

The Social Structure of Competition

RONALD S. BURT

My starting point is this: A player brings capital to the competitive arena and walks away with profit determined by the rate of return where the capital was invested. The market production equation predicts profit: Invested capital, multiplied by the going rate of return, equals the profit to be expected from the investment. Investments create an ability to produce a competitive product. For example, capital is invested to build and operate a factory. Rate of return is an opportunity to profit from the investment.

Rate of return is keyed to the social structure of the competitive arena and is the focus here. Each player has a network of contacts in the arena. Certain players are connected to certain others, trusting certain others, obligated to support certain others, dependent on exchange with certain others. Something about the structure of the player's network and the location of the player's contacts in the social structure of the arena create a competitive advantage in getting higher rates of return on investment. This chapter is about that advantage. It is a description of the way in which social structure renders competition imperfect by creating entrepreneurial opportunities for certain players and not for others.¹

1. OPPORTUNITY AND SOCIAL CAPITAL

A player brings three kinds of capital to the competitive arena. There are more, but three are sufficient here. First, the player has financial capital: cash in hand, reserves in the bank, investments coming due, lines of credit. Second, the player has human capital: natural abilities—charm,

This material is taken from a book entitled *Structural Holes*, which will be published in 1992 by Harvard University Press. Permission to reproduce the material is gratefully acknowledged. My thanks also to Professor Richard Swedburg, who condensed the original chapter to a size suitable for this anthology.

health, intelligence, and looks—combine with skills acquired in formal education and job experience to equal the ability to excel at certain tasks. Third, the player's relationships with other players are social capital: through friends, colleagues, and general contacts the player receives opportunities to use his or her financial and human capital. I refer to opportunities broadly, but I certainly mean to include the obvious examples of job promotions, participation in significant projects, influential access to important decisions, and so on. The social capital of people aggregates into the social capital of organizations. In a firm that provides services—for example, advertising, brokerage, or consulting—there are people valued for their ability to deliver a quality product. Then there are the “rainmakers,” valued for their ability to deliver clients. The former do the work and the latter make it possible for all to profit from the work. The former represent the financial and human capital of the firm; the latter represent its social capital. More generally, property and human assets define the firm's production capabilities. Relations within and beyond the firm are social capital.

1.1 Distinguishing Social Capital

Financial and human capital are distinct in two ways from social capital. First, they are the property of individuals. They are owned in whole or in part by a single individual defined in law as capable of ownership, typically a person or corporation. Second, they concern the investment term in the market production equation. Whether held by a person or the fictive person of a firm, financial and human capital gets invested to create production capabilities. Investments in supplies, facilities, and people serve to build and operate a factory. Investments of money, time, and energy produce a skilled manager. Financial capital is needed for raw materials and production facilities. Human capital is needed to craft the raw materials into a competitive product.

Social capital is different on both counts. First, it is a thing owned jointly by the parties to a relationship. No one player has exclusive ownership rights to social capital. If you or your partner in a relationship withdraws, the connection dissolves with whatever social capital it contained. If a firm treats a cluster of customers poorly and they leave, the social capital represented by the firm-cluster relationship is lost. Second, social capital concerns rate of return in the market production equation. Through relations with colleagues, friends, and clients come the opportunities to transform financial and human capital into profit.

Social capital is the final arbiter of competitive success. The capital invested to bring your organization to the point of producing a superb product is as rewarding as the opportunities to sell the product at a profit. The investment to make you a skilled manager is as valuable as the opportunities and the leadership positions in which you get to apply your managerial skills. The investment to make you a skilled scientist with state-of-

the-art research facilities is as valuable as the opportunities and the projects in which you get to apply those skills and facilities.

More accurately, social capital is as important as competition is imperfect and investment capital is abundant. Under perfect competition, social capital is a constant in the production equation. There is a single rate of return because capital moves freely from low-yield to high-yield investments until rates of return are homogeneous across alternative investments. Where competition is imperfect, capital is less mobile and plays a more complex role in the production equation. There are financial, social, and legal impediments to moving cash between investments. There are impediments to reallocating human capital, both in terms of changing the people to whom you have a commitment and in terms of replacing those people with new. Rate of return depends on the relations in which capital is invested. Social capital is a critical variable. This is all the more true where financial and human capital are abundant—which in essence reduces the investment term in the production equation to an unproblematic constant.

These conditions are generic to the competitive arena, making social capital a factor as routinely critical as financial and human capital. Competition is never perfect. The rules of trade are ambiguous in the aggregate and everywhere negotiable in the particular. The allocation of opportunities is rarely made with respect to a single dimension of abilities needed for a task. Within an acceptable range of needed abilities, there are many people with financial and human capital comparable to your own. Whatever you bring to a production task, there are other people who could do the same job; perhaps not as well in every detail, but probably as well within the tolerances of the people for whom the job is done. Criteria other than financial and human capital are used to narrow the pool down to the individual who gets the opportunity. Those other criteria are social capital. New life is given to the proverb of success being determined less by what you know than by who you know. As a senior colleague once remarked, "Publishing high quality work is important for getting university resources, but friends are essential." Only a select few of equally qualified people get the most rewarding opportunities. Only some of comparably high quality products come to dominate their markets. So the question is how.

1.2 The Who and the How

The competitive arena has a social structure: players trusting certain others, obligated to support certain others, dependent on exchange with certain others, and so on. Against this backdrop, each player has a network of contacts: everyone you now know, everyone you have ever known, and everyone who knows you even though you don't know them. Something about the structure of the player's network and the location of the player's contacts in the social structure of the arena adds up to a competitive advantage in getting higher rates of return on investment.

There are two routes into the social capital question. The first describes a network as your access to people with specific resources, creating a correlation between theirs and yours; the second describes social structure as capital in its own right. The idea for the first approach has circulated as power, prestige, social resources, and—more recently—social capital. Nan Lin and his colleagues provide an exemplar for this line of work, showing how the occupational prestige of a person's job is contingent on the occupational prestige of a personal contact leading to the job (Lin 1982; Lin, Ensel, and Vaughn 1981; Lin and Dumin 1986). Related empirical results appear in Campbell, Marsden, and Hurlbert (1986), De Graaf and Flap (1988), Flap and De Graaf (1989), and Marsden and Hurlbert (1988). Coleman (1988) discusses the transmission of human capital across generations. Flap and Tazelaar (1989) provide a thorough review with special attention to social network analysis.

Empirical questions in this line of work concern the magnitude of association between contact resources and your own resources, and variation in the association across kinds of relationships. Granovetter's (1973) weak-tie metaphor, discussed in detail shortly, is often invoked to distinguish kinds of relationships.

Network analysts will recognize this as an example of social contagion analysis. Network structure doesn't predict attitudes or behaviors directly. It predicts similarity between attitudes and behaviors. The research tradition is tied to the Columbia Sociology survey studies of social influence conducted during the 1940s and 1950s. In one of the first well-known studies, for example, Lazarsfeld, Berelson, and Gaudet (1944) show how a person's vote is associated with the party affiliations of friends. Persons claiming to have voted for the presidential candidate of a specific political party tend to have friends affiliated with that party. Social capital theory developed from this line of work describes the manner in which resources available to any one person in a population are contingent on the resources available to individuals socially proximate to the person.

Empirical evidence is readily available. People develop relations with people like themselves (for example, Fischer 1982; Marsden 1987; Burt 1990). Wealthy people develop ties with other wealthy people. Educated people develop ties with one another. Young people develop ties with one another. There are reasons for this. Socially similar people, even in the pursuit of independent interests, spend time in the same places. Relationships emerge. Socially similar people have more shared interests. Relationships are maintained. Further, we are sufficiently egocentric to find people with similar tastes attractive. Whatever the etiology for strong relations between socially similar people, it is to be expected that the resources and opinions of any one individual will be correlated with the resources and opinions of their close contacts.

A second line of work describes social structure as capital in its own right. Where the first line describes the network as a conduit, the second line describes how networks are themselves a form of social capital. This line of work is much less developed than the first. Indeed, it is little

developed beyond intuitions in empirical research on social capital. Network range, indicated by size, is the primary measure. For example, Boxman, De Graaf, and Flap (1991) show that people with larger contact networks obtain higher paying positions than people with small networks. A similar finding in social support research shows that persons with larger networks tend to live longer (Berkman and Syme 1979).

Both lines of work are essential to a general definition of social capital. Social capital is at once the structure of contacts in a network and resources they each hold. The first term describes how you reach. The second describes who you reach.

For two reasons, however, I ignore the question of "who" to concentrate on "how." The first is generality. The question of "who" elicits a more idiographic class of answers. Predicting rate of return depends on knowing the resources of a player's contacts. There will be interesting empirical variation from one kind of activity to another, say job searches versus mobilizing support for a charity, but the empirical generalization is obvious: Doing business with wealthy clients, however wealth is defined, has a higher margin than doing business with poor clients. I want to identify parameters of social capital that generalize beyond the specific individuals connected by a relationship.

The second reason is correlation. The two components in social capital should be so strongly correlated that I could reconstruct much of the phenomenon from whichever component more easily yields a general explanation. To the extent that people play an active role in shaping their relationships, then a player who knows how to structure a network to provide high opportunity knows who to include in the network. Even if networks are passively inherited, the manner in which a player is connected within social structure says much about contact resources. *I will show that players with well-structured networks obtain higher rates of return.* Resources accumulate in their hands. People develop relations with people like themselves. Therefore how a player is connected in the social structure indicates the volume of resources held by the player and the volume to which the player is connected.

The nub of the matter is to describe network benefits in competition so as to be able to describe how certain structures enhance those benefits. The benefits are of two kinds, information and control. First I'll describe information benefits because they are more familiar; then I'll examine control benefits, showing how both kinds of benefits are enhanced by the same element of social structure.

2. INFORMATION

Opportunities spring up everywhere; new institutions and projects that need leadership, new funding initiatives looking for proposals, new jobs for which you know of a good candidate, valuable items entering the market for which you know interested buyers. The information benefits

of a network define who knows about these opportunities, when they know, and who gets to participate in them. Players with a network optimally structured to provide these benefits enjoy higher rates of return to their investments because such players know about, and have a hand in, more rewarding opportunities.

2.1 Access, Timing, and Referrals

Information benefits occur in three forms: access, timing, and referrals. Access refers to receiving a valuable piece of information and knowing who can use it. Information doesn't spread evenly through the competitive arena. It isn't that players are secretive, although that too can be an issue. The issue is that players are unevenly connected with one another, are attentive to the information pertinent to themselves and their friends, and are all overwhelmed by the flow of information. There are limits to the volume of information you can use intelligently. You can keep up with only so many books, articles, memos, and news services. Given a limit to the volume of information that anyone can process, the network becomes an important screening device. It is an army of people processing information who can call your attention to key bits—keeping you up to date on developing opportunities, warning you of impending disasters. This second-hand information is often fuzzy or inaccurate, but it serves to signal something to be looked into more carefully.

Related to knowing about an opportunity is knowing who to bring into it. Given a limit to the financing and skills that we possess individually, most complex projects will require coordination with other people as staff, colleagues, or clients. The manager asks, "Who do I know with the skills to do a good job with that part of the project?" The capitalist asks, "Who do I know who would be interested in acquiring this product or a piece of the project?" The department head asks, "Who are the key players needed to strengthen the department's position?" Add to each of these the more common question, "Who do I know who is most likely to know the kind of person I need?"

Timing is a significant feature of the information received by the network. Beyond making sure that you are informed, personal contacts can make you one of the people informed early. It is one thing to find out that the stock market is crashing today. It is another to discover that the price of your stocks will plummet tomorrow. It is one thing to learn the names of the two people referred to the board for the new vice-presidency. It is another to discover that the job will be created and your credentials could make you a serious candidate for the position. Personal contacts get significant information to you before the average person receives it. That early warning is an opportunity to act on the information yourself or invest it back into the network by passing it on to a friend who could benefit from it.

These benefits involve information flowing from contacts. There are also benefits in the opposite flow. The network that filters information

coming to you also directs, concentrates, and legitimates information about you going to others.

In part, this does no more than alleviate a logistics problem. You can be in only a limited number of places within a limited amount of time. Personal contacts get your name mentioned at the right time in the right place so opportunities are presented to you. Their referrals are a positive force for future opportunities. They are the motor expanding the third category of people in your network, the players whom you don't know but who are aware of you. I'm thinking of that remark so often heard in recruitment deliberations: "I don't know her personally, but several people whose opinion I trust have spoken well of her."

Beyond logistics, there is an issue of legitimacy. Even if you know about an opportunity and could present a solid case for why you should get it, you are a suspect source of information. The same information has more legitimacy when it comes from someone inside the decision-making process who can speak to your virtues. Speaking about my own line of work, which I expect in this regard is typical, candidates offered the university positions with the greatest opportunity are people who have a strong personal advocate in the decision-making process, a person in touch with the candidate to ensure that all favorable information, and responses to any negative information, gets distributed during the decision.

2.2 Benefit-Rich Networks

A player with a network rich in information benefits has (1) contacts established in the places where useful bits of information are likely to air, (2) providing a reliable flow of information to and from those places.

The second criterion is as ambiguous as it is critical. It is a matter of trust, of confidence in the information passed and the care with which contacts look out for your interests. Trust is critical precisely because competition is imperfect. The question is not whether to trust, but who to trust. In a perfectly competitive arena, you can trust the system to provide a fair return on your investments. In the imperfectly competitive arena, you have only your personal contacts. The matter comes down to a question of interpersonal debt. If I do for her, will she for me? There is no general answer. The answer lies in the match between specific people. If a contact feels that he is somehow better than you—a sexist male dealing with a woman, a racist white dealing with a black, an old-money matron dealing with an upwardly mobile ethnic—your investment in the relationship will be taken as your proper obeisance to a superior. No debt is incurred. We use whatever cues can be found for a continuing evaluation of the trust in a relation, but really don't know until the trusted person helps when you need it. With this kind of uncertainty, players are cautious about extending themselves for people whose reputation for honoring interpersonal debt is unknown. The importance of this point is

illustrated by the political boundary around senior management for outsider managers trying to break through the boundary (Burt 1992, Ch. 4).

We do know from social science research that strong relations and mutual relations tend to develop between people with similar social attributes such as education, income, occupation, and age (for example, Fischer 1982; Burt 1986, 1990; Marsden 1987; see also note 2). Whether egocentrism, cues from presumed shared background and interests, or confidence in mutual acquaintances to enforce interpersonal debt, the operational guide to the formation of close, trusting relationships seems to be that a person more like me is less likely to betray me. For the purposes here, I put the whole issue to one side as person-specific and presumed resolved by the able player.

That leaves the first criterion, establishing contacts where useful bits of information are likely to air. Everything else constant, a large, diverse network is the best guarantee of having a contact present where useful information is aired.

Size is the more familiar criterion. Bigger is better. Acting on this understanding, people can expand their networks by adding more and more contacts. They make more cold calls, affiliate with more clubs, attend more social functions. Numerous books and self-help groups can assist you in "networking" your way to success by putting you in contact with a large number of potentially useful, or helpful, or like-minded people. The process is illustrated by the networks at the top of Figure 2-1. The four-contact network at the left expands to sixteen contacts at the right. Relations are developed with a friend of each contact in network A, doubling the contacts to eight in network B. Snowballing through friends of friends, there are sixteen contacts in network C, and so on.

Size is a mixed blessing. More contacts can mean more exposure to valuable information, more likely early exposure, and more referrals. But increasing network size without considering diversity can cripple the network in significant ways. What matters is the number of nonredundant

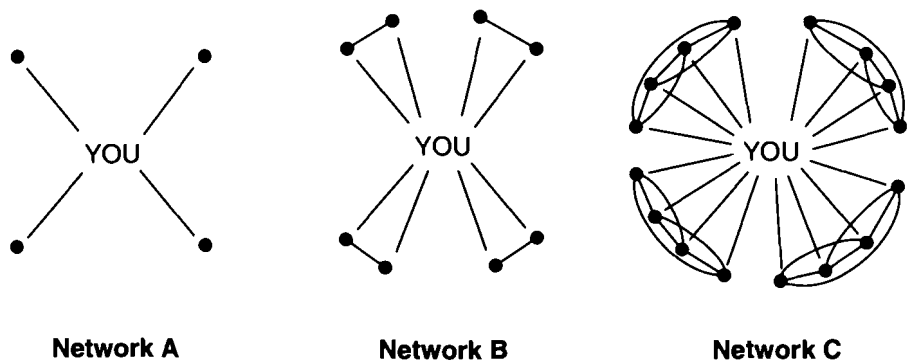


Figure 2-1
Network Expansion

contacts. Contacts are redundant to the extent that they lead to the same people and so provide the same information benefits.

Consider two four-contact networks, one sparse and the other dense. There are no relations between the contacts in the sparse network, and strong relations between every contact in the dense network. Both networks cost whatever time and energy is required to maintain four relationships. The sparse network provides four nonredundant contacts, one for each relationship. No one of the contacts gets you to the same people reached by the other contacts. In the dense network, each relationship puts you in contact with the same people you reach through the other relationships. The dense network contains only one nonredundant contact. Any three are redundant with the fourth.

The sparse network provides more information benefits. It reaches information in four separate areas of social activity. The dense network is a virtually worthless monitoring device because the strong relations between people in the network means that each person knows what the other people know, so they'll discover the same opportunities at the same time.

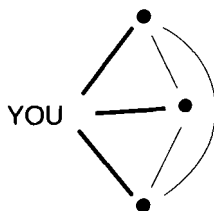
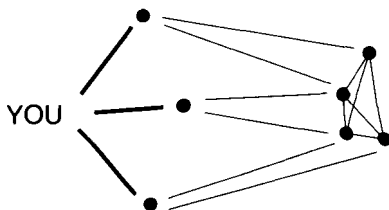
The issue is opportunity costs. At minimum, the dense network is inefficient in the sense that it returns less diverse information for the same cost as the sparse network. A solution is to put more time and energy into adding nonredundant contacts to the dense network. But time and energy are limited, which means that inefficiency translates into opportunity costs. Taking four relationships has an illustrative limit on the number of strong relations that a player can maintain, the player in the dense network is cut off from three-fourths of the information provided by the sparse network.

3. STRUCTURAL HOLES

It will be convenient to have a term for the separation between nonredundant contacts. I use the term "structural hole." Nonredundant contacts are connected by a structural hole. A structural hole is a relationship of nonredundancy between two contacts. The hole is a buffer, like an insulator in an electric circuit. As a result of the hole between them, the two contacts provide network benefits that are in some degree additive rather than overlapping.

3.1 Empirical Indicators

Nonredundant contacts are disconnected in some way—either directly in the sense of no direct contact with one another, or indirectly in the sense of one having contacts that exclude the others. The respective empirical conditions that indicate a structural hole are cohesion and structural equivalence. Both conditions define holes by indicating where they are absent.

**Redundancy
by Cohesion****Redundancy
by Structural
Equivalence****Figure 2-2**

Structural Indicators of Redundancy

Under the cohesion criterion, two contacts are redundant to the extent that they are connected by a strong relationship. A strong relationship indicates the absence of a structural hole. Examples would be father and son, brother and sister, husband and wife, close friends, people who have been partners for a long time, people who frequently get together for social occasions, and so on. You have easy access to both people if either is a contact. Redundancy by cohesion is illustrated at the top of Figure 2-2. The three contacts are connected to one another, and so provide the same network benefits. The presumption here—routine in network analysis since Festinger, Schachter, and Back's (1950) analysis of information flowing through personal relations and Homans' (1950) theory of social groups—is that the likelihood of information moving from one person to another is proportional to the strength of their relationship. Empirically, strength has two independent dimensions: frequent contact and emotional closeness (see Marsden and Hurlbert 1988; Burt 1990).

Structural equivalence is a useful second indicator for detecting structural holes. Two people are structurally equivalent to the extent that they have the same contacts. Regardless of the relation between structurally equivalent people, they lead to the same sources of information and so are redundant. Where cohesion concerns direct connection, structural equivalence concerns indirect connection by mutual contact. Redundancy by structural equivalence is illustrated at the bottom of Figure 2-2. The three contacts have no direct ties with one another. They are nonredundant by cohesion. But each leads you to the same cluster of more distant players. The information that comes to them, and the people to whom

they send information, are redundant. Both networks in Figure 2-2 provide one nonredundant contact at a cost of maintaining three.

The indicators are neither absolute nor independent. Relations deemed strong are only strong relative to others. They are our strongest relations. Structural equivalence rarely reaches the extreme of complete equivalence. People are more or less structurally equivalent. Also, the criteria are correlated. People who spend a lot of time with the same other people often get to know one another. The mutual contacts responsible for structural equivalence set a stage for the direct connection of cohesion. The empirical conditions between two players will be a messy combination of cohesion and structural equivalence, present to varying degrees, at varying levels of correlation.

Cohesion is the more certain indicator. If two people are connected with the same people in a player's network (making them redundant by structural equivalence), they can still be connected with different people beyond the network (making them nonredundant). But if they meet frequently and feel close to one another, then they are likely to communicate and probably have contacts in common. More generally, and especially for fieldwork informed by attention to network benefits, the general guide is the definition of a structural hole. There is a structural hole between two people who provide nonredundant network benefits. Taking the cohesion and structural equivalence conditions together, redundancy is most likely between structurally equivalent people connected by a strong relationship. Redundancy is unlikely, indicating a structural hole, between total strangers in distant groups. After control benefits have been introduced, I'll return to this issue to discuss the depth of a hole.

3.2 The Efficient-Effective Network

Balancing network size and diversity is a question of optimizing structural holes. The number of structural holes can be expected to increase with network size, but the holes are the key to information benefits. The optimized network has two design principles: efficiency and effectiveness.

3.2.1 Efficiency The first principle concerns efficiency, and it says that you should maximize the number of nonredundant contacts in the network to maximize the yield in structural holes per contact. Given two networks of equal size, the one with more nonredundant contacts provides more benefits. There is little gain from a new contact redundant with existing contacts. Time and energy would be better spent cultivating a new contact to unreached people.² Maximizing the nonredundancy of contacts maximizes the structural holes obtained per contact.³

Efficiency is illustrated by the networks in Figure 2-3. These reach the same people reached by the networks in Figure 2-1, but in a different way. What expands in Figure 2-1 is not the benefits, but the cost of maintaining the network. Network A provides four nonredundant contacts.

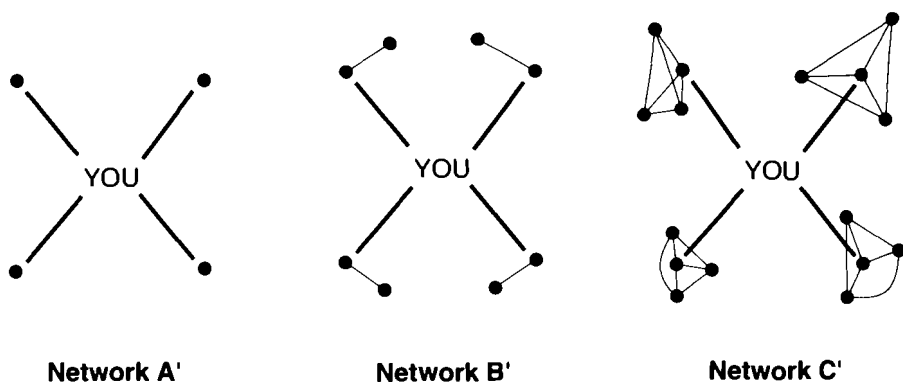


Figure 2-3

Strategic Network Expansion

Network *B* provides the same number. The information benefits provided by the initial four contacts are redundant with benefits provided by their close friends. All that has changed is the doubled number of relationships maintained in the network. The situation deteriorates even further with the sixteen contacts in network *C*. There are still only four nonredundant contacts in the network, but their benefits are now obtained at a cost of maintaining sixteen relationships.

With a little network surgery, the sixteen contacts can be maintained at one-fourth of the cost. As illustrated in Figure 2-3, select one contact in each cluster to be a primary link to the cluster. Concentrate on maintaining the primary contact, and allow direct relationships with others in the cluster to weaken into indirect relations through the primary contact. These players reached indirectly are secondary contacts. Among the redundant contacts in a cluster, the primary contact should be the one most easily maintained and most likely to honor an interpersonal debt to you in particular. The secondary contacts are less easily maintained or less likely to work for you (even if they might work well for someone else). The critical decision obviously lies in selecting the right person to be a primary contact. That is the subject of trust discussed earlier. With a good primary contact, there is little loss in information benefits from the cluster and a gain in the reduced effort needed to maintain the cluster in the network.

Repeating this operation for each cluster in the network recovers effort that would otherwise be spent maintaining redundant contacts. By reinvesting that saved time and effort in developing primary contacts to new clusters, the network expands to include an exponentially larger number of contacts while expanding contact diversity. The sixteen contacts in network *C* of Figure 2-1, for example, are maintained at a cost of four primary contacts in network *C'* of Figure 2-3. Some portion of the time spent maintaining the redundant other twelve contacts can be reallocated to expanding the network to include new clusters.

3.2.2 Effectiveness The second principle for the optimized network requires a further shift in perspective. Distinguish primary from secondary contacts and focus resources on preserving the primary contacts. Here contacts are not people on the other end of your relations; they are ports of access to clusters of people beyond. Guided by the first principle, these ports should be nonredundant so as to reach separate, and therefore more diverse, social worlds of network benefits. Instead of the player maintaining relations with all contacts, the task of maintaining the total network is delegated to primary contacts. The player at the center of the network is then free to focus on properly supporting relations with primary contacts and expanding the network to include new clusters. Where the first principle concerns the average number of people reached with a primary contact, the second concerns the total number of people reached with all primary contacts. The first principle concerns the yield per primary contact. The second concerns the total yield of the network. More concretely, the first principle moves from the networks in Figure 2-1 to the corresponding networks in Figure 2-3. The second principle moves from left to right in Figure 2-3. The target is network C' in Figure 2-3; a network of few primary contacts, each a port of access to a cluster of many secondary contacts.

Figure 2-4 illustrates some complexities in unpacking a network to maximize structural holes. The BEFORE network contains five primary contacts and reaches a total of fifteen people. However, there are only two clusters of nonredundant contacts in the network. Contacts 2 and 3 are redundant in the sense of being connected with each other and reaching the same people (cohesion and structural equivalence criteria). The same is true of contacts 4 and 5. Contact 1 is not connected directly to contact 2 but reaches the same secondary contacts, so contacts 1 and 2 provide redundant network benefits (structural equivalence criterion). Illustrating the other extreme, contacts 3 and 5 are connected directly, but they are nonredundant because they reach separate clusters of secondary contacts (structural equivalence criterion). In the AFTER network, contact 2 is used to reach the first cluster in the BEFORE network, and contact 4 is used to reach the second cluster. The time and energy saved by withdrawing from relations with the other three primary contacts are reallocated to primary contacts in new clusters. The BEFORE and AFTER networks are both maintained at a cost of five primary relationships, but the AFTER network is dramatically richer in structural holes, and so network benefits.

Network benefits are enhanced in several ways. There is a higher volume of benefits because more contacts are included in the network. Beyond volume, diversity enhances the quality of benefits. Nonredundant contacts ensure exposure to diverse sources of information. Each cluster of contacts is an independent source of information. One cluster, no matter how numerous its members, is one source of information because people connected to one another tend to know about the same things at about the same time. The information screen provided by multiple clus-

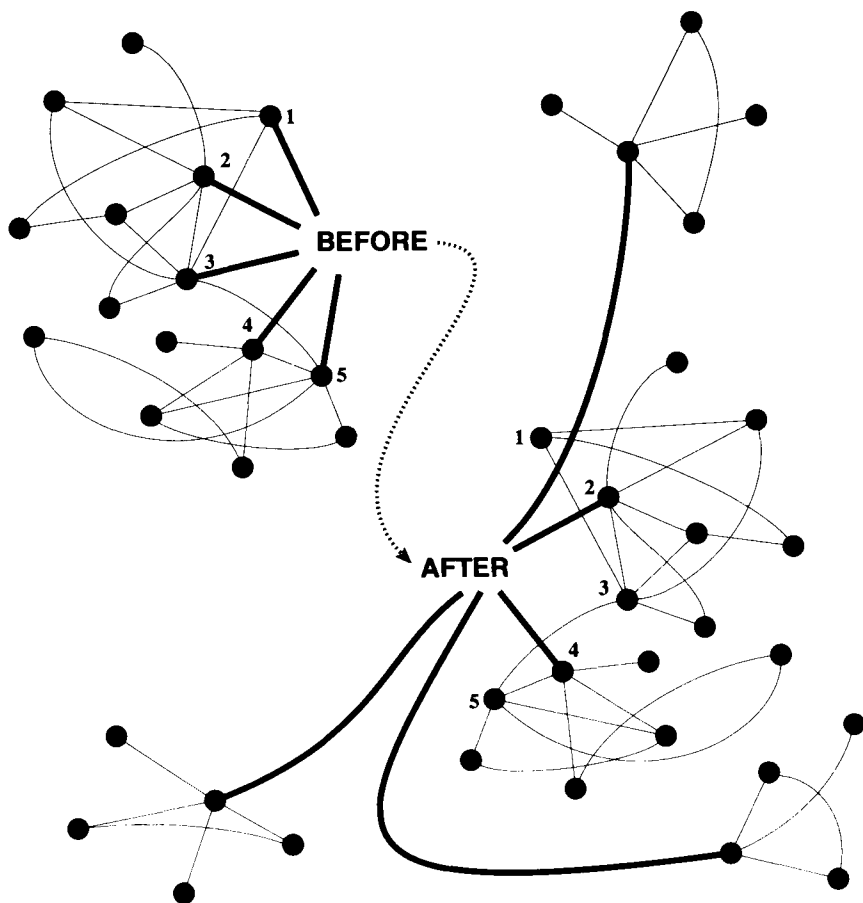


Figure 2-4
Optimizing for Structural Holes

ters of contacts is broader, providing better assurance of the player being informed of opportunities and impending disasters. Further, since non-redundant contacts are linked only through the central player, you are assured of being the first to see new opportunities created by needs in one group that could be served by skills in another group. You become the person who first brings together people, giving you the opportunity to coordinate their activities. These benefits are compounded by the fact that having a network that yields such benefits makes you even more attractive as a network contact to other people, easing the task of expanding the network to best serve your interests.

3.2.3 Growth Patterns A more general sense of efficiency and effectiveness is illustrated with network growth. In Figure 2-5, the number of contacts in a player's network increases from left to right on the horizontal axis. The number who are nonredundant increases up the vertical axis.

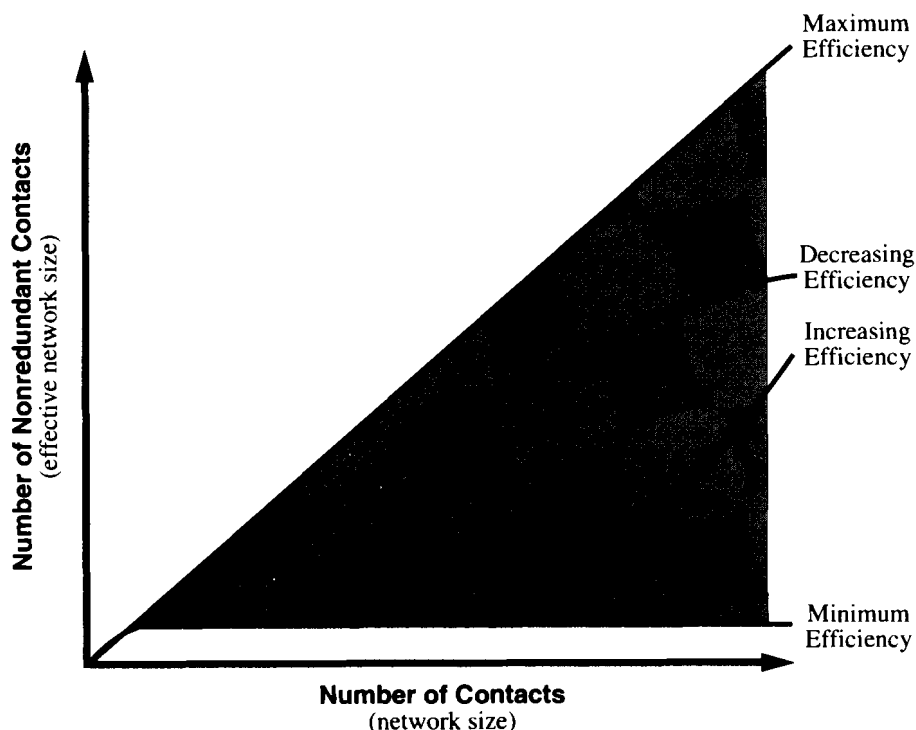


Figure 2-5
Efficiency and Effectiveness

Observed network size increases on the horizontal; effective size increases up the vertical. Networks can be anywhere in the gray area. The maximum efficiency line describes networks in which each new contact is completely nonredundant with other contacts. Effective size equals actual size. Efficient-effective networks are in the upper right portion of the graph. The minimum efficiency line describes networks in which each new contact is completely redundant with other contacts; effective size equals one regardless of multiple contacts in the network.

The two lines between the extremes illustrate more probable growth patterns. The decreasing-efficiency line shows players building good information benefits into their initial network, then relaxing to allow increasing redundancy as the network gets large. Friends of friends begin to be included. Comparisons across networks of different sizes suggest that this is the growth pattern among senior managers (Burt 1992, Ch. 4).

The increasing efficiency line illustrates a different growth pattern. Initial contacts are redundant with one another. A foundation is established with multiple contacts in the same cluster. After the foundation is established, the player's network expands to include contacts in other clusters and effective size begins to increase. There are two kinds of clusters in which optimizing for saturation is wiser than optimizing for

efficiency. The first is obvious. Leisure and domestic clusters are a congenial environment of low-maintenance, redundant contacts. Efficiency mixes poorly with friendship. Judging friends for efficiency is an interpersonal flatulence from which friends will flee. The second exception is a cluster of contacts where resources are dense. For the CEO, the board of directors would be such a cluster. The university provost is similarly tied to the board of trustees. For the more typical manager, the immediate work group is such a cluster, especially with respect to funding authority within the group. These clusters are so important to the vitality of the rest of the network that it is worth treating each person in them as a primary contact regardless of redundancy. Saturation minimizes the risk of losing effective contact with the cluster and minimizes the risk of missing an important opportunity anywhere in the cluster.

The general point is that the probability of receiving network benefits from a cluster has two components: the probability that a contact will transmit information to you and the probability that it will be transmitted to the contact. I count on dense ties within a cluster to set the second probability to one. The probability of having a benefit transmitted to you therefore depends only on the strength of your relationship with a contact in the cluster. However, where the density of ties in an opportunity-rich cluster lowers the probability of your contact knowing about an opportunity, there is value in increasing the number—and so the redundancy—of contacts in the cluster so that total coverage of the cluster compensates for imperfect transmission within it.

3.3 Structural Holes and Weak Ties

In 1973, Mark Granovetter published his now famous article “The Strength of Weak Ties.” The weak-tie argument is elegantly simple. The stage is set with results familiar from the social psychology of Festinger and Homans circa 1950, the results I discussed in section 3.1 with respect to cohesion indicators of structural holes. People live in a cluster of others with whom they have strong relations. Information circulates at a high velocity within these clusters. Each person tends to know what the other people know. Therefore, and this is the insight of the argument, the spread of information on new ideas and opportunities must come through the weak ties that connect people in separate clusters. The weak ties so often ignored by social scientists are in fact a critical element of social structure. Hence the strength of weak ties. Weak ties are essential to the flow of information that integrates otherwise disconnected social clusters into a broader society.

The idea and its connection with structural holes is illustrated in Figure 2-6. There are three clusters of players. Strong ties, indicated by solid lines, connect players within clusters. Dashed lines indicate two weak ties between players in separate clusters. You, as one of the players, have a unique pattern of four ties: two strong ties within your cluster and a weak tie to a contact in each in the other clusters. There are three classes

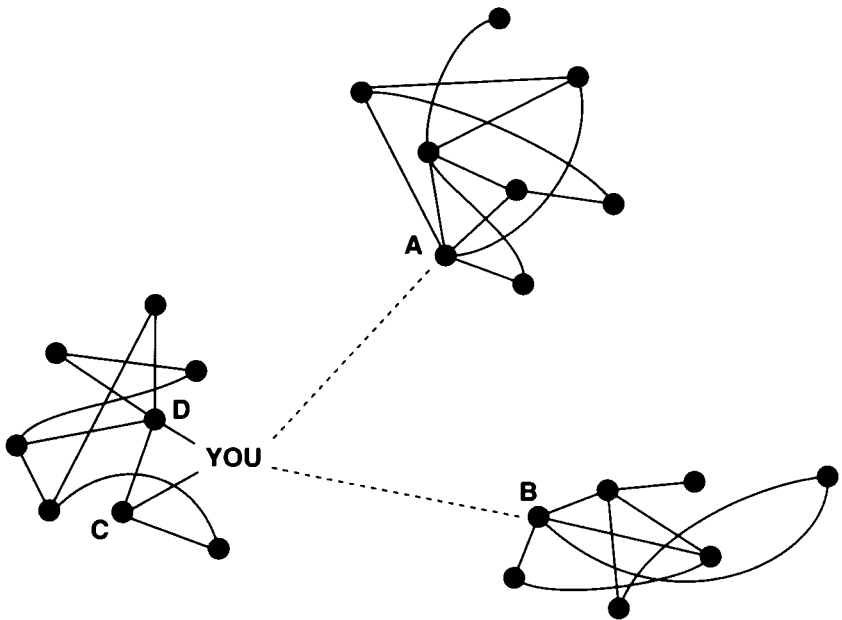


Figure 2-6
Structural Holes and Weak Ties

of structural holes in your network: (1) holes between the cluster around contact *A* and everyone in your own cluster—for example, the hole between contacts *A* and *C*; (2) holes between the cluster around contact *B* and everyone in your own cluster—for example, the hole between contacts *B* and *C*; and (3) the hole between contacts *A* and *B*.

Weak ties and structural holes seem to describe the same phenomenon. In Figure 2-6, for example, they predict the same ranking of information benefits. You are best positioned for information benefits, contacts *A* and *B* are next, followed by everyone else. You have two weak ties, contacts *A* and *B* have one each, and everyone else has none. You have the largest volume of structural holes between your contacts, contacts *A* and *B* have fewer, and everyone else has few or none.

The weak-tie argument is simpler than my argument and already well known. Why complicate the situation with the structural-hole argument?

There are two reasons. First, the causal agent in the phenomenon is not the weakness of a tie but the structural hole it spans. Tie weakness is a correlate, not a cause. The structural-hole argument captures the causal agent directly, providing a stronger foundation for theory and a clearer guide for empirical research. Second, by shifting attention away from the structural hole responsible for information benefits to the strength of the tie providing them, the weak-tie argument obscures the control benefits of structural holes. Control benefits augment, and in some ways are more important than, the information benefits of structural holes. Building both benefits into the argument more clearly speaks to the generality of the

phenomenon under study. I'll elaborate the first point, then move to the second in the next section.

The weak-tie argument is about the strength of relationships at the same time that it is about their location. The two dashed lines in Figure 2-6 are bridges. They are the only connection between two otherwise separate clusters of strongly interconnected players (compare Granovetter 1973:1065 on weak ties as bridges). A bridge is at once two things. It is a chasm spanned and the span itself. By title and subsequent application, the weak-tie argument is about the strength of relationships that span the chasm between two social clusters. The structural-hole argument is about the chasm spanned. It is the latter that generates information benefits. Whether a relationship is strong or weak, it generates information benefits when it is a bridge over a structural hole.

Consider a cross-tabulation of ties by their strength and location. Your relationships can be sorted into two categories of strength. Strong ties are your most frequent and close contacts. Weak ties are your less frequent, less close contacts. Between these two categories you have a few strong ties and many weak ties.

Now sort by location, redundant ties within your social cluster versus nonredundant ties to people in other clusters. The nonredundant ties are your bridges to other clusters. From what we know about the natural etiology of relationships, bridges are less likely to develop than ties within clusters. The category of redundant ties includes your strong ties to often-met close friends and colleagues, but it also includes their friends and friends of their friends, whom you meet only occasionally if at all. As you expand your inventory from your closest, most frequent contacts to your more distant ones, contacts tend to be people like yourself before you reach a sufficiently low level of relationship to include people from completely separate social worlds. This tendency varies from one person to the next, but it is in the aggregate the substance of the well-documented tendency already discussed for relations to develop between socially similar people. In the Figure 2-6 illustration, you are one of nine people in your social cluster. You have strong ties to two people. Through those two, you have weak ties to the other six people in the cluster. To keep the sociogram simple, I deleted the dashed lines for those ties and their equivalent inside the other clusters. The other six people in your cluster are friends of friends, people whom you know and sometimes meet but don't have the time or energy to include among your closest contacts. The cluster is clearly held together by strong ties. Everyone has two to five strong ties to others within the cluster. All nine people are likely to know about the same opportunities as expected in a cohesive cluster. Of the thirty-six possible connections among the nine people in the cluster, however, only twelve are solid-line strong ties. The remaining two-thirds are weak ties between redundant friends of friends.

Now cross-tabulate the two classifications and take expected values. The result is given in Table 2-1. Information benefits vary across the columns of the table, higher through nonredundant ties. This is accurately

Table 2-1
The Natural Distribution of Relationships

	LOCATION IN SOCIAL STRUCTURE		Total
	Redundant Tie Within Cluster	Nonredundant Tie Beyond Cluster	
Strength			
Weak Tie	many	some	more
Strong Tie	some	rare	less
Total	more	less	

represented in both the weak-tie and the structural-hole argument. But quick reading of the weak-tie argument, with its emphasis on the strength of a relationship, has led some to test the idea that information benefits covary inversely with the strength of ties. This is a correlation between the rows and columns of Table 2-1, which is no correlation at all. In fact, the typical tie in Table 2-1 is weak and provides redundant information. The correlation in a study population depends on the distribution of ties in the table, but there is no theoretical reason to expect a strong correlation between the strength of a relationship and the information benefits it provides.

The weak-tie argument is about the two cells in the second column of the table. It predicts that nonredundant ties, the bridges that provide information benefits, are more likely weak than strong. In the second column of Table 2-1, weak-tie bridges are more likely than strong-tie bridges. To simplify his argument, Granovetter (1973:1063) makes this tendency absolute by ruling out strong-tie bridges (the "rare" cell in Table 2-1, the "forbidden triad" in Granovetter's argument). As Granovetter puts it: "A strong tie can be a bridge, therefore, only if neither party to it has any other strong ties, unlikely in a social network of any size (though possible in a small group). Weak ties suffer no such restriction, though they are certainly not automatically bridges. What is important, rather, is that all bridges are weak ties" (1064).

Bridge strength is an aside in the structural-hole argument. Information benefits are expected to travel over all bridges, strong or weak. Benefits vary between redundant and nonredundant ties, the columns of Table 2-1. Thus structural holes capture the condition directly responsible for the above-described information benefits. The task for a strategic player building an efficient-effective network is to focus resources on the maintenance of bridge ties. Otherwise, and this is the correlative substance of the weak-tie argument, bridges will fall into their natural state of being weak ties.

4. CONTROL AND THE *TERTIUS GAUDENS*

I've described how structural holes can determine who knows about opportunities, when they know, and who gets to participate in them. Players with a network optimized for structural holes enjoy higher rates

of return on their investments because they know about, and have a hand in, more rewarding opportunities.

They are also more likely to secure favorable terms in the opportunities they choose to pursue. The structural holes that generate information benefits also generate control benefits, giving certain players an advantage in negotiating their relationships. To describe how this is so, I break the negotiation into structural, motivational, and outcome components. The social structure of the competitive arena defines opportunities; a player decides to pursue an opportunity and is sometimes successful.

4.1 *Tertius Gaudens*

Beginning with the outcome, sometimes you will emerge successful from negotiation as the *tertius gaudens*. Taken from the work of Georg Simmel, the *tertius* role is useful here because it defines successful negotiation in terms of the social structure of the situation in which negotiation is successful. The role is the heart of Simmel's (1922) later analysis of the freedom an individual derives from conflicting group affiliations (see Coser 1975 for elaboration).⁴ The *tertius gaudens* is "the third who benefits" (Simmel 1923:154, 232).⁵ The phrase survives in what I am told is a well-known Italian proverb: *Far i due litiganti, il terzo gode* (Between two fighters, the third benefits). It has moved north to a more jovial Dutch phrase: *de lachende derde* (the laughing third).⁶ *Tertius*, *terzo*, or *derde*—the phrase describes an individual who profits from the disunion of others.

There are two *tertius* strategies: being the third between two or more players after the same relationship, and being the third between players in two or more relations with conflicting demands. The first, and simpler, strategy is the familiar economic bargaining between buyer and seller. Where two or more players want to buy something, the seller can play their bids against one another to get a higher price. The strategy extends directly to social commodities: a woman with multiple suitors, or a professor with simultaneous offers of positions in rival institutions.

The control benefits of having a choice between players after the same relationship extends directly to choice between the simultaneous demands of players in separate relationships. The strategy can be seen between hierarchical statuses in the enterprising subordinate under the authority of two or more superiors—for example, the student who strikes her own balance between the simultaneous demands of imperious faculty advisers.⁷ The bargaining isn't limited to situations of explicit competition. In some situations, emerging as the *tertius* depends on creating competition. In proposing the concept of a role-set, for example, Merton (1968:393–394) identifies this as a strategy to resolve conflicting role demands. Make simultaneous, contradictory demands explicit to the people posing them, and ask them to resolve their—now explicit—conflict. Even where it doesn't exist, competition can be produced by defining issues such that contact demands become contradictory and must be resolved before you

can meet their requests. Failure is possible. You might provide too little incentive for the contacts to resolve their differences. Contacts drawn from different social strata need not perceive one another's demands as carrying equal weight. Or you might provide too much incentive. Now aware of one another, the contacts could discover sufficient reason to cooperate in forcing you to meet their mutually agreed-on demands (Simmel 1902:176, 180–181 calls attention to such failures). But if the strategy is successful, the pressure on you is alleviated and replaced with an element of control over the negotiation. Merton states the situation succinctly: the player at the center of the network, ". . . originally at the focus of the conflict, virtually becomes a more or less influential bystander whose function it is to high-light the conflicting demands by members of his role-set and to make it a problem for them, rather than for him, to resolve *their* contradictory demands" (1968:430).

The strategy holds equally well with large groups. Under the rubric "divide and rule," Simmel (1902:185–186) describes institutional mechanisms through which the Incan and Venetian governments obtained advantage by creating conflict between subjects. The same point is illustrated more richly in Barkey's (1990) comparative description of state control in early seventeenth-century France and Turkey. After establishing the similar conditions in the two states at the time, Barkey asks why peasant-noble alliances developed in France against the central state while no analogous or substitutable alliances developed in Turkey. The two empires were comparable in many respects that scholars have cited to account for peasant revolt. They differed in one significant respect correlated with revolt—not in the structure of centralized state control, but in control strategy. In France, the king sent trusted representatives as agents to collect taxes and affect military decisions in provincial populations. The intrusion by these outside agents, *intendants*, affecting fundamental local decisions was resented by the established local nobility. Local nobility formed alliances with the peasantry against the central state. In Turkey, the sultan capitalized on conflict between leaders in the provinces. When a bandit became a serious threat to the recognized governor, a deal was struck with the bandit making him the legitimate governor. As Barkey puts it, "At its most extreme, the state could render a dangerous rebel legitimate overnight. This was accomplished by the striking of a bargain which ensured new sources of revenue for the rebel and momentary relief from internal warfare and perhaps, an army or two for the state" (1990:18). The two empires differed in their use of structural holes. The French king ignored them, assuming he had absolute authority. The Turkish sultan strategically exploited them, promoting competition between alternative leaders. Conflict within the Turkish empire remained in the province, rather than being directed against the central state. As is characteristic of the control obtained via structural holes, the resulting Turkish control was more negotiated than the absolute control exercised in France. It was also more effective.

4.2 The Essential Tension

There is a presumption of tension here. Control emerges from *tertius* brokering tension between other players. No tension, no *tertius*.

It is easy to infer that the tension presumed is one of combatants. Certainly there is a *tertius*-rich tension between combatants. Governors and bandits in the Turkish game played for life-or-death stakes. Illustrating this inference, a corporate executive listening to my argument expressed skepticism. Her colleagues, she explained, took pride in working together in a spirit of partnership and goodwill. The *tertius* imagery rang true to her knowledge of many firms, but not her own.

The reasoning is good, but the conclusion is wrong. I referred the skeptical executive to an analysis of hole effects that by coincidence was an analysis of managers at her level, in her firm (Burt 1992, Ch. 4). Promotions are strongly correlated, and illuminatingly so for women, with the structural holes in a manager's network.

The tension essential to the *tertius* is merely uncertainty. Separate the uncertainty of control from its consequences. The consequences of the control negotiation can be life or death in the extreme of combatants, or merely a question of embarrassment. Everyone knows you made an effort to get that job, but it went to someone else. The *tertius* strategies can be applied to control with severe consequences or to control of little consequence. What is essential is that control is uncertain, that no one can act as if they have absolute authority. Where there is any uncertainty about whose preferences should dominate a relationship, there is an opportunity for the *tertius* to broker the negotiation for control by playing demands against one another. There is no long-term contract that keeps a relationship strong, no legal binding that can secure the trust necessary to a productive relationship. Your network is a pulsing swirl of mixed, conflicting demands. Each contact wants your exclusive attention, your immediate response when a concern arises. Each, to warrant their continued confidence in you, wants to see you measure up to the values against which they judge themselves. Within this preference webwork, where no demands have absolute authority, the *tertius* negotiates for favorable terms.

4.3 The Connection with Information Benefits

This brings me back to information benefits. Structural holes are the setting for *tertius* strategies. Information is the substance. Accurate, ambiguous, or distorted information is moved between contacts by the *tertius*. One bidder is informed of a competitive offer in the first *tertius* strategy. A player in one relationship is informed of demands from other relationships in the second *tertius* strategy.

The two kinds of benefits augment and depend on each other. Application of the *tertius* strategies elicits additional information from contacts interested in resolving the negotiation in favor of their own preferences.

The information benefits of access, timing, and referrals enhance the application of strategy. Successful application of the *tertius* strategies involves bringing together players who are willing to negotiate, have sufficiently comparable resources to view one another's preferences as valid, but won't negotiate with one another directly to the exclusion of the *tertius*. Having access to information means being able to identify where there is an advantage to bringing contacts together and is the key to understanding the resources and preferences being played against one another. Having that information early is the difference between being the one who brings together contacts versus being just another person who hears about the negotiation. Referrals further enhance strategy. It is one thing to distribute information between two contacts during negotiation, another thing to have people close to each contact endorsing the legitimacy of the information you distribute.

5. ENTREPRENEURS

Behavior of a specific kind converts opportunity into higher rates of return. Information benefits of structural holes might come to a passive player, but control benefits require an active hand in the distribution of information. Motivation is now an issue. The *tertius* plays conflicting demands and preferences against one another, building value from their disunion. You enter the structural hole between two players to broker the relationship between them. Such behavior is not to everyone's taste. A player can respond in ways ranging from fully developing the opportunity to ignoring it. When you take the opportunity to be the *tertius*, you are an entrepreneur in the literal sense of the word—a person who generates profit from being between others. Both terms will be useful in these precise meanings; entrepreneur refers to a kind of behavior, the *tertius* is a successful entrepreneur.⁸

Motivation is often traced to cultural beliefs and psychological need. For example, in *The Protestant Ethic and the Spirit of Capitalism*, Weber describes the seventeenth-century bourgeois Protestant as an individual seeking—in religious duty, in Calvinist “calling”—the profit of sober, thrifty, diligent exploitation of opportunities for usury and trade (1905: especially 166ff). Psychological need is another motive. McClelland (1961) describes the formation of a need to achieve in childhood as critical to later entrepreneurial behavior (a need that can be cultivated later if desired: McClelland 1975). Schumpeter (1912) stresses nonutilitarian motives:

First of all, there is the dream and the will to found a private kingdom, usually, though not necessarily, also a dynasty. . . . Then there is the will to conquer: the impulse to fight, to prove oneself superior to others, to succeed for the sake, not of the fruits of success, but of the success itself. . . . Finally, there is the joy of creating, of getting things done, or simply of exercising one's energy and ingenuity (93).⁹

5.1 Opportunity and Motivation

These are powerful frameworks for understanding competition, but I don't wish to detour into the beliefs behind entrepreneurial behavior. I propose to leap over the motivation issue by taking, for three reasons, a player's network as simultaneously an indicator of entrepreneurial opportunity and motivation.

First, there is the clarity of an opportunity. Players can be pulled to entrepreneurial action by the promise of success. I do not mean that players are rational creatures expected to calculate accurately and act in their own interest. Nor do I mean to limit the scope of the argument to situations in which players act as if they were rational in that way. I mean simply that between two opportunities, any player is more likely to act on the one with the clearer path to success. The clarity of opportunity is its own motivation. As the number of entrepreneurial opportunities in a network increases, the odds of some being clearly defined by deep structural holes increases, so the odds of entrepreneurial behavior increase. To be sure, a person whose abilities or values proscribe entrepreneurial behavior is unlikely to act, and someone inclined to entrepreneurial behavior is more likely to act, even taking the initiative to create opportunities. Regardless of ability or values, however, within the broad range of acceptable behaviors a person is unlikely to take entrepreneurial action if the probability of success is low. You might question the propriety of a scholar negotiating between two universities that offer a position, but the question is not an issue for the player with one offer.

There are also network analogues to the psychological and cultural explanations of motive. Beginning with psychological need, a person with a taste for entrepreneurial behavior is prone to building a network configured around such behavior. If I find a player with a network rich in the structural holes that make entrepreneurial behavior possible, I have a player willing and able to act entrepreneurially. But it is the rare person who is the sole author of her or his network. Networks are more often built in the course of doing something else. Your work, for example, involves meeting people from very different walks of life, so your network ends up composed of contacts who without you have no contact with one another. Even so, the network is its own explanation of motive. As the volume of structural holes in a player's network increases—regardless of the process that created them—the entrepreneurial behavior of making and negotiating relations between others becomes a way of life. This is a network analogue to the cultural explanation of motive. If all you know is entrepreneurial relationships, the motivation question is a nonissue. Being willing and able to act entrepreneurially is how you understand social life.

I will treat motivation and opportunity as one and the same. For reasons of a clear path to success, or the tastes of the player as the network's author, or the nature of the player's environment as author of the network, a network rich in entrepreneurial opportunity surrounds a

player motivated to be entrepreneurial. At the other extreme, a player innocent of entrepreneurial motive lives in a network devoid of entrepreneurial opportunity.

5.2 Measurement Implications

Entrepreneurial motivation highlights a complexity that might otherwise obscure the association between structural holes and rates of return. Consider the graph in Figure 2-7. Players are defined by their rate of return on investments (vertical axis) and the entrepreneurial opportunities of structural holes in their networks (horizontal axis).

The sloping line in the graph describes the hole effect of players rich in structural holes (horizontal axis) getting higher rates of return on investments (vertical axis). The increasingly positive slope of the line captures the increasing likelihood of *tertius* profit. A player invests in certain relationships. They need not all be high-yield relationships. The higher the proportion of relationships enhanced by structural holes, the more likely and able the entrepreneurial player, and so the more likely that the

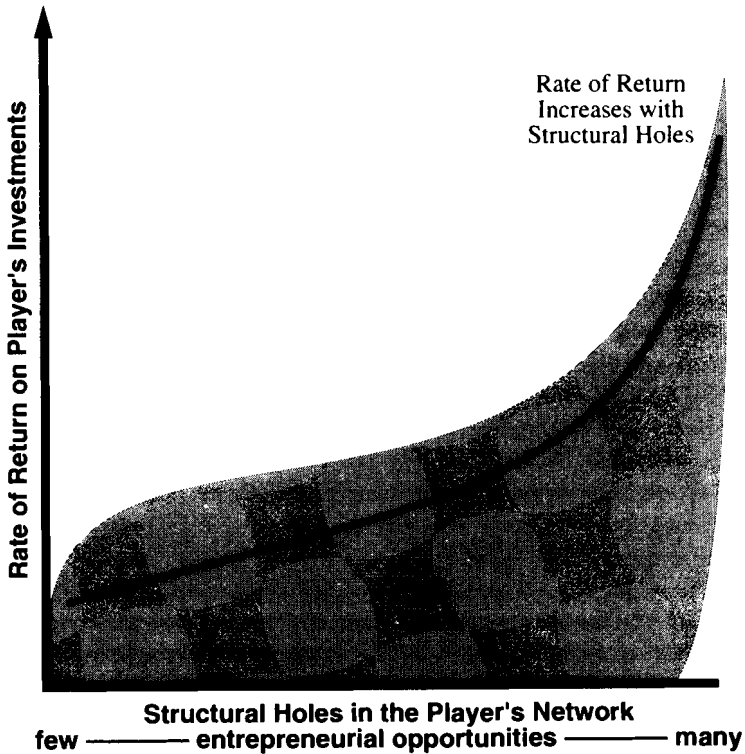


Figure 2-7
Rate of Return and Structural Holes

player's investments are in high-yield relationships. The result is a higher aggregate rate of return on investments.

I have shaded the area in the graph to indicate how I expect data to be distributed around the line of association. There is no imperative that says players have to take advantage of the benefits provided by structural holes. Players rich in entrepreneurial opportunity may choose to develop opportunities (and so appear in the upper right corner of the graph) or ignore them (and so appear in the lower right corner of the graph). Some players in Figure 2-7 are above the line. Some are below. If players were perfectly rational, observations would be clustered around the line. Players would take advantage of any entrepreneurial opportunity presented to them. A control for differences in player motivation, such as a McClelland measure of need for achievement, would have the same effect. The point is not the degree of deviation from the line of association; it is the greater deviation below the line. Variable motivation creates deviations below the true hole effect on rate of return.

This emphasizes the relative importance for empirical research of deviations above and below the line of association. Observations in the lower right corner of the graph, players under-utilizing their entrepreneurial opportunities, might be due to variation in motivation. Observations in the upper left corner are a severe test of the argument. Players who have opportunities can choose whether to develop them. Players without opportunities don't have that choice. Within the limits of measurement error, there should be no observations in the upper left corner of the graph.

6. STRUCTURAL AUTONOMY

I can now summarize the argument with a concept defining the extent to which a player's network is rich in structural holes, and so entrepreneurial opportunity, and so information and control benefits. That concept is structural autonomy. I will present the concept in a general way here (see Burt 1992, Ch. 2, for formal details).

I began with a generic production equation. Profit equals an investment multiplied by a rate of return. The benefits of a relationship can be expressed in an analogous form: time and energy invested to reach a contact multiplied by a rate of return. A player's entrepreneurial opportunities are enhanced by a relationship to the extent that (1) the player has invested substantial time and energy to secure a connection with the contact, and (2) there are many structural holes around the contact ensuring a high rate of return on the investment. More specifically, rate of return concerns how and who you reach with the relationship. Time and energy invested to reach a player with more resources generates more social capital. For the sake of argument, as explained in the discussion of social capital, I assume that a player with a network optimized for structural holes can identify suitably endowed contacts such that I don't have

to carry the issue of who as another variable in the analysis. My concern is the how of a relationship, defined by the structure of a network and its connection with the social structure of the competitive arena. Thus the rate of return keyed to structural holes is a product of the extent to which there are (1) many primary structural holes between the contact and others in the player's network, and (2) many secondary structural holes between the contact and others outside the network who could replace the contact.

There is also the issue of structural holes around the player. As the holes around contacts provide information and control benefits to the player, holes around the player can be developed by contacts for their benefit. Developing entrepreneurial opportunities depends on having numerous structural holes around your contacts and none attached to yourself.

These considerations come together in the concept of structural autonomy. Players with relationships free of structural holes at their own end and rich in structural holes at the other end are structurally autonomous. These are the players best positioned for the information and control benefits that a network can provide. These are the players to the far right of the graph in Figure 2-7. Structural autonomy summarizes the action potential of the *tertius's* network. The budget equation for optimizing structural autonomy has an upper limit set by the *tertius's* time and energy, and a trade-off between the structural holes a new contact provides versus the time and energy required to maintain a productive relationship with the contact.¹⁰ The summary conclusion is that players with networks optimized for structural holes—that is to say, players with networks providing high structural autonomy—enjoy higher rates of return on their investments because they know about, have a hand in, and exercise control over, more rewarding opportunities.

7. THE BROADER CONTEXT

The structural-hole argument has four signature qualities:

1. Competition is a matter of relations, not player attributes.
2. Competition is a relation emergent, not observed.
3. Competition is a process, not just a result.
4. Imperfect competition is a matter of freedom, not just power.

These qualities are not individually unique to the structural-hole argument. They are jointly characteristic of it.

First, competition is a matter of relationships, not player attributes. The structural-hole argument escapes the debilitating social science practice of using player attributes for explanation. The unit of analysis in which structural holes have their causal effect is the network of relations that intersect in a player. The intersection is known by various names as

a role, a market, or a position in social structure. The players in which relations intersect are physical and legal entities; a person, an organization, or a broader aggregation of physical and legal entities. The attributes of the players in whom the relations intersect—black, white, female, male, old, young, rich, poor—are an empirical curiosity irrelevant to the explanation. Competition is not about being a player with certain physical attributes; it is about securing productive relationships. Physical attributes are a correlate, not a cause, of competitive success. Causation resides in the intersection of relations. Holes can have different effects for people with different attributes or for organizations of different kinds, but that is because the attributes and organizational forms are correlated with different positions in social structure. The manner in which a structural hole is an entrepreneurial opportunity for information and control benefits is the bedrock explanation that carries across player attributes, populations, and time. The task for the analyst is to cut past the spurious correlation between attributes and outcomes to reach the underlying social structural factors that cause the outcome. This point is developed at length elsewhere (Burt 1992, Ch. 5).

Second, competition is a relation emergent, not observed. The structural holes in which competition develops are invisible relations of non-redundancy, relations visible only by their absence. Consider the atavistic driver experiment. You're on the freeway. There is a car ahead of you going 65 MPH. Pull up so your front wheels are parallel to his. Stay there. This won't take long. If he speeds up, speed up. If he slows down, slow down. You feel the tension in yourself as you know it's building in the next car. He looks over. Is this a sexual come-on or a threat? Deciding against sex, he may slow down, hoping you'll go away. If that doesn't work, and he doesn't feel that his car can escape yours, his anger will be apparent only on his face. If he is more confident, he'll accelerate to get away from you. Let him.

For the moment when you two stood in common time and place, you were competitors. Break the parallelism, and the competition is gone. There is no behavioral relationship between the drivers that is competition. Competition is an intense, intimate, transitory, invisible relationship created between players by their visible relations with others. It's the cheek-by-jowl with respect to the passing environment that makes the drivers competitors.

The task of analyzing competition is made more difficult by the fact that the structural holes in which competition thrives don't connect the players we see. They connect invisible pieces of players; the pieces we see in any one of the many roles and markets in which the person or firm is a player. I see one piece of you in the office, another on the street, another at home. Each piece has an attendant network of relations with relevant others. The causal force of structural holes resides in the pattern of relations that intersect in each network. That intersection happens in players, but where it occurs is distinct from the causal force released by its occurrence. This is another view of my first point, people and organizations

are not the source of action so much as they are the vehicles for structurally induced action.

These qualities make it very difficult to capture competition without having the conceptual and research tools to represent the social structure of the competitive arena. Understanding competition will be one of the important returns from the work invested during the 1970s and 1980s in network analysis. The social structure of competition is not about the structure of competitive relations. It is about the social structure of the relations for which players compete. The structural-hole argument is not a theory about competitive relationships. It is a theory about competition for the benefits of relationships. To explain variation in competitive success, I have to look beyond the competitors themselves to the circumstances of the relations for which they compete. The terrain on which competition plays out lies beyond the competitors themselves. It lies in their efforts to negotiate relations with other players. Where those relations are positioned in social structure such that there is little room to negotiate, the margin between success and failure is slim. The social structure of competition is about the negotiability of the relationships on which competitors survive. That is the essence of the structural autonomy concept.

Third, competition is a process, not just a result. With important exceptions, most competition theories are about what is left when competition is over. They are an aside in efforts to answer the practical question of how to maximize producer profit. Answering the question requires a definition of how price varies with output. It is convenient to assume that there is a condition of "competition" such that price is constant with output. The presumed competition exists when (1) there are an infinite number of buyers and sellers known to one another, (2) goods can be divided for sale to any number of buyers, and (3) buyers and sellers are free to exchange without interference from third parties. When goods are exchanged under these conditions, conditions of "perfect" competition, equilibrium prices can be derived that will clear the market. An architecture of powerful economic theory about price and production follows.¹¹

The alternative is to start with the process of competition and work toward its results. This is a less elegant route for theory, but one that veers closer to the reality of competition as we experience it. The structural-hole argument is not about the flow of goods. No mechanism is proposed to define the prices that "clear" the imperfectly competitive market. Such a mechanism could be proposed, but not here. This chapter is about the competitive process by which the price and occurrence of transactions is decided. If you will, it is about the players who form the deal, not the lawyers who write the contracts. The social structure of competition is about negotiating the relationships on which competitors survive. Structural holes determine the extent to which, and manner in which, certain players have a competitive advantage in that negotiation.

Fourth, imperfect competition is a matter of freedom, not just power.

The structural-hole argument is a theory of competition made imperfect by the freedom of individuals to be entrepreneurs. In this, the theory cuts across the usual axis of imperfect competition.

In the perfectly competitive arena, any party to a transaction has unlimited choice between alternative partners. Numerous alternatives exist and players are free to choose. The fact of that choice drives price to a minimum. The significance of any one player as an entrepreneur is zero. The structural image is one of relational chaos. Players are free to withdraw from existing relations to join with anyone who better serves their interests. Obligation stops with the execution of the transaction.

Deviations from this image measure imperfect competition, usually defined by the extent to which choice is concentrated in the hands of the strongest player. As Stigler (1957:262) concludes his historical review: "If we were free to redefine competition at this late date, a persuasive case could be made that it should be restricted to meaning the absence of monopoly power in a market." At the extreme of perfect competition, every player has unlimited choice among alternative relationships. At the other extreme, choice is concentrated in the hands of a dominant player. Everyone else is assigned to relations by the dominant player. Familiar images are monopoly, cults, village kinship systems, political machines, fascist bureaucracies. The structural image is one of a completely and rigidly interconnected system of people and establishments within a market. High obligation relations, with obligation enforced by authority or convention, allow neither negotiation nor the strategic replacement of partners.

Observed behavior lies between these extremes. Control is never absolute; it is negotiated—whether exercised through competitive price, bureaucratic authority, or some other social norm. In the most regulated arena, there are special relationships through which certain players move to get around the dicta of the governing mechanism. In the most competitive of arenas, there are relations between certain players that provide them special advantages. Competition is omnipresent and everywhere imperfect.

The extremes of perfect and regulated competition are more similar on a critical point than either is to the reality of observed behavior between them. They are both images of dominance. Players are homogeneously trivial under competitive market pricing and, at the other extreme, homogeneously trivial under the dicta of the dominant player. The dominant player defines fair exchange in the regulated market. Buyer and seller are locked into exchange relations by the dicta of the dominant player. The press of numbers defines fair exchange in the perfectly competitive market. Competition between countless buyers and sellers involves negotiation between alternative relations, not within a relationship. Any one partner in a relationship is a faceless cog, readily replaced with someone else. At either extreme, the lack of negotiation within a relationship denies the individuality of buyer and seller.

But their individuality is the key to understanding competition. The

substantive richness of competition lies in its imperfections, the jostling of specific players against one another looking for a way to make a difference. In the substantive details of imperfect competition lie the defining parameters of competition. They are the parameters of player individuality. Competition is imperfect to the extent that any player can affect the terms of any particular relationship. Oligopoly, the extent to which multiple players together constitute a monopoly, is an insufficient answer. The central question for imperfect competition is how players escape domination either by the market or by another player.

That is the focus of the structural-hole argument—a theory of freedom instead of power, of negotiated control instead of absolute control. It is a description of the extent to which the social structure of a competitive arena contains entrepreneurial opportunities for individual players to affect the terms of their relationships.

Notes

1. This chapter is approximately two-thirds of a chapter by the same name in a book, *Structural Holes*, to be published in 1992 by Harvard University Press. Professor Richard Swedberg, who skillfully condensed the original material for the purposes of this anthology, called my attention to an assumption implicit in my argument. I refer to people and organizations in the competitive arena as "players." Professor Swedberg felt I used the term to denote a very active actor, seeking out contacts and opportunities. He gently suggested that the term had a touch of frivolity and that I might do well to substitute with a more neutral term such as "actor." In earlier writing I have used the more neutral term for general discussion (Burt 1982), but for the topic of competition I prefer to use player. It better fits my felt-reality of the phenomenon. More than implying activity, it is a term of peer recognition: "Yes, he's a player." He's a presence in the game. If you have the motivation, resources, and skills to compete, you're a player; otherwise, you're scenery. Everyone is a player in some arenas, scenery in most. This chapter is about the social structural conditions that give certain players a competitive advantage.
2. This point is significant because it contradicts the natural growth of contact networks. Left to the natural course of events, a network will accumulate redundant contacts. Friends introduce you to their friends and expect you to like them. Business contacts introduce you to their colleagues. You will like the people you meet in this way. The factors that make your friends attractive make their friends attractive because like seeks out like. Your network grows to include more and more people. These relations come easily, they are comfortable, and they are easy to maintain. But these easily accumulated contacts do not expand the network so much as they fatten it, weakening its efficiency and effectiveness by increasing contact redundancy and tying up time. The process is amplified by spending time in a single place—in your family, in your neighborhood, or in the office. The more time you spend with any specific primary contact, the more likely you will be introduced to their friends. Evidence of these processes can be found in studies of balance and transitivity in social relations (see Burt 1982:55–60 for review) and in studies of the tendency for redundant relations to develop among physically proximate people (for example, the suggestively detailed work of Festinger, Schachter, and Back 1950; or the work with more definitive data on social contexts by Fischer 1982, and on social foci by Feld 1981, 1982). For the purposes here, I ignore the many day-to-day tactical issues critical to maintaining

a network. Thorough treatment requires considerable discussion and didactic devices. This is the function of the seminars offered by the Denver firm, Strategic Connections. I discuss tactical issues in a short book, *The Network Entrepreneur*, written in 1987 for distribution from the firm.

3. The number of structural holes is not increased directly but is likely to increase. The presumption through all this is that the time and energy to maintain relationships is limited and the constant pressure to include new contacts will use all time and energy available (as in the preceding footnote). Although structural holes are not increased directly by maximizing nonredundant contacts, they can be expected to increase indirectly from the reallocation of time and energy from maintaining redundant contacts to acquiring new nonredundant contacts (as illustrated in Figure 2-4).
4. This theme is often grouped with Durkheim's (1893) argument for the liberating effect of a division of labor, but it is useful to distinguish the two arguments for the present purposes. Simmel focuses on the liberating quality of competition between multiple affiliations, which is the concern here. Durkheim focuses on the liberating quality of interdependent affiliations. Integration, rather than competition, is Durkheim's theme. That theme continues in Blau's (1977) analysis of cross-cutting social circles, in which he argues that conflict between strata becomes increasingly difficult as affiliations provide people with alternative stratification hierarchies. Flap (1988) provides a network-oriented review of such work, building from anthropology and political science, to study the "crisscross" effect inhibiting violence.
5. Georg Simmel introduced this phrase in papers on the importance of group size, translated and published by Albion Small in the *American Journal of Sociology* (Simmel 1896:393-394, 1902:174-189). A later version was translated by K. H. Wolff (Simmel 1923:154-169, 232-234).
6. I am grateful to Anna DiLellio for calling my attention to the Italian proverb and to Hein Schreuder for calling my attention to the Dutch expression. The idea of exploiting a structural hole is viscerally familiar to all audiences, but interestingly varied across cultures in phrasing the profit obtained (an interesting site for a Zelizer-1989 kind of analysis).
7. This point is nicely exemplified in Simmel's (1896:394) discussion of subordination comparing the freedom of two medieval subordinate positions: the bondsman ("unfree") and the vassal:

An essential difference between the medieval "unfree" men and the vassals consisted in the fact that the former had and could have only one master, while the latter could accept land from different lords and could take the oath of fealty to each. By reason of the possibility of placing themselves in the feudal relation to several persons the vassals won strong security and independence against the individual lords. The inferiority of the position of vassalage was thereby to a considerable degree equalized.

8. A substantial block of material was deleted between here and the next section on (1) the literal meaning of entrepreneurs, (2) the importance of structural holes within the clusters of secondary contacts, (3) market boundaries, and (4) a more careful discussion of holes defined by cohesion versus structural equivalence. If the leap to structural autonomy seems awkward here, consider looking at the full discussion (Burt 1992, Ch. 1).
9. I am grateful to Richard Swedberg for giving me the benefit of his careful study of Schumpeter in calling my attention to these passages. Their broader scope and context are engagingly laid out in his biography of Schumpeter (Swedberg 1991). The passages can also be found in the Schumpeter selection included in Parsons et al., *Theories of Society* (1961:513).
10. This sentence is the starting point for an optimization model in which the benefits of a contact are weighed against the cost of maintaining a relation

with the contact, subject to a time and energy budget constraint on the aggregate of contacts in a network. The work is beyond the scope of this discussion, but I want to remove an ostensible barrier to such work, and in the process highlight a scope limitation to my argument. Marks (1977) provides a cogent argument against the energy scarcity metaphor so often used to justify discussions of role negotiations. Instead of viewing roles as energy debilitating, Marks argues for an "expansion" view in which energy is created by performing roles (compare Sieber 1974). Marks and Sieber discuss the advantages of performing multiple roles. Both are responding to the energy-scarcity arguments used to motivate discussions of mechanisms by which people manage role strain (most notably, Merton 1957; Goode 1960). To quote Goode (1960:485), a person "... cannot meet all these demands to the satisfaction of all the persons who are part of his total role network. Role strain—difficulty in meeting given role demands—is therefore normal. In general, the person's total role obligations are overdemanding." I have borrowed the theme of overdemanding role obligations. The *tertius* budget constraint concerns both the time and energy cost of maintaining existing relations and the opportunity costs of contacts lost because of redundancy. However, my argument only concerns negotiations within a single role. The mechanisms used to manage role strain, such as segregating role relations in time and space, could also be used by the *tertius* to manage conflict to his or her own advantage, but I am ignoring that possibility, and so limiting the scope of my argument, to focus on the situation in which *tertius* negotiates conflicting demands that have to be met simultaneously.

11. This paragraph owes much to Stigler's (1957) review of the evolution of competition in economic theory. He provides the simple profit question that calls for an assumption of competition. The three conditions for perfect competition are adapted from Edgeworth (1881:17–19), but I appreciated their evolutionary significance only in the context of alternatives laid out in Stigler's (1957) review. Beyond providing context, the clarity of Stigler's presentation, here and with respect to Edgeworth on marginal utility, offers a great improvement over the original. At the same time, as always, the original has value. Edgeworth's characterization of free choice in terms of no intrusive third parties is the key to the social structure of competition. Structural holes are the variable determining the extent to which there are no intrusive third parties to a relationship. Stigler's (1957:247) recoding of that to be the "complete absence of limitations upon individual self-seeking behavior" states the original thought in terms more compatible with subsequent developments in economic theory, but obscures the social structural insight in the original.

References

- Barkey, Karen. 1990. "Rebellious Alliances: The State and Peasant Unrest in Early Seventeenth Century France and the Ottoman Empire." Paper presented at the 1989 annual meetings of the American Sociological Association.
- Berkman, Lisa F., and S. Leonard Syme. 1979. "Social Networks, Host Resistance, and Mortality: A Nine-Year Follow-Up Study of Alameda County Residents." *American Journal of Epidemiology* 109:186–204.
- Blau, Peter M. 1977. *Heterogeneity and Inequality*. New York: Free Press.
- Boxman, Ed A. W., Paul M. De Graaf, and Hendrik D. Flap. 1991. "The Impact of Social and Human Capital on the Income Attainment of Dutch Managers." *Social Networks* 13:in press.
- Burt, Ronald S. 1982. *Toward a Structural Theory of Action*. New York: Academic Press.

- . 1986. "A Note on Sociometric Order in the General Social Survey Network Data." *Social Networks* 8:149–174.
- . 1990. "Kinds of Relations in American Discussion Networks." In C. Calhoun, M. W. Meyer, and W. R. Scott, eds., *Structures of Power and Constraint*. New York: Cambridge University Press, pp. 411–451.
- . 1992. *Structural Holes*. Cambridge: Harvard University Press.
- Campbell, Karen E., Peter V. Marsden, and Jeanne S. Hurlbert. 1986. "Social Resources and Socioeconomic Status." *Social Networks* 8:97–117.
- Coleman, James S. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94, supplement: S95–120.
- Coser, Rose Laub. 1975. "The Complexity of Roles as a Seedbed of Individual Autonomy." In L. A. Coser, ed., *The Idea of Social Structure*. New York: Harcourt, Brace, Jovanovich.
- De Graaf, Nan D., and Hendrik D. Flap. 1988. "With a Little Help from My Friends." *Social Forces* 67: 453–472.
- Durkheim, Émile [1893] 1933. *The Division of Labor in Society*, G. Simpson, tr. New York: Free Press.
- Edgeworth, F. Y. 1881. *Mathematical Psychics*. London: C. Kegan Paul.
- Feld, Scott. L. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86:1015–1035.
- . 1982. "Social Structural Determinants of Similarity." *American Sociological Review* 47:797–801.
- Festinger, Leon, Stanley Schachter, and Kurt W. Back. 1950. *Social Pressures in Informal Groups*. Stanford: Stanford University Press.
- Fischer, Claude S. 1982. *To Dwell Among Friends*. Chicago: University of Chicago Press.
- Flap, Hendrik D. 1988. *Conflict, Loyalty, and Violence*. New York: Verlag Peter Lang.
- Flap, Hendrik D., and Nan D. De Graaf. 1989. "Social Capital and Attained Occupational Status." *Netherlands' Journal of Sociology* 22:145–161.
- Flap, Hendrik D., and F. Tazelaar. 1989. "The Role of Informal Social Networks on the Labor Market: Flexibilization and Closure." In H. Flap, ed., *Flexibilization of the Labor Market*. Utrecht, Holland: ISOR, University of Utrecht, pp. 99–118.
- Goode, William J. 1960. "A Theory of Role Strain." *American Sociological Review* 25:483–496.
- Granovetter, Mark S. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–1380.
- Homans, George C. 1950. *The Human Group*. New York: Harcourt, Brace and World.
- Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1944. *The People's Choice*. New York: Columbia University Press.
- Lin, Nan. 1982. "Social Resources and Instrumental Action." In P. V. Marsden and Nan Lin, eds., *Social Structure and Network Analysis*. Beverly Hills: Sage.
- Lin, Nan, and Mary Dumin. 1986. "Access to Occupations Through Social Ties." *Social Networks* 8:365–385.
- Lin, Nan, Walter M. Ensel, and John C. Vaughn. 1981. "Social Resources and Strength of Ties." *American Sociological Review* 46:393–405.
- Marks, Stephen R. 1977. "Multiple Roles and Role Strain: Some Notes on Human Energy, Time and Commitment." *American Sociological Review* 42:921–936.
- Marsden, Peter V. 1987. "Core Discussion Networks of Americans." *American Sociological Review* 52:122–131.
- Marsden, Peter V., and Jeanne S. Hurlbert. 1988. "Social Resources and Mobility Outcomes: A Replication and Extension." *Social Forces* 67:1038–1059.
- McClelland, David C. 1961. *The Achieving Society*. Princeton: Van Nostrand.
- . 1975. *Power*. New York: Irvington.
- Merton, Robert K. [1957] 1968. "Continuities in the Theory of Reference Group

- Behavior." In Robert K. Merton, ed., *Social Theory and Social Structure*. New York: Free Press, pp. 335-440.
- Parsons, Talcott, Edward Shils, Kaspar D. Naegle, and Jesse R. Pitts. 1961. *Theories of Society*. New York: Free Press.
- Schumpeter, Joseph A. [1912] 1961. *The Theory of Economic Development*, R. Opie, tr. Cambridge, MA: Harvard University Press.
- Sieber, Sam D. 1974. "Toward a Theory of Role Accumulation." *American Sociological Review* 39:567-578.
- Simmel, Georg. 1896. "Superiority and Subordination as Subject-Matter of Sociology, II," A. Small, tr. *American Journal of Sociology* 2:392-415.
- . 1902. "The Number of Members as Determining the Sociological Form of the Group, II," A. Small, tr. *American Journal of Sociology* 8:158-196.
- . [1922] 1955. *Conflict and Web of Group Affiliations*, K. H. Wolff and R. Bendix, trs. New York: Free Press.
- . [1923] 1950. *The Sociology of Georg Simmel*, K. H. Wolff, tr. New York: Free Press.
- Stigler, George J. [1957] 1965. "Perfect Competition, Historically Contemplated." In G. J. Stigler, ed., *Essays in the History of Economics*. Chicago: University of Chicago Press, pp. 234-267.
- Swedberg, Richard. 1991. *Schumpeter—A Biography*. Princeton: Princeton University Press.
- Weber, Max. [1904-1905] 1930. *The Protestant Ethic and the Spirit of Capitalism*, T. Parsons, tr. New York: Charles Scribner's Sons.
- Zelizer, Viviana A. 1989. "The Social Meaning of Money: 'Special Monies.'" *American Journal of Sociology* 95:342-377.