j=l. Pattison and Wasserman (forthcoming) extend this to a more complete multivariate framework, and they suggest a wide class of dependence models for such data. Finally, this approach may also be extended to multiple networks (Anderson, Wasserman, and Crouch 1999: 55).

The p^* approach has the advantages of generality and flexibility; however, there are some disadvantages. Many of the disadvantages also apply to most rigorous network studies, such as the absence of a direct retrieval of standard errors of the estimates of parameters, or the assumption (required for pseudo-likelihood estimation) regarding the independence of observations. As this latter assumption can also be stated as an assumption that all dependencies between observations have been taken into account in the model, it is actually an assumption made by other models, including a QAP model. Perhaps of greater practical importance, then, is the exponential complexity that results when a number of relations, especially polychotomous, are taken as predictors of another relation (see Robins, Pattison and Wasserman, forthcoming, for an extension to valued data). While it should be possible in principle to extend the model to include continuous dyadic covariates, this has not yet been done.

The p_2 Model

A different extension of the p_1 framework is found in the p_2 model of van Duijn and Snijders (forthcoming); also see Lazega and van Duijn (1997). In essence, the (fixed but unknown) expansiveness and attractiveness parameters of the p_1 model are made linear functions of a set of individual level covariates and random errors. The p_2 model also parameterizes the density and the reciprocity as a function of dyadic parameters. As a result, the p_2 model can be used to study logistic regressions between dyadic variables. Like the QAP framework (and unlike the p_1 model), there is an explicit introduction of correlated errors associated with each person, leading to correlation of errors across dyads, but the p_2 model differs from a QAP implementation in a number of ways.

For example, given a dependent variable **Y** and two independent variables **X** and **Z**, within the p_2 framework we model the density of **Y** $\mu_{ij}=\mu + X_{ij}\delta_1 + Z_{ij}\delta_2$. (For simplicity of exposition, I ignore the possibility of modeling the reciprocity parameter, and instead, constrain this parameter to be zero in the logarithmic metric.) Covariates on the individual level (such as those employed in the example below) are taken into account separately. Because of the random error structure assumed for the residual attractiveness and expansiveness not explained by covariates, the resulting model is quite close to a QAP logistic model. So while p_2 can allow for the fitting of logistic models with continuous covariates that are difficult to fit within the p^* framework, it cannot fit other models, and its use is restricted to those with GAUSS. It is also not implemented for sets of dyads from more than one network.

A GENERALIZED QAP FRAMEWORK

In sum, there is no convenient way to fit logistic and related models to network data, especially not to multiple networks. As a result, I introduce a generalized QAP framework that allows for the fitting of linear and non-linear models to data from multiple networks. This approach is made actual in a program for Windows NT/95/98, Dyadic Analysis for Multiple Networks. In addition to linear regressions, the program DAMN fits logistic, probit, poisson, and negative-binomial regressions.[2] I first discuss the larger QAP routine, then the model fitting, and then particularities.

The outer shell of the algorithm is a QAP permutation routine, which randomly permutes individuals within their own group. In other words, we reconceive the QAP procedure as reproducing a distribution in which it is impossible that persons could have been in different groups. From this reconstructed distribution, an unbiased estimate of an exact test can be retrieved. This means that any given number of permutations may be more (or less) adequate to characterize this distribution for groups that are smaller (or larger) than others. Otherwise, the basic approach is the same as for a single group. For each permutation, the model in question is fit, and a tally kept of how many times coefficients equal to or larger than those found in the actual data are observed, leading to a p-value.[3]

When we fit data from more than one group, our model is technically a multilevel model. As is well known, the standard errors for group level variables in such models tend to be biased due to autocorrelation problems (Mason et al. 1983, DiPrete and Forristal 1994). But the QAP test for such group level variables is wholly meaningless, since all the permutations occur *within* groups. Thus neither the conventional test nor the QAP test gives us a good measure of such group level variables. The methods for correcting such models (see Bryk and Raudenbush 1992) require the covariance matrix for the parameters, which the permutation test does not produce. As a result, there is no good test of the significance of such group level variables, nor whether the multilevel nature of the model biases other coefficients, most importantly, interactions between dyadic and group level variables. An extremely conservative test would be a fixed effects model, in which we add a dummy variable for each group save one. Such a fixed effects model is automated in DAMN for when one wants to carry out “worst case” comparisons.

The model fit for each permutation may be either a conventional linear regression, or a logistic, probit, Poisson, or negative binomial regression. Each of these requires iterative fitting, and hence leads to a dramatic increase in the time necessary to complete a QAP analysis, in contrast to the linear regression, which has a closed form solution.[4] For large data sets, the time required can be significant (on the order of 10s of minutes, say).[5] Further, with such iterative fitting, different statistical routines can lead to different coefficient estimates. Hence the results obtained from this program may differ from those obtained from other programs. These differences may arise due to differences in standards of tolerance for convergence, differences in the way that near-infinite estimates are treated, and the precise nature of the maximization routine. Such differences are likely to be largest when *N*s are small but coefficients are large, such as when some data are extremely skewed. In such cases, the likelihood surface may be rather flat, in that there is a range of parameter values which produce more or less the same predicted frequencies. In such cases, results differ greatly according to the algorithm used. Experience suggests that the logistic routine is more robust than the probit routine in such circumstances.

The combination of multiple permutations and iterative fitting can lead to a lengthy process when data from a number of different groups are analyzed at one time. In such cases, doing the traditional 3000 or 5000 permutations can bog down a computer for something easily on the order of half an hour. However, it is possible to avoid such a lengthy process in many cases, since it may be that after just 300 permutations we are close to the true p-value. The problem is, of course, that we do not know in advance when 300 permutations are adequate and when we need 3000 or more. DAMN solves this by compartmentalizing its permutation results into separate logical “bins,” and comparing the p-values across bins. When the range is less than some specified amount, it considers its process to have “converged.” Further savings in time can be made by ignoring the coefficients for the constant and the fixed effects discussed above. Finally, since we are often interested in certain critical tests (such as p<.05 as opposed to p>.05), DAMN can restrict its attention to those coefficients which seem to be straddling some critical test value. Other p-values are ignored, which is especially helpful as it is non-significant coefficients which are most likely to change with sampling-via-permutations. As a result, it is sometimes the case that a complex model can be fit without too long a wait.

Other Features of the Program
DAMN is intended as a general platform for dealing with dyadic data from multiple networks, but not for manipulating the data. Recoding and similar transformations must be done in another environment, and the final file saved in a standard tab delimited ASCII file with variable names. However, to avoid the production of large data files, DAMN will automatically flip independent variables around, so that they refer to alter's reports, and not ego's reports. This can be done with the dependent variable as well, to determine reciprocity effects. DAMN not only carries out the analyses discussed above, but can be used to generate network files that can be read by other programs (such as UCINET and KRACKPLOT). In addition, DAMN permits the viewing of the data for each group, both as a matrix and as a graph.

EXAMPLE

For purposes of brief illustration, I take some illustrative data from the Urban Communes Data Set (Zablocki 1980), a set of network data from over 40 naturally occurring communities in the 1970s.[6] I will first use one group of 18 persons, and then a combined set of 40 groups, to examine some of the sources of social influence. Respondents were asked to name those in the group that they thought were "influential"; for any dyad, then, either ego did or did not name alter as influential. Were such attributions of influentiality tied to a *hierarchical* ordering of persons, or was there a tendency towards *mutuality* of such attributions?

Table 1
Results for Commune 57

	Model 1	**Model 2**	**Model 3**
Ego's Status	-0.429	0.35	0.999
	[p=.540]	[p=.324]	[p=.227]
	{rg=.400}	{rg=.100}	{rg=.100}
Alter's Status	3.733***	2.955**	2.921*
	[p<.001]	[p=.032]	[p=.024]
		{rg=.030}	{rg=.017}
Ego Defers to Alter		1.714***	1.713***
		[p<.001]	[p<.001]
Alter Claims Power		1.124**	1.108***
		[p=.002]	[p<.001]
		{rg=.010}	

Reciprocity			-152.848
			[p=1.000]
Constant	-3.568	-4.363	-4.292
Log-Likelihood	-43.388	-39.12	-38.82

If attributions of influentiality are tied to persons' positions in a strictly *hierarchical* order, we would assume that those alters who are of higher status would be more likely to receive attributions of influentiality than those of lower status. To measure status, I use that latent measure that is retrieved by the "symmetric" model for observed responses to a question on the balance of interpersonal power that is discussed in Martin (1998). I note that such a case can be treated within the p_2 model, but not within the p^* (because of the continuous covariates). Our first model has attributions of influentiality as a dependent variable, and ego's and alter's statuses as independent variables. The results are displayed in Table 1, Model 1.

For each coefficient, we have the value in logarithmic terms, the QAP p-value, and, for coefficients that were not taken into account in the convergence decision, the "range" of estimated p-values across bins. Thus the coefficient for ego's status is clearly not significant, and while the estimates of the exact p-value varied widely (over a range of .4), none fell on the other side of the line p=.1. While ego's status is not related to influentiality, then, alter's status most decidedly *is*. The p-value here was actually *zero*—no permutations had values for this coefficient as great as that which was observed. It does seem that there is a strong hierarchical component to influentiality—high status people are more likely to be chosen as influential.

Of course, there is a possible confounding factor, since an $(N-1)^{th}$ part of the measure of alter's status is ego's power relation with alter. (This is because the alter's status measure is based on the set of N-1 dyadic power relationships between alter and the other members of the group.) It might be that it is simply the dyadic relation between ego and alter that is important, and not alter's overall standing within the group. To test this, model 2 adds dummy variables for ego deferring to alter (saying alter has more power than ego), and alter claiming to have more power than ego. Hence we have two closely related measures of the existence of a power relation between ego and alter. As we find out, both of these measures are indeed related to the likelihood that ego will see alter as influential—the dyadic power relation *does* matter. But it does not make the effect of alter's status insignificant, though it is decreased slightly. I note that while the same general pattern of findings is retrieved by the p_2 model (constrained to have no reciprocity effects), the coefficients are not always twice their standard errors.[7]

In other words, in this commune, we find a strong hierarchical component to the organization of influence. But what about the effects of reciprocity? Do these also exist? It is, after all, possible that there are both hierarchical and reciprocal effects. Model 3 enters this effect, which produces a preposterous logged coefficient of –152. This is because there simply are no dyads with mutual attributions of influentiality in this group. In general, this would seem a strong finding. But influentiality is a relatively rare attribution (with 14 attributions in the group as a whole, there is less than one per member on average). Unless there was a strong tendency towards mutuality, we might never see it with simply this one group, since by chance, there are unlikely to be mutual attributions even were there no negative effect. (And the p-value of 1.00 confirms that this large number is more or less meaningless.[8])

But if we assume that the magnitude of the effects is the same across all 40 groups in this sample, by pooling the dyads, we can get an answer to the question of a tendency towards reciprocity in attributions of influentiality. Table 2 contains replications of models 1-3 for this larger set of around 3500 (asymmetric) dyads. Model 1 demonstrates that while the effect for alter's status is more or less the same in all groups together as it is in the one group examined, the net effect of ego's status in the larger sample seems to be mildly, but quite significantly, positive. Model 2 also confirms that while adding measures of the dyadic power relation reduces the magnitude of the coefficient for alter's status, it does not eliminate it. Finally, Model 3 enters the reciprocity effect, and it is significantly positive. While not as large as the effect of alter's status, it is roughly of the magnitude of the effect of the dyadic power relation (ego deferring to alter). Further, because we have controlled for ego and alter's status, we know that this is not actually a "ceiling" effect masking itself as reciprocity. (That is, if alter's status, but not ego's status, led to a high probability of ego attributing influentiality to alter, there would be an apparent reciprocity effect since the high status members would have no one else to choose but each other.)

In sum, by pooling the dyads, we are able to get considerable insight as to the dynamics by which group members perceive others as influential, though we might not be able to determine these by examining one single group.

Table 2
Results for 40 Combined Groups

	Model 1	**Model 2**	**Model 3**
Ego's Status	.455***	.974***	.649*
	[p<.001]	[p<.001]	[p=.010]
			{rg=.010}
Alter's Status	2.444***	1.784***	1.801***
	[p<.001]	[p<.001]	[p<.001]
Ego Defers to Alter		.975***	.938***
		[p<.001]	[p<.001]
Alter Claims Power Reciprocity		.531**	.532**
		[p=.002]	[p=.008]
			.942***
			[p<.001]
Constant	-2.247	-2.61	-2.745
Log-Likelihood	-1154.61	-1116.58	-1097.85

CONCLUSION

Increases in computing power make possible the substitution of processor-hungry routines for mathematically elegant ones. For social network researchers, this means that it is possible to use a permutation-based test for the significance of coefficients for a wide variety of models which otherwise might not be applied. A freely-available and user-friendly program allows the implementation of the most widely used models for single dependent variable regression type models.[9]

REFERENCES

Anderson, C.J., S. Wasserman and B. Crouch. 1999. A P* Primer: Logit Models for Social Networks. *Social Networks*. 21(1): 37-66.

Baker, Frank B. and L.J. Hubert. 1981. The Analysis of Social Interaction Data: A Nonparametric Technique. *Sociological Methods and Research* 9:339-361.

Blau, Judith R. and R.D. Alba. 1982. Empowering Nets of Participation. *Administrative Science Quarterly* 27: 363-379.

Borgatti, Stephen P., M. Everett, and L. Freeman. 1992. *UCINET IV Version 1.0.* Columbia: Analytic Technologies.

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

DiPrete, Thomas A. and J.D. Forristal. 1994. Multilevel Models: Methods and Substance. *Annual Review of Sociology* 20:331-357.

Goldstein, Harvey. 1991. Nonlinear multilevel models, with an Application to Discrete Response Data. *Biometrika* 78:45-51.

Gulati, Ranjay, and M. Gargiulo. 1999. Where Do Interorganizational Networks Come From? *American Journal of Sociology* 104:1439-1493.

Hanushek, Eric A., and J.E. Jackson. 1977. *Statistical Methods for Social Scientists.* New York: Academic Press.

Holland, P.W. and S. Leinhardt 1981. An Exponential Family of Probability Distributions for Directed Graphs. *Journal of the American Statistical Association* 76:33-65.

Hubert, Lawrence. 1985. Combinatorial Data Analysis: Association and Partial Association. *Psychometrika* 50:449-467.

Hubert, Lawrence and J. Schultz. 1976. Quadratic Assignment as a General Data Analysis Strategy. *British Journal of Mathematical and Statistical Psychology* 29:190-241.

Krackhardt, David. 1987. QAP Partialling as a Test of Spuriousness. *Social Networks* 9:171-186.

Krackhardt, David. 1988. Predicting with Networks: Nonparameteric Multiple Regression Analysis of Dyadic Data. *Social Networks* 10:359-381.

Krackhardt, David. 1992. A Caveat on the Use of the Quadratic Assignment Procedure. *Journal of Quantitative Anthropology* 3:279-296.

Krackhardt, David and L.W. Porter. 1986. The Snowball Effect: Turnover Embedded in Communication Networks. *Journal of Applied Psychology* 71:50-55.

Lazega, Emmanuel, and M.A.J. van Duijn. 1997. Position in Formal Structure, Personal Characteristics and Choices of Advisors in a Law Firm: A Logistic Regression Model for Dyadic Network Data. *Social Networks* 19:375-397.

Longford, N. T. 1994. Logistic Regression with Random Coefficients. *Computational Statistics and Data Analysis* 17:1-15.

Martin, John Levi. 1998. Structures of Power in Naturally Occurring Communities. *Social Networks* 20:197-225.

Mason, William M., G.Y. Wong, and B. Entwhistle. 1983. Contextual Analysis through the Multilevel Linear Model. *Sociological Methodology* 13:72-103.

Pattison, Philippa, and S. Wasserman. forthcoming. Logit Models and Logistic Regressions for Social Networks: II. Multivariate Relations. *British Journal of Mathematical and Statistical Psychology.*

Robins, Garry, P. Pattison, and S. Wasserman. 1999. Logit Models and Logistic Regressions for Social Networks: III. Valued Relations. *Psychometrika* 64(3):371-394.

van Duijn, Marijtje A. J. 1999. *p2 Computer Program*. Available from the author.

van Duijn, Marijtje A. J., and T.A.B. Snijders. Forthcoming. p2: a random effects model with Covariates for Directed Graphs.

Wasserman, Stanley, C.J. Anderson, and B. Crouch. 1999. A p* Primer: Logit Models for Social Networks. *Social Networks* 21:37-66.

Wasserman, Stanley, and P. Pattison. 1996. Logit Models and Logistic Regressions for Social Networks: I. An Introduction to Markov Graphs and p*. *Psychometrika* 61:401-425.

Zablocki, Benjamin. 1980. *Alienation and Charisma: A Study of Contemporary American Communes.* New York: The Free Press.

Appendix A

Sample Input File for DAMN

GROUP	EGO	ALTER	INFLUE	STATUS	STATUSA	EGODEFER	BLANK
19	172	173	0	.4633	-.2325	0	.5
19	172	174	0	.4633	.501	0	.5
19	172	175	0	.4633	-.04	0	.5
19	172	176	0	.4633	-.6918	0	.5
19	173	172	1	-.2325	.4633	1	.5
19	173	174	1	-.2325	.501	1	.5
19	173	175	0	-.2325	-.04	1	.5
19	173	176	0	-.2325	-.6918	0	.5
19	174	172	0	.501	.4633	0	.5
19	174	173	0	.501	-.2325	0	.5
19	174	175	0	.501	-.04	0	.5
19	174	176	0	.501	-.6918	0	.5
19	175	172	0	-.04	.4633	0	.5
19	175	173	0	-.04	-.2325	0	.5
19	175	174	0	-.04	.501	0	.5
19	175	176	0	-.04	-.6918	0	.5
19	176	172	0	-.6918	.4633	0	.5
19	176	173	0	-.6918	-.2325	0	.5
19	176	174	0	-.6918	.501	0	.5
19	176	175	0	-.6918	-.04	0	.5
42	196	197	1	.6011	.1988	0	.5
42	196	198	0	.6011	-.5532	0	.5
42	196	199	0	.6011	-.4858	0	.5
42	196	200	0	.6011	.239	0	.5
42	197	196	0	.1988	.6011	1	.5
42	197	198	0	.1988	-.5532	0	.5
42	197	199	0	.1988	-.4858	0	.5
42	197	200	0	.1988	.239	0	.5
42	198	196	1	-.5532	.6011	1	.5
42	198	197	0	-.5532	.1988	0	.5
42	198	199	0	-.5532	-.4858	0	.5
42	198	200	0	-.5532	.239	1	.5
42	199	196	0	-.4858	.6011	1	.5
42	199	197	1	-.4858	.1988	1	.5
42	199	198	1	-.4858	-.5532	0	.5
42	199	200	0	-.4858	.239	0	.5
42	200	196	0	.239	.6011	0	.5
42	200	197	0	.239	.1988	0	.5
42	200	198	0	.239	-.5532	0	.5
42	200	199	1	.239	-.4858	0	.5

Appendix B

Sample Output File for DAMN

```
Logfile for DAMN program.
DATE : 11/10/1999
TIME : 1:33PM
+-----------------+
|      DAMN       |
+-----------------+
Data Analysis for Multiple Networks.
Version BETA.2
(This version fixes permutation for low numbers)


Load data file ex2gp.DAT?
Press to read in   40 cases; otherwise enter a different
number
Reading file ex2gp.DAT


Press to read in   4 variables; otherwise enter a
different number
Reading  4 variables.
DAMN read  40 dyads from   2 groups.
Choose one of the following:
(F) Create a set of group-level files
(A) Analyze using QAP
(N) New Data File
(D) Pass on DOS commands
(X) Exit
> a
Do only some groups?
?
** ANALYSIS MENU **
(R) Linear Regression
(L) Logistic Regression
(P) Probit Regression
(S) Poisson Regression (the S is for Simeon Denis)
(N) Negative Binomial Regression
(C) Change settings
(G) Just change group being analyzed
(X) Exit to main menu.
> L
QAP modeling done by default--type 'N' to override.
?
Select a dependent variable to use.
> influe
Input IVs by name; separate with blanks; ignore
end of line (don't press until all done).
> status statusa
Do fixed effects model?
?
Converged at permutation      50
```

```
*** REGRESSION PARAMETERS ***
STATUS    coefficient=  -0.86009  p-value=0.20000
 range=0.20000
STATUSA   coefficient=   0.63684  p-value=0.38000
 range=0.10000
constant  coefficient=  -1.64139  p-value=0.40000
Final Likelihood:      -17.776853

** ANALYSIS MENU **
(R) Linear Regression
(L) Logistic Regression
(P) Probit Regression
(S) Poisson Regression (the S is for Simeon Denis)
(N) Negative Binomial Regression
(C) Change settings
(G) Just change group being analyzed
(X) Exit to main menu.
> L
QAP modeling done by default--type 'N' to override.
?
Select a dependent variable to use.
Press to re-use the following (or enter new):
influe
>
Press to re-use the following (or enter new):
STATUS STATUSA
> STATUS STATUSA egodefer
Do fixed effects model?
?
Maximum estimate deviation is 0.200000 at permutation
number 50
Maximum estimate deviation is 0.200000 at permutation
number 100
{****etc.}
Maximum estimate deviation is 0.094737 at permutation
number 950
Converged at permutation    1000

*** REGRESSION PARAMETERS ***
STATUS    coefficient=  -0.20313  p-value=0.37500
 range=0.05000
STATUSA   coefficient=  -0.62174  p-value=0.29300
range=0.09500
EGODEFER  coefficient=   2.52461  p-value=0.00100
range=0.00500
constant  coefficient=  -2.34395  p-value=0.14100
Final Likelihood: -15.379511

** ANALYSIS MENU **
(R) Linear Regression
(L) Logistic Regression
(P) Probit Regression
(S) Poisson Regression (the S is for Simeon Denis)
(N) Negative Binomial Regression
(C) Change settings
```

```
(G) Just change group being analyzed
(X) Exit to main menu.
> x
Choose one of the following:
(F) Create a set of group-level files
(A) Analyze using QAP
(N) New Data File
(D) Pass on DOS commands
(X) Exit
> x
Session ended relatively normally
```

CONNECTIONS 22(2): 61-70

Non-Parametric Standard Errors and Tests for Network Statistics [1]

Tom A.B. Snijders
ICS, University of Groningen

Stephen P. Borgatti
Carroll School of Management, Boston College

Two procedures are proposed for calculating standard errors for network statistics. Both are based on resampling of vertices: the first follows the bootstrap approach, the second the jackknife approach. In addition, we demonstrate how to use these estimated standard errors to compare statistics using an approximate t-test and how statistics can also be compared by another bootstrap approach that is not based on approximate normality.

INTRODUCTION

In social network analysis, we are used to calculating descriptive statistics for networks, but not so used to accompanying these statistics with standard errors. Yet the general arguments for the benefits of standard errors do apply to social network analysis: it is useful to have an indication of how precise a given description is, particularly when making comparisons between groups. The problem is that there are no established, widely applicable, ways of calculating standard errors for network statistics. Our objective in this paper is to develop some practical procedures for doing so.

In the general (non-network) case, there are, roughly speaking, two approaches to calculating standard errors. The first is to take some descriptive statistic as the point of departure and find a way to calculate a standard error that requires a minimum of assumptions — the simplest example is the commonly calculated standard error of the mean of a simple random sample, which is based only on the assumption that the sample is simple random. The second approach is to formulate some statistical model for the observations, estimate the parameters of this model, and calculate the standard error of these parameter estimates. This paper presents an elaboration of the first of these two approaches for the case of network data. We assume that a researcher is interested in some descriptive statistic — the density of the network, an index for transitivity, network centralization, or any other network property — and wishes to have a standard error for this descriptive statistic without making implausibly strong assumptions about how the network came about.

Two general-purpose non-parametric methods have been proposed in the statistical literature to construct standard errors for complicated or poorly understood statistics: the jackknife (Tukey,

1958) and the bootstrap (Efron, 1979). Both methods are based on resampling, i.e., constructing many artificial data sets out of the observed data set, and using the variability between these artificial data sets. These methods have been theoretically elaborated and have led to a number of applications. A number of useful reviews exist in the literature, including ones by LePage and Billard (1992) and Shao and Tu (1995). A user-friendly DOS program for bootstrap and jackknife analysis, *BOJA*, is available (Boomsma, 1991).

In the literature, these methods are developed for the usual random samples with rectangular data matrices but not for network data with their typical square data matrices and empty diagonal. This note proposes extensions of these methods to the network situation. It is assumed, at least initially, that a network data set Y with N nodes is available, there is interest in a statistic Z that is calculated from Y, and we would like to have a standard error for Z. All of our examples are based on network density, but it is important to remember that the procedures are very general, and could be used with virtually any network statistic, including centralization measures, number of cliques, fit to a core/periphery model, and so on.

THE JACKKNIFE

The basic idea of the jackknife is that, given a dataset of N sample elements, N artificial datasets are created by deleting each sample element in turn from the observed dataset. The new datasets are quite similar, but the variability among them does give an indication of the variability that may be expected between independent replicates of the data set.[2] The jackknife standard error for statistics Z defined for rectangular data matrices corresponding to simple random samples is defined by

$$\sqrt{\frac{N-1}{N}\sum_{i=1}^{N}(Z_{-i}-Z_{-*})^2} \qquad (1)$$

where Z_{-i} is the statistic obtained for the data set from which case i is deleted, and $Z_{-\cdot}$ is the average of $Z_{-1}, ..., Z_{-N}$. The fact that the sum of squares is multiplied by a factor close to 1, instead of being divided by N-1, reflects the fact that these N artificial data sets are much more similar than would be N independent replicates.

The jackknife principle was studied by Frank and Snijders (1994) for network statistics. They found that the multiplication factor (N-1)/N is not adequate for network statistics. The reason is that this multiplication factor is based on the property, valid for most statistics based on simple random samples, that their variance is inversely proportional to the sample size (or approximately so). This is not so for network statistics. The number of relevant elements of an $N\times N$ adjacency matrix with no reflexive ties is $N(N-1)$, and the variance of Z will more likely be inversely proportional to $N(N-1)$ than to N. Accordingly, Frank and Snijders (1994) proposed for network statistics the jackknife variance estimate defined by

$$s.e._J(Z) = \sqrt{\frac{N-2}{2N}\sum_{i=1}^{N}(Z_{-i}-Z_{-*})^2} \qquad (2)$$

where Z_{-i} is the network statistic obtained for the data set from which vertex i is deleted (so that a network on N-1 vertices remains), and $Z_{-.}$ is again the average of Z_{-1}, ..., Z_{-N}. For the statistics studied by Frank and Snijders (estimates for the number of vertices in an unknown graph, based on a snowball sample), the jackknife estimate of standard error performed quite well.

THE BOOTSTRAP

The basic idea of the bootstrap is that the observed data are treated as a population in itself, and that artificial samples of size N are drawn with replacement from the observed data. Thus, each artificial sample will contain multiple copies of some elements of the observed data, whereas other observed data points will be missing from the artificial sample. For network data, the obvious analogy is to draw a sample with replacement from the vertices.

To specify this more explicitly, suppose that the data consist of a network on N vertices denoted $i = 1, ..., N$, where the tie between vertices i and j is denoted Y_{ij}. (The network could be a graph, a directed graph, or a graph with valued edges.) A large number M of bootstrap samples is to be drawn. Each single bootstrap sample is drawn in the following way. A random sample with replacement is drawn from the vertices, and denoted $i(1)$, ..., $i(N)$. This means that all these $i(k)$ are independent draws from the numbers 1, ..., N. The artificial network Y^* is the network induced by these vertices $i(1)$ to $i(N)$. If vertices k and h in the artificial network correspond to different original vertices $i(k)$ and $i(h)$, this means simply that

$$Y^*_{kh} = Y_{i(h)i(h)} \quad , \quad for\ i(k) \neq i(h) \tag{3}$$

i.e., in the artificial network the tie between vertices k and h is the same as the tie between vertices $i(k)$ and $i(h)$ in the observed network.

It is not obvious what to do for the ties between those artificial vertices that correspond to the same real vertex, at least in networks where reflexive ties are not defined. The idea of resampling vertices is that the procedure is meant to leave the basic network structure intact, and any scheme for filling in ties between artificial vertices corresponding to the same real vertex runs the risk of mixing up this structure. (One source of comfort is that, as the number of vertices grows larger, the expected fraction of such doubtfully determined ties will get closer to 0.) As an expedient solution, we propose a dyad-based bootstrap for these ties. This means that the values for the dyads defined by artificial vertices corresponding to the same real vertex are chosen, with replacement, from the set of all $N(N$-1)/2 dyads. The order of the two elements within the dyad is also determined randomly.

The bootstrap standard error is then determined as follows. The described procedure of generating an artificial network is repeated M times independently, where M is large, e.g., 1,000. For each artificial network drawn in this way, the statistic of interest is calculated. Denote these artificial statistics by $Z^{*(1)}$ to $Z^{*(M)}$. This means that $Z^{*(m)}$ is calculated on the basis of the m'th artificially generated network Y^*. The artificial networks are regarded as networks that might have been observed instead of the actually observed one, so that $Z^{*(1)}$ to $Z^{*(M)}$ is regarded as a synthetic sample from the distribution of Z. Accordingly, the bootstrap standard error is

$$s.e._B(Z) = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}(Z^{*(m)} - Z^{*(\cdot)})^2} \qquad (4)$$

where $Z^{*(\cdot)}$ is the mean of the $Z^{*(m)}$.

The main assumption of this bootstrap standard error is that it makes sense to regard the vertices as interchangeable, since the observed vertices are indeed treated interchangeably in the sampling process. Thus, for a network composed of a class of school children this might be more reasonable than for a network existing within a hierarchically structured organization (given that the hierarchy matters for the observed network).

EVALUATING A NETWORK STATISTIC

In this section we consider the problem of comparing an observed network statistic Z with a theoretical value μ. For example, consider a social relation that serves as a conduit for the transmission of a virus.[3] The greater the density of the network, the faster and more certainly the virus is spread. Based on theoretical considerations, we might postulate the existence of a threshold value of network density below which the infection cannot sustain itself, and above which there is danger of developing an epidemic.[4] For any given network, the question is whether the density of ties falls within the safe range.

A standard approach in this situation is to define a null hypothesis stating that the network density is less than the parameter μ, and reject the hypothesis if the observed statistic is sufficiently larger than the parameter, relative to the standard error of the sampling distribution. Hence, we calculate

$$t = \frac{Z - \mu}{\sigma} \qquad (5)$$

and reject the null if t is larger than 1.65, which is the critical value associated with a maximum Type 1 error of 0.05 in a 1-tailed test.

As an empirical example, consider the network of friendship ties among 67 prison inmates collected by Gagnon in the 1950s, reported by MacRae (1960), and available as part of the UCINET 5 software package (Borgatti, Everett and Freeman, 1999). Let us assume that the theoretical "tipping point" that separates epidemic from extinction occurs at density 3%. The observed density for this network is 0.0412. The standard error, as estimated by the bootstrap method with 1,000 samples, is 0.0060 (it is 0.0036 using the jackknife method). Converting to standard error units, we obtain (0.0412 - 0.03)/0.0060 = 1.87. This value is larger than 1.65 and in fact corresponds to a 1-tailed significance level of 0.03. Therefore, we reject the null hypothesis and provisionally conclude that the population is in danger.

It should be noted that in this approach we have assumed that the shape of the sampling distribution is approximately normal, and therefore use the bootstrap sampling distribution only to estimate the variance of this distribution. An alternative approach that does not make this assumption is to use the bootstrap sampling distribution directly to calculate the probability of obtaining an observed density as large as actually observed given the null hypothesis. Hence we would like to simply count the proportion of bootstrap samples that have a test statistic larger than the observed. However, the bootstrap distribution is centered on (or near) the observed statistic, Z, rather than the theoretical parameter. This is because the bootstrap samples from the data rather than from the null distribution. Therefore, as Noreen (1989) suggests, we need to subtract the mean of the bootstrap sampling distribution ($Z^{*(.)}$) from each value of $Z^{*(m)}$ and then add back the theoretical value. We then count the proportion of $Z^{*(m)}$ values that are larger than the observed Z. In effect, we assume the shape of the bootstrap distribution is correct, but simply mis-centered, and we use the center that corresponds to our null hypothesis.

In our example, the mean bootstrap density was 0.0406. We therefore subtracted 0.0406 from each bootstrap density, and added 0.03. We then counted the number of samples in which $Z^{*(m)}$-0.0406+0.03 was greater than or equal to the observed value of 0.0412. Adding 1 to this count and dividing by M+1 gives an estimate of the proportion of samples (including the observed) which would equal or exceed the observed value – in short, the significance level. In this case, we obtained a significance of 0.038, which agrees well with our previous estimate.

Table 1
Comparison of Approaches for the One Sample Case

	Classical Estimate (σ/√n)	Bootstrap-Assisted SE*	Bootstrap Direct Method*
SE	0.0030	0.0060	NA
T-Statistic	3.73	1.87	NA
1-Tailed Significance	< 0.001	0.031	0.038

* Using 5,000 bootstrap samples

Table 1 compares the significance levels obtained via the two bootstrap methods, and the (inappropriate) classical approach, which estimates the standard error of the sampling distribution from standard deviation of the sample variable. Note that the two bootstrap methods agree closely, while the classical method, whose assumptions are violated by network data, yields very different values.

COMPARING TWO NETWORKS

Another important application area is the comparison of a network statistics for two different groups. For example, Ziegler et al (1985) report the corporate interlocks among the major German business entities (15 in total). Stokman et al (1985) report interlocks among the major Dutch business entities (16 in total). The data are available as part of the UCINET 5 (Borgatti, Everett & Freeman, 1999) software package. The question we pose is this: is the level of interlock (i.e., the density of ties) different in the two countries?

The observed density of the Dutch network was 0.5, while the density of the German network was 0.6381, for an observed difference of 0.1381.To test the significance of this difference, we can construct a bootstrap or jackknife-based t-test. Assuming a null hypothesis of no difference, the standard approach[5] is to calculate

$$t = \frac{Z_1 - Z_2}{\sqrt{SE_1^2 + SE_2^2}} \qquad (6)$$

where SE_1 and SE_2 are standard errors usually estimated from the standard deviation of the measured variable in each sample. In the network case, we substitute the jackknife or bootstrap-derived standard errors as outlined above. Selecting the bootstrap as our method of choice, the standard errors for the Dutch and German networks are 0.0902 and 0.1083 respectively. The t-statistic works out to

$$t = \frac{0.5 - 0.6381}{\sqrt{0.0902^2 + 0.1083^2}} = \frac{-0.1381}{0.140943} = 0.9798$$

which is clearly not significant. Thus, we cannot conclude that the Dutch and German economies have developed different levels of corporate interlock.

PAIRED SAMPLES

Next we consider the case of the same network observed at two points in time. For example, Jean Bartunek[6] collected work relationships at two different times among faculty and staff at a school that included elementary, middle, and high school sub-units. The school had had a history of autonomous action within these levels and inadequate coordination among them. Time 1 was at the beginning of the school year, just before a new staff position was implemented with the explicit purpose of increasing coordination, and therefore work relationships, among the faculty staff. The data at Time 2 were collected at the end of the school year, 9 months after the position was created and a person hired to fill it. One of the questions posed by the school was whether the new person was successful in increasing coordination among the administration and faculty (in network terms, whether work ties were increased). In short, we would want to know whether the density at Time 2 is significantly greater than the density at Time 1.

This situation is different from the last one in that the two samples here are not independent: instead, we have two sets of measurements of the same relation on the same set of players, and the relationship between a pair of players at Time 1 is unlikely to be independent of their relationship at Time 0. Therefore, to conduct a t-test, we must construct a different approach, analogous to the classical paired-sample t-test, for estimating the standard error of the difference.

Using the bootstrap, we propose two approaches to the paired sample case, just as we did with the independent samples case. The first is as follows. Given the set of *N* actors in the observed network(s), a random sample of size *N* is drawn with replacement. For this artificial set of actors, two separate networks, one for each time period, are then constructed using the procedures outlined earlier. The density of each is computed, and the difference between them recorded. This is repeated *M* times, and the S.E. of the difference (SE_d) is computed as in Equation 4, where Z refers to the difference in density. This is then used to calculate a t-statistic as shown in Equation 7.

$$t = \frac{Z_1 - Z_2}{s.e._B} \quad (7)$$

This approach is convenient but as noted before assumes that the shape of the sampling distribution is approximately normal. The second approach dispenses with this assumption by directly counting the proportion of bootstrap samples that yield a difference as extreme as actually observed. Once again, however, the bootstrap distribution is (asymptotically) centered on the observed difference rather than the theoretical expectation (usually, zero). Therefore, we subtract the mean of the bootstrap distribution from each bootstrap difference score, and then count the proportion of mean-centered bootstrap differences which are as large as the difference actually observed.

As an empirical example, we use Kapferer's (1972) tailor shop data.[7] He recorded "sociational" (friendship, emotional) data on 39 members of a tailor shop in Zambia at two points in time. After the first set of observations there was a failed strike attempt. After the second set, there was a successful strike. A theory of collective movements might suggest that strikes cannot be successfully organized unless members of the group are well enough connected to enable a single view to emerge (rather than a multiplicity of views held in disparate corners of the network). The hypothesis one would want to test then is that the density at Time 2 is higher than at Time 1.

Table 2.
Bootstrap-Assisted Paired Sample T-Test

	Time 2	**Time 1**	**Difference**
Density	0.3009	0.2132	0.0877
Bootstrap SE (5000 samples)	0.0306	0.0271	0.0245
T-Statistic			3.5773
Significance			< 0.001

As shown in Table 2, the observed density at Time 2 is clearly higher than at Time 1, and the difference is significant when we run a paired sample t-test using the bootstrap-derived *SE* for the difference. The null hypothesis is rejected, and the research hypothesis is supported.[8]

As an aside, it is interesting to note that, had we assumed independent samples, the t-statistic would have been 2.15, which is much smaller than the 3.58 we obtain with the paired samples formula, though still significant.

Table 3.
Direct Bootstrap Approach to Comparing
Difference in Densities for Two Measurements
On the Same Actors

	Difference in Density
Observed	0.0877
Avg. of bootstrap distribution	0.0841
Prop. of bootstrap samples with mean-adjusted difference as large as observed	0.0004

Table 3 gives the results of using the direct bootstrap approach. As shown in the table, the average density difference in the bootstrap sampling distribution was 0.0841. Subtracting this quantity from each bootstrap density and counting the proportion of samples with mean-centered differences greater than the observed difference of 0.0877, we obtain a significance value of 0.0004, which agrees well with the t-test computation.

DISCUSSION

We have proposed non-parametric standard errors, based on an extension of the bootstrap and jackknife principles to network data. In addition we have examined direct methods of evaluating specific hypotheses that are also based on the resampling principle, but which do not assume normality of the artificial sampling distribution. All of these techniques are computer-intensive but, in principle, easy to apply.

Standard errors and statistical tests are inevitably based on considerations that the data — in our case, the network — "could have been different". These differences could occur because of observation errors, unreliability of measurement, the contingent — or probabilistic — nature of the processes that gave rise to the observed relations, sampling of vertices, choice of the observation moment (i.e., sampling in time), or many other reasons. Even in the study of entire networks, such considerations often are realistic. You cannot have a statistical test without the assumption that the data could have been different; the question is, which differences would have been likely?

The main basis for both the bootstrap and jackknife approaches is the assumption that the vertices are interchangeable. Expressing this more intuitively, it is assumed that for different vertices i and j, the ith row and column of the adjacency matrix are "just as good" as the jth row and column, and that the network in which the ith row and column was replaced by the jth "could also have been observed" — ignoring, for the moment, the (i, j) and (j, i) elements. In still other words, the essential structure of the network would not be changed to an important extent by this replacement. If it is reasonable to take this viewpoint, then the bootstrap standard error and the bootstrap tests are reasonable also.

For the jackknife standard error, there is an additional assumption: the variance of the statistic is assumed to be approximately inversely proportional to $N(N\text{-}1)$, where N is the number of vertices. This assumption is not always well-founded, and it can be checked only if the distribution of the network is known, i.e., in theoretical cases. It seems more realistic to assume that, in general, the variance of network statistics is inversely proportional to something between N and $N(N\text{-}1)$. This makes the bootstrap standard error more reliable than the jackknife standard error, except in those cases where evidence for the reliability of the jackknife standard error has specifically been established.

As such, the proposed methods bear some relation to the permutation technique for testing relations between networks, also known as QAP ("Quadratic Assignment Procedure") correlation, proposed by Hubert and Baker (1978) and elaborated and popularized, e.g., by Krackhardt (1988). In contrast to the bootstrap and jackknife procedures, this permutation technique requires that two (or more) networks with the same vertex set are available, and a null hypothesis about their statistical independence (possibly partialling out other variables) is being tested. This null hypothesis is understood as follows: the identity of the N vertices is immaterial for the relation between the two networks, implying that a data set with permuted vertices could just as well have been observed. Similar to the bootstrap and jackknife procedures, the permutation technique is based on artificial data sets: in the simplest case where conformity of two adjacency matrices is tested, one matrix is left as it is while in the other the rows and columns are permuted correspondingly. The QAP technique provides the distribution and standard errors under the null hypothesis of independence. Thus, it is non-parametric because it does not make assumptions about the probability distribution of the networks considered separately, but it is restricted to the null hypothesis that the two networks are statistically independent.

It should be noted that there exist some simple (and usually unrealistic) null models under which the standard errors of certain statistics have been calculated. For distributions that imply exchangeable vertices (in other words, permutation invariance), it makes sense to compare the standard errors derived under such a distribution to the bootstrap and jackknife standard errors. Examples are the $U \mid L$ distribution, where the graph or digraph is random under the condition of a fixed number of arcs; and the $U \mid M,A,N$ distribution, where the dyad count of a digraph is supposed to be given but for the rest the digraph is random (see Wasserman and Faust, 1994, Chapter 13). For some network statistics, standard errors under such distributions have been calculated. E.g., Holland and Leinhardt (1975) give the standard sampling variance of arbitrary linear combinations of the triad census under the $U \mid M,A,N$ distribution and Snijders (1981) gives the sampling variance of the degree variance under the $U \mid L$ distribution. For the majority of graphs generated under such a distribution with exchangeable vertices, it may be expected that the bootstrap and jackknife standard errors are of the same order of magnitude as the standard errors calculated by such formulae, because all three are appropriate in such cases. However, for graphs that have a low probability under these simple distributions — in other words: for which such a simple distribution does not give a good fit — the formulae are untrustworthy and the bootstrap standard error is more reliable, since it requires only the exchangeability of the vertices and not the large degree of randomness that is inherent to these simple distributions. It may be expected, because of this large degree of randomness assumed by the simple distributions, that the bootstrap standard errors will tend to be larger than the standard errors derived for the simple distributions.

Let us end by mentioning that the basis for these non-parametric standard errors and probabilities is mainly intuitive, and that it would be interesting to see research devoted to their reliability. In the meantime, we suggest that it is reasonable to use the techniques we propose, since (a) there seem to be no alternatives in the general case, and (b) it is better to have a rough impression of the uncertainty or variability associated with observed network statistics than none at all. Therefore we hope that especially the bootstrap standard error will be applied widely by network analysts. [9]

REFERENCES

Anderson, R.M. (1982). *The Population Dynamics of Infectious Diseases.* London: Chapman Hall.

Boomsma, A. (1991). *BOJA. A Program for Bootstrap and Jackknife Analysis*. User's guide. Groningen: ProGAMMA (information at http://www.gamma.rug.nl)

Borgatti, S.B., Everett, M.G., and Freeman, L.C. (1999). *Ucinet 5 for Windows: Software for Social Network Analysis*. Boston: Analytic Technologies.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics*, 7, 1-26

Frank, O. & Snijders, T.A.B. (1994) Estimating hidden populations using snowball sampling. *Journal of Official Statistics*, 10: 53-67.

Holland, P.W., and Leinhardt, S. (1975). The Statistical Analysis of Local Structure in Social Networks. In Heise, D.R. (ed.), *Sociological Methodology, 1976*, 1-45. San Francisco: Jossey-Bass

Hubert, L.J., and Baker, F.B. (1978). Evaluating the Conformity of Sociometric Measurements. *Psychometrika*, 43, 31-41

Kapferer, B. (1972). *Strategy and transaction in an African factory*. Manchester: Manchester University Press.

Krackhardt, D. (1988). Predicting with Networks: Nonparametric Multiple Regression Analysis of Dyadic Data. *Social Networks*, 10, 359 - 381

LePage, F., and Billard, L. (eds.) (1992). *Exploring the Limits of Bootstrap*. New York: Wiley

MacRae, J. (1960). Direct factor analysis of sociometric data. *Sociometry*. 23:360-371.

Noreen, E.W. (1989). *Computer-Intensive Methods for Testing Hypotheses*. New York: Wiley

Shao, J., and Tu, D. (1995). *The Jackknife and Bootstrap*. New York: Springer

Snijders, T.A.B. (1981), The Degree Variance: An Index of Graph Heterogeneity. *Social Networks, 3*, 163-174.

Stevenson, W., & Bartunek, J. (1996). Power, interaction, position, and the generation of cultural agreement in organizations. *Human Relations* 49: 75-104.

Stokman, F., Wasseur, F., and Elsas, D. (1985) The Dutch network: Types of interlocks and network structure. In F. Stokman, R. Ziegler & J. Scott (eds.) *Networks of Corporate Power*. Cambridge: Polity Press.

Tukey, J. (1958). Bias and Confidence in Not Quite Large Samples (abstract). *Annals of Mathematical Statistics*, 29, 614

Wasserman, S., and Faust, K. (1994). *Social Network Analysis. Methods and Applications*. Cambridge etc.: Cambridge University Press.

Ziegler, R., Bender, R., and Biehler H. (1985) Industry and banking in the German corporate network. In F. Stokman, R. Ziegler & J. Scott (eds.) *Networks of Corporate Power*. Cambridge: Polity Press.