# Appendix B
# Measuring Access to Structural Holes

Structural holes are opportunities to broker connections between people. Access to structural holes indicates your opportunities to broker connections. When everyone you know is connected with one another, you have no opportunities to make connections. When you know a lot of people disconnected from one another, then you have opportunities to make connections between people otherwise disconnected, connections in terms of coordination between the disconnected people, and connections in terms of ideas or resources derived from exposure to contacts who differ in opinion or the way they behave. In this, access to structural holes can be said to measure the extent to which a manager's network gives him a vision advantage in detecting and developing opportunities, the extent to which the network puts him at risk of productive accident. This appendix is about the access people have to structural holes. Industry access is discussed in Chapter 5 as an extension of the discussion here.

"Opportunities" should be emphasized in these sentences. None of the network measures to be discussed index brokerage behavior. They index opportunities for brokerage. There are reliability, cost, and precedence reasons to measure brokerage opportunity instead of behavior, as discussed around Figure 2.4 in Chapter 2. Reasons notwithstanding, measuring brokerage behavior by its opportunities rather than its occurrence has implications, again as discussed in Chapter 2.

The implication relevant to this Appendix is that three brokerage terms are used as synonyms in current practice. I want to be clear about it to avoid confusion when the measures are discussed: Access to structural holes is discussed as synonymous with brokerage opportunities, both of which are discussed as synonymous with brokerage. All three terms are about the advantage created when connections are made between disconnected people, connections in terms of coordination between the disconnected people, or connections in terms of ideas or resources derived from exposure to contacts who differ in opinion or the way they behave.

**Bridge Counts**

Bridge counts are an intuitively appealing measure. The relation between two people is a bridge if there are no indirect connections between the two people through mutual contacts. Associations with performance have been reported measuring brokerage with a count of bridges (e.g., Burt, Hogarth, and Michaud, 2000:Appendix; Burt, 2002).

**Constraint**

I measure brokerage opportunities with a summary index, network constraint which measures the extent to which a manager's time and energy are concentrated in a single group of interconnected colleagues – which means no access to structural holes (Burt, 1992: Chap. 2):

$$C_i = \sum_j c_{ij},\ i \neq j \qquad \text{(B1)}$$

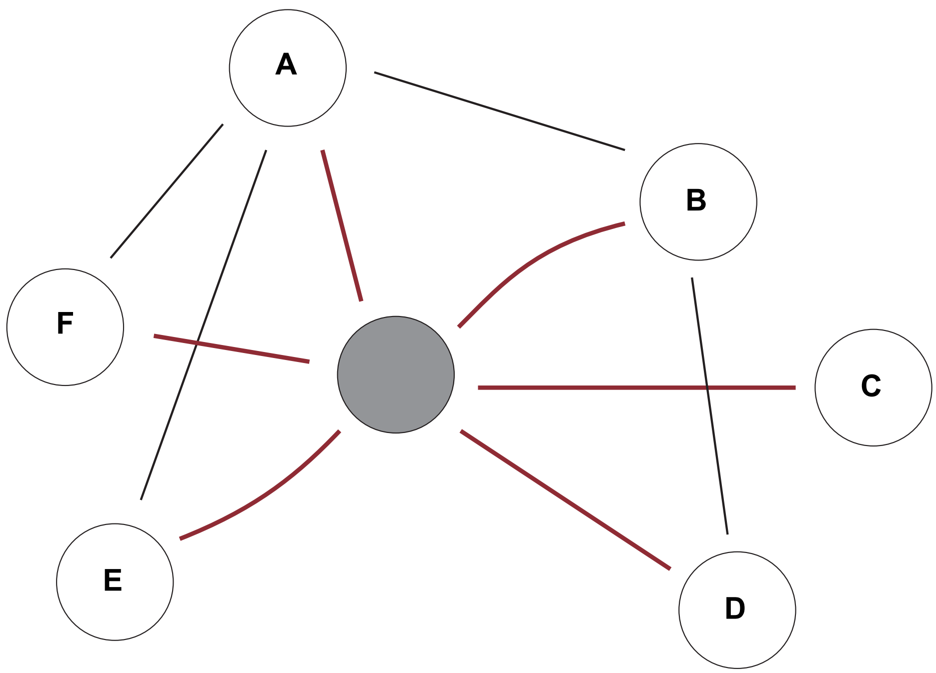
where $C_i$ is network constraint on manager i, and $c_{ij}$ is a measure of i's dependence on contact j:

$$c_{ij} = (p_{ij} + \Sigma_q p_{iq} p_{qj})^2,\ \ i \neq q \neq j \qquad \text{(B2)}$$

where $p_{ij}$ is the proportion of manager i's network time and energy spent on contact j, $p_{ij} = z_{ij} / \Sigma_q z_{iq}$, and variable $z_{ij}$ measures the strength of connection between contacts i and j, so the contact-specific constraint $c_{ij}$ varies from 0 to 1 with the extent to which i's network time and energy is directly ($p_{ij}$) or indirectly ($\Sigma_q p_{iq} p_{qj}$) spent on colleague j. Connection $z_{ij}$ measures the lack of a structural hole so it is made symmetric before computing $p_{ij}$ in that a hole between i and j is unlikely to the extent that either i or j feels that they spend a lot of time in the relationship (strength of connection "between" i and j versus strength of connection "from" i to j; see Burt, 1992:51). Network constraint, as the sum of $c_{ij}$, measures the extent to which the manager's network of colleagues is like a straightjacket around the manager, limiting his or her vision of alternative ideas and sources of support. I multiply scores by 100 to discuss integer levels of constraint. The calculation is illustrated in Figure B1.

Network constraint varies with three network dimensions: size, density, and hierarchy. Constraint on a person is high if the person has few contacts (small network) and those contacts are strongly connected to one another, either directly (as in a dense network), or through a central, mutual contact (as in a hierarchical network). The index,

# Figure B1
# Computing Network Constraint




Network constraint measures the extent to which your network time and energy is concentrated in a single group. There are two components: (direct) a contact consumes a large proportion of your network time and energy, and (indirect) a contact controls other people who consume a large proportion of your network time and energy. The proportion of i's network time and energy allocated to j, $p_{ij}$, is the ratio of $z_{ij}$ to the sum of i's relations, where $z_{ij}$ is the strength of connection between i and j, here simplified to zero versus one.

$$c_{ij} = (p_{ij} + \Sigma_q \, p_{iq} p_{qj})^2 \quad q \neq i,j$$

contact-specific constraint (x100):

| | |
|---|---|
| **A** | 15.1 |
| **B** | 8.5 |
| **C** | 2.8 |
| **D** | 4.9 |
| **E** | 4.3 |
| **F** | 4.3 |
| total | 39.9 |

100(1/36)

total 39.9 = aggregate constraint ($C = \Sigma_j \, c_{ij}$)

network data

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| A | . | 1 | 0 | 0 | 1 | 1 | 1 |
| B | 1 | . | 0 | 1 | 0 | 0 | 1 |
| C | 0 | 0 | . | 0 | 0 | 0 | 1 |
| D | 0 | 1 | 0 | . | 0 | 0 | 1 |
| E | 1 | 0 | 0 | 0 | . | 0 | 1 |
| F | 1 | 0 | 0 | 0 | 0 | . | 1 |
| gray dot | 1 | 1 | 1 | 1 | 1 | 1 | . |

C, can be written as the sum of three variables: $\Sigma_j(p_{ij})^2 + 2\Sigma_j p_{ij}(\Sigma_q p_{iq} p_{qj}) + \Sigma_j(\Sigma_q p_{iq} p_{qj})^2$. The first term in the expression, C-size in Burt (1998a), is a Herfindahl index measuring the extent to which manager i's relations are concentrated in a single contact. The second term, C-density in Burt (1998a), is an interaction between strong ties and density in the sense that it increases with the extent to which manager i's strongest relations are with contacts strongly tied to the other contacts. The third term, C-hierarchy in Burt (1998a), measures the extent to which manager i's contacts concentrate their relations in one central contact. See Burt (1992:50ff.; 1998a:Appendix) and Borgatti, Jones, and Everett (1998) for further discussion of the components in network constraint.

**Size**

Network size, N, is the number of contacts in a network. In graph-theory discussions, the size of the network around a person is discussed as "degree." Isolates, that is, people with no contacts, are a special case discussed below. For non-zero network size, other things equal, more contacts mean that a manager is more likely to receive diverse bits of information from contacts and is more able to play their individual demands against one another. Network constraint is lower in larger networks because the proportion of a manager's network time and energy allocated to any one contact ($p_{ij}$ in equation B2) decreases on average as the number of contacts increases.

**Density**

Density is the average strength of connection between contacts: $\Sigma\, z_{ij} / N^*(N-1)$, where summation is across all contacts i and j. Density is sometimes discussed as a proportion because in studies limited to binary network data (people are connected or not), the average strength of connection between contacts equals the proportion of contact pairs connected. Dense networks are more constraining since there are more connections ($\Sigma_q p_{iq} p_{qj}$ in equation B2). Connections between contacts increase the probability that the contacts know the same information and eliminate opportunities to broker information between contacts. Dense networks offer less of the information and control advantage associated with spanning structural holes. Density is only one form of network closure, but it is a form often discussed as closure. Contacts in a dense network are in close communication so they can readily enforce sanctions against individuals who violate shared beliefs or norms of behavior.

Hypothetical networks in Figure B2 illustrate how constraint varies with size, density, and hierarchy. Relations are binary and symmetric in Figure B2. The graphs display relations between contacts. Relations with the respondent are not presented. The first column contains sparse (minimum density) networks. No contact is connected with other contacts. The second column of the figure contains maximum-density networks. Every contact has a strong connection with each other contact. At each network size, constraint is lower in the sparse-network column.

**Hierarchy**

Density is a form of closure in which contacts are equally connected. Hierarchy is another form of closure in which a minority of contacts, typically one or two, stand above the others for being more the source of closure. The extreme is to have a network organized around one contact. For people in job transition, such as M.B.A. students, that one contact is often the spouse. In organizations, hierarchical networks are often built around the boss.

Hierarchy and density both increase, but in different ways, the indirect connection component in network constraint ($\Sigma_q p_{iq} p_{qj}$ in equation B2). Where network constraint measures the extent to which contacts are redundant, network hierarchy measures the extent to which the redundancy can be traced to a single contact in the network. The central contact in a hierarchical network gets the same information available to the manager and cannot be avoided in manager negotiations with each other contact. More, the central contact can be played against the manager by third parties because information available from the manager is equally available from the central contact since manager and central contact reach the same people. In short, the manager whose network is built around a central contact runs a risk of playing Tonto to the central contact's Lone Ranger. Network constraint increases with both density and hierarchy, but density and hierarchy are empirically distinct measures and fundamentally distinct with respect to social capital because it is hierarchy that measures social capital borrowed from a partner (the central point in Chapter 7).

The Coleman-Theil inequality index has attractive qualities as a measure of hierarchy (Burt, 1992:70ff.). Applied to contact-specific constraint scores, the index is the ratio of $\Sigma_j r_j \ln(r_j)$ divided by N ln(N), where N is number of contacts, $r_j$ is the ratio of contact-j constraint over average constraint, $c_{ij}/(C/N)$. The ratio equals zero if all contact-specific constraints equal the average, and approaches 1.0 to the extent that all constraint is from one contact. Again, I multiply scores by 100 and report integer points of hierarchy.

In the first and second columns of Figure B2, no one contact is more connected than others, so all of the hierarchy scores are zero. Non-zero hierarchy scores occur in the third column of Figure B2, where one central contact is connected to all others who are otherwise disconnected from one another. The hierarchy can be seen in the relative levels of constraint posed by individual contacts. Contact A poses more severe constraint than the others because network ties are concentrated in A (cf. contact A in Figure B1). The Coleman-Theil index increases with the number of people connected to the central contact (the difference between minimum and maximum constraint is larger in larger hierarchical networks). Hierarchy is 7 in the third column of Figure B2 for the three-contact hierarchical network, 25 for the five-contact network, and 50 for the ten-contact network. This feature of hierarchy increasing with the number of people in the hierarchy turns out to be important for measuring the network advantage of outsiders because it measures the volume of a strategic partner's network (Chapter 7), which strengthens the hierarchy association with performance (Burt, 1998a:Table 1).

Note that constraint increases with hierarchy and density such that evidence of density correlated with performance can be evidence of a hierarchy effect (as illustrated in Chapter 7). Constraint is high in the dense and hierarchical three-contact networks (93 and 84 points respectively). Constraint is 65 in the dense five-contact network, and 59 in the hierarchical network; even though density is only 40 in the hierarchical network. In the ten-contact networks, constraint is lower in the dense network than the hierarchical network (36 versus 41), and density is only 20 in the hierarchical network. Density and hierarchy are correlated, but distinct, components in network constraint.

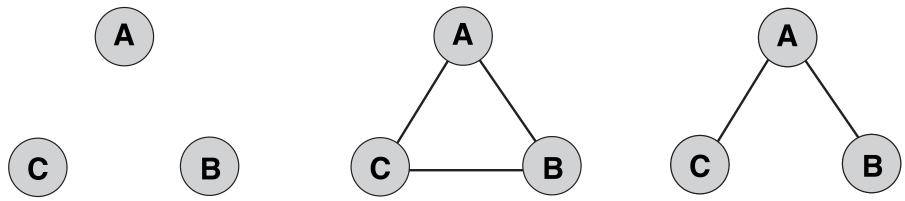
**Betweenness**

Freeman's (1977, 1979) betweenness index is an intuitively appealing measure. The index is a count of, or ratio of possible, monopoly opportunities for brokerage. If you know two disconnected people, then you have one opportunity to broker a connection between people. If you know four people disconnected from one another, then you

# Figure B2
# Network Size, Density, Hierarchy, and Constraint

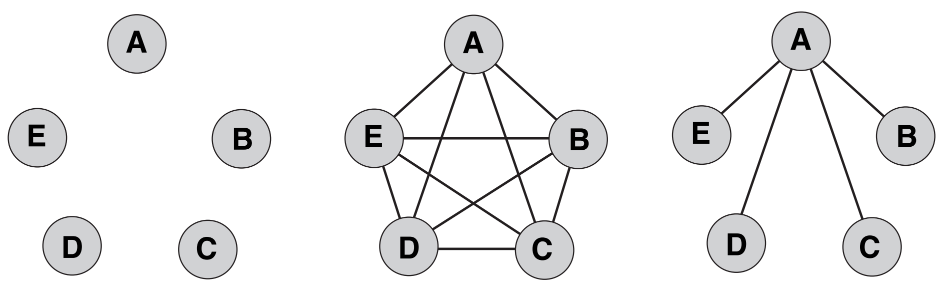
| | **Sparse Networks** | **Dense Networks** | **Hierarchical Networks** |
|---|---|---|---|

**Small Networks**




| | Sparse Networks | Dense Networks | Hierarchical Networks |
|---|---|---|---|
| size | 3 | 3 | 3 |
| density | 0 | 100 | 67 |
| hierarchy | 0 | 0 | 7 |
| constraint | 33 | 93 | 84 |
| from: | | | |
| A | 11 | 31 | 44 |
| B | 11 | 31 | 20 |
| C | 11 | 31 | 20 |

**Larger Networks**




| | Sparse Networks | Dense Networks | Hierarchical Networks |
|---|---|---|---|
| size | 5 | 5 | 5 |
| density | 0 | 100 | 40 |
| hierarchy | 0 | 0 | 25 |
| constraint | 20 | 65 | 59 |
| from: | | | |
| A | 4 | 13 | 36 |
| B | 4 | 13 | 6 |
| C | 4 | 13 | 6 |
| D | 4 | 13 | 6 |
| E | 4 | 13 | 6 |

**Still Larger Networks**

| | Sparse Networks | Dense Networks | Hierarchical Networks |
|---|---|---|---|
| size | 10 | 10 | 10 |
| density | 0 | 100 | 20 |
| hierarchy | 0 | 0 | 50 |
| constraint | 10 | 36 | 41 |

NOTE — Network scores other than size are multiplied by 100.

have six opportunities to broker a connection between people. Where a structural hole is defined to occur when two people are disconnected, then betweenness is a count of the structural holes to which a person has monopoly access.

In the network around the manager in Figure 2.1, for example, there are two indirect connections between contacts 1 and 3, one through contact 2 and the other through the manager. Since one of the two shortest paths between 1 and 3 goes through the manager, the manager has one half of an opportunity to broker the connection between 1 and 3. That is his only brokerage opportunity. Half of a brokerage opportunity divided by the three connections possible among three contacts means that the manager has the betweenness score of .17 reported in Figure 2.1.

There are N(N-1)/2 possible opportunities for brokerage between pairs of N contacts. The banker in Figure 2.1 is in a position to broker 23 connections between colleagues. Divided by the 28 connections possible among eight contacts, the banker has a betweenness score of .82, which says that he very nearly has a monopoly on all brokerage opportunities in the network (until the banker network is extended in Figure 2.2 to include indirect connections).

The index does not distinguish brokerage opportunities to which a person has direct access versus opportunities to which the person has indirect access. A brokerage opportunity between two of your close friends has the same weight in the index as a brokerage opportunity between two strangers far away in the network. Betweenness was introduced to describe centrality in small, five-person task groups in laboratory experiments (Freeman, 1977). Distant contacts were not an issue. When the index is applied to even modest-size populations such as the ones analyzed in this book, it can include opportunities in the index that are probably unrealistic to include — especially given the results in Chapters 3 and 4 showing that returns to brokerage are limited to opportunities among direct contacts (as was the case in the five-person task groups in which betweenness was initially measured). I am not describing a problem with the index. I am describing a caution in its use. Ideally, betweenness would be computed for the immediate network around a person, then for indirect contacts further removed.

**The Special Case of Isolates**

An isolate is a person who has no contacts in a network. When asked to name the people with whom they discuss their work, for example, isolates name no one and no one names them. In fact, virtually everyone at a managerial rank discusses their work with someone, but that someone could be a relative, a bartender at a favorite pub, a subordinate who did not make it into the study population, or some other colleague who did not make it into the study population. People can have a local circle of discussion partners at the same time that they are isolated within management. I use the term "isolate" in the specific network sense of the term.

Isolates pose a unique problem for certain measures of brokerage. Network size is unambiguously zero. A count of bridge relations is unambiguously zero. Betweenness as a count of brokered relations is unambiguously zero. However, measures of "average" relationship such as network density, or network constraint, are undefined. When N is zero, there is no average.

One way to proceed is to eliminate isolates by research design. For example, reporting relations are included among discussion relations in three study populations analyzed in this book. Since everyone reports to someone, no one is an isolate. Snowball sampling is another way to go. Two study populations in this book, the investment bankers and the analysts, are define by snowball sampling. A banker or analyst is only included in the study population when they cite or are cited by colleagues. The supply-chain population was defined by response to a network survey, augmented by snowball sampling to include contacts cited by two or more of the survey respondents. The respondents all named discussion partners and the additional contacts were only include if they were cited, so there would be no isolates even without reporting relations being included among the discussion relations. There were in fact a large number of supply-chain managers in the company who were isolated from other supply-chain managers (Burt, 2004), but the isolates were ignored for this analysis.

In general, populations defined by a non-network criterion — all students in a classroom, or all managers in an organization — will contain isolates. If isolates are included in the analysis, they have they have to be assigned density or constraint scores in keeping with the network concept being measured. If performance is being predicted from access to structural holes (as in Chapters 3 to 5), then each isolate is a closed network unto him or herself. Having no access to structural holes, isolates correspond to density and constraint scores of one. Assigning constraint scores of one to isolates is also consistent with the strong correlation between network size and the log of constraint (e.g., -.86 in the product-launch network, -.77 across the supply-chain managers, -.87 in the HR organization, -.70 across the bankers, -.76 across the analysts). On the other hand, if stability and trust are being predicted from social regulation within a closed network (as in Chapter 6), then each isolate, like a hermit, is free from social regulation. Neither constrained nor supported by a surrounding network of colleagues, isolates correspond in social regulation to density and constraint scores of zero. A moral here is that, if isolates are included in an analysis of social capital, it is wise to test effect robustness by adding to any prediction a dummy variable distinguishing isolates.

**Indirect Network Constraint**

The network around each of a manager's direct contacts poses some level of constraint and opportunity, on the contact directly, and on the manager indirectly through the contact. I measure indirect network constraint by aggregating constraint in networks around each of the manager's contacts,

$$IC_i = \sum_j \delta_{ij} C_j, \ i \neq j \qquad \text{(B3)}$$

where $C_j$ is direct network constraint on contact j (equation B1), and $\delta_{ij}$ is a weight for pooling contact networks. There is low indirect constraint on a manager connected to brokers (low $C_j$ scores average to a low $IC_i$ score). A manager subject to low indirect constraint is connected to colleagues whose networks are rich in brokerage opportunities, through whom the manager has indirect access to structural holes.

I tried measuring indirect network constraint as the arithmetic average across a manager's contacts ($\delta_{ij}$ = 1/N, where N is the number of the manager's contacts). This is the measure on the horizontal axis of Figure 4.5. I also tried the constraint on the manager's boss, under the assumption that the chain of command is the primary source for opportunities ($\delta_{ij}$ = 1 for manager's boss, 0 for all other contacts; Figure 4.6), and constraint on the manager's best-connected colleague, under the assumption that every contact need not be a source of opportunity, but you need at least one ($\delta_{ij}$ = 1 for the contact with the lowest network constraint, which means the largest, least redundant, network; 0 for all other contacts; Figure 4.7). These three aggregations yield the same result: strong zero-order association with performance and negligible partial association.

More sophisticated measures could be productive in other study populations, however, the simple arithmetic average is strongly correlated with more sophisticated measures in the manager populations studied here (also see Table 5.1 for tests with alternatives predicting performance, and Table E6 in Appendix E for correlations among the alternatives). For example, I computed indirect constraint as the weighted average of constraint with weights proportional to the constraint posed by each contact. The 1/n weight for alter j in the arithmetic mean is replaced with $c_{ij}/C$, where $c_{ij}$ is the level of constraint posed on ego i by alter j and C is the network constraint score for ego. This weighting emphasizes the networks around the direct contacts who most constrain ego. The weighted measure of indirect network constraint is correlated with the arithmetic mean .84, .78, .97, and .95 respectively for the analysts, bankers, managers, and the product-launch employees. I also tried weighting inverse to $c_{ij}$ to emphasize networks around the contacts most likely to be bridges. Again the weighted measure is strongly correlated with the arithmetic mean and yields the same associations with performance.

Average Within Ignores Across

Indirect constraint on ego measured by average constraint on alters has two properties to note for future research. First, it does not measure total indirect constraint. The total has two components: a component defined by connections within the network around each alter, and a component defined by connections across the networks around each alter. Averaging constraint scores across alters captures the first component plus some unknown portion of the second component (larger portion to the extent that the contacts for one alter are the same for other alters). The difference is illustrated in Figure 6.1b, where person 1 is part of a group separate from persons 2 and 3. Constraint is high within the network around each of ego's three contacts, but lower in the combined network across the three contacts.

I feel comfortable averaging within contact networks because within-neighbor network constraint measures the extent to which each neighbor has no direct access to structural holes and returns to brokerage are concentrated in direct access. Moreover, Model F in Table 5.1 shows that where there is brokerage spillover from neighbor networks — between industry networks — the results with total indirect constraint are similar to the results with indirect constraint averaged within contact networks. Total indirect constraint adds negligible prediction to indirect constraint averaged within contact networks. As a further check, I ran some tests with the product-launch and supply-chain managers. These are the populations most balkanized into subgroups, so connections across networks would be most likely in these populations to be valuable for coordinating information across the networks. I created a “total indirect network constraint” measure, TC, analogous to the measured used in Model F in Table 5.1 so I could add ln(TC) to the predictions in Chapter 3 to see whether something important had been lost by ignoring connections across neighbor networks. For each manager with M indirect contacts, I constructed an (M,M) matrix of connections among the indirect contacts, and computed network constraint from equation (B1) as if the manager had a connection with each of the M indirect contacts. The more interconnected the indirect contacts, the higher the total indirect network constraint on the manager. This measure of total indirect network constraint includes constraint from connections within the network around each of the manager’s contacts, plus constraint from connections across the networks. Across the product-launch managers, ln(TC) is correlated .45 with the log average indirect constraint measure used in the text and adds negligible prediction to the performance equations in Table 3.1 (-1.02 t-test for compensation, .75 for z-score annual evaluation). Across the supply-chain managers, ln(TC) is correlated .57 with the log average indirect constraint measure used in the text, and adds negligible prediction to the performance equations in Table 3.2 (-1.62 t-test for compensation, .01 for annual evaluation, and -1.79 for value of best idea). These results for total indirect network constraint are not equally negligible, but they are similar to average indirect constraint in being negligible, and all pale in comparison to the strong performance associations with direct network constraint. In short, there is no evidence of spillover missed by ignoring connections across neighbor networks to focus on structure within the networks.

Unproductive If Pushed Too Far

A second property to note for future research is that the average-alter measure can be unproductive in describing distant alters. Specifically, where each person in a

# Table B1
# Average Network Constraint on Increasingly Distant Contacts

| Maximum Path Distance to Averaged Alters | Standard Deviation in Indirect Constraint | Correlation between Direct and Indirect Constraint |
|---|---|---|
| 1 | 4.68 | .31 |
| 2 | 1.51 | -.09 |
| 3 | .92 | -.52 |
| 4 | .25 | -.76 |
| 5 | .11 | -1.00 |

population can reach every other person by some number of intermediaries, each person is indirectly constrained by N-1 alters (everyone else in the population) and indirect constraint averaged across all alters equals the population average excluding ego. In such a population, as alters further removed are included in alter averages, variance in indirect constraint decreases and the correlation between direct and indirect constraint approaches negative one.

Illustrative results are given above in Table B1 for the bankers in Chapter 4. The first column is the length of the path distance from ego to alters included in the network around ego, the second column is the standard deviation of indirect constraint, and the third column is the correlation between direct and indirect network constraint. The first row is the measure used in Chapters 3 and 4: Indirect constraint is the average network constraint on ego's direct contacts (alters one step distant from ego). The bottom row corresponds to the longest path distance, which in the banker population is 5 steps. As the network around ego expands to include more distant alters (down the rows), the indirect-constraint standard deviation decreases and the correlation between direct and indirect constraint approaches negative one. I am comfortable with alter averaging in this book because indirect constraint is limited to direct contacts and direct contacts are few relative to the number of people in each study population. In other populations, convergence could be an issue to consider. Ceteris paribus, the convergence to negative one will be faster in smaller, more-connected populations.

**Positional Measures**

Gathering data on relations between contacts complicates a routine survey (Table A2 in Appendix A). Inter-contact data are the most difficult and time-consuming survey network data to obtain. It is tempting to leave them out of the survey. There is the further incentive that useful work on social capital can be published that does not take into account ties between contacts. For example, Meyerson (1994) predicts executive salary in a selection of Swedish firms from a count of an executive's sociometric contacts outside the firm, and the proportion of the cited relations that are strong. Executives with stronger ties outside the firm enjoy higher salaries. Uzzi (1996) predicts failures among New York apparel contractors from distributions of business across contacts. Failure is less likely for contractors who have an exclusive business

---

*Uzzi's measures warrant comment because the work is so engaging but the measures can be understood to measure the extent to which relations are embedded in dense networks, whereupon their association with survival would be misinterpreted. The measures are computed from data that describe contractor sales of apparel components to manufacturers who assemble and market finished clothing (Uzzi, 1996: 696). Two measures Uzzi discusses as embeddedness are associated with contractor survival (a third, "social capital embeddedness," a dummy variable distinguishing contractors affiliated with a business group, has negligible association with survival). Uzzi (1996:686, italics in original) begins with an exclusive contractor-manufacturer tie; "The degree to which a firm uses embedded ties to link to its network is measured with the variable <u>first-order network coupling</u>." The variable is a Herfindahl index of concentration (sum of squared proportions) measuring the extent to which all of a contractor's sales are to a single manufacturer (Uzzi, 1996: Eq. 1). The other network variable (Uzzi, 1996:687) measures the average extent to which the contractor (focal firm) is the only contractor selling to its manufacturers (network partners); "<u>Second-order network coupling</u> measures the degree to which a focal firm's network partners maintain arm's length or embedded ties with their network partners." The variable is a Herfindahl index measuring the extent to which all of a manufacturer's purchases are from the contractor, averaged across the manufacturers to which the contractor sold goods (Uzzi, 1996: Eq. 3). The two Herfindahl indices are associated with contractor failure: failure is less likely to the extent that a contractor sells exclusively to a single manufacturer, and the manufacturers to which it sells only buy from that contractor. Thus my summary statement in the text that failure is less likely for contractors that have an exclusive business relationship with a preferred partner. Relations between manufacturers and between contractors are unknown, so there is no measure of the density and hierarchy of the network in which contractor-manufacturer relations were embedded. Uzzi's (1996) results are conceptually the same as, though substantively more detailed than, Meyerson's (1994) and Gabbay's (1997) results showing how important it is to span a structural hole with a strong, reliable relationship. The structural hole from which Uzzi's contractors and manufacturers profit is the division between people who make garment components and people who assemble the components into clothing.

relationship with a preferred partner.* Baum, Calabrese, and Silverman (2000) predict the performance of biotechnology start-ups from the diversity of the start-up's alliance network at founding. Alliances are sorted into nine categories according to the alliance partner (e.g., non-rival biotechnology firm, government laboratory, marketing company, etc.) and a start-up has a diverse network to the extent that it has alliances equally in all nine categories (Baum et al., 2000:276-277). This is diversity distinct from the image of spanning structural holes. For example, a start-up that has alliances with five nonrival biotechnology firms that do not have alliances with one another has a network that spans structural holes between the alliance partners (network constraint score would be 20, as illustrated in the first column of Figure B2), but the start-up would be coded by the Baum et al. measure as having zero network diversity because all five partners are the same kind of partner (non-rival biotechnology firm). Measurement complications notwithstanding, the brokerage concept is robust: Baum et al.'s network diversity measure does well predicting start-up differences in patent production (Baum et al. 2000:283). There has even been productive work at the radical extreme of measuring social capital without network data. Belliveau et al. (1996) infer relations from background similarities between people, as do Ancona and Caldwell (1992a) and Reagans and Zuckerman (2001) in their suggestive work on the external networks of teams (though Reagans, Zuckerman, and McEvily, 2004, build on Reagans and McEvily, 2003, to show the enhanced performance prediction that network data can add to predictions from team demography). Leana and Pil (2006) measure school social capital with teacher opinions about information sharing, trust, and shared goals within the school to report a correlation between high student test scores and positive teacher opinions. Perhaps most widely known is Coleman's (1988; 1990:590-597) analysis of social capital in which he infers network closure from family demography (children in families with two parents and few children have more closed networks), family mobility (children who have lived in the same neighborhood all their lives have more closed networks), and school (children in Catholic and other religious private high schools have more closed networks).

Inferences about social capital can be made in the absence of data on relations between contacts if data are available on the positions contacts hold in the broader social system beyond the network under analysis. People who occupy the same position in the broader social system are exposed to similar ideas, skills, and resources, and so are to some extent redundant contacts. They are redundant by structural equivalence (see Figure 8.5 in Chapter 8 for numerical illustration). Therefore, social capital can be inferred from the positions to which a person is connected. This is the foundation for positional measures of social capital (see Lin, Fu, and Hsung, 2001; van der Gaag, Snijders, and Flap, 2008, for methods review, and see Lin and Erickson, 2008, for recent applications to questions of access, trust, and inequality).

Positional measures are defined in two steps. The first step sorts potential contacts into categories according to their position in some broader social system. For example, contacts in different occupational statuses have access to different resources (Laumann, 1966; Lin, Ensel, and Vaughn, 1981; Lin and Dumin, 1986; Erickson, 1996), relations in broken homes are different from relations in intact families (Coleman, 1990), people long with the firm are different from new hires (Ancona and Caldwell, 1992a; Reagans and Zuckerman, 2001), contacts inside a firm are different from contacts outside (Meyerson, 1994), contacts in one division or function of a company are different from contacts in another (Ancona and Caldwell, 1992b; Hansen, 1999), contacts in one academic school of thought are different from contacts in another (Collins, 1998), alliances can be distinguished by kind of alliance partner (Baum, Calabrese, and Silverman, 2000; Powell, White, Koput, and Owen-Smith, 2005), or positions can be inferred from patterns of interaction (Walker, Kogut, and Shan, 1997). This first step for positional measurement is akin to the name generators in survey network data. Contacts are elicited for kinds of relationships by name generators and research design involves selecting an appropriate set of generators. Here, contacts are elicited for kinds of positions and research design involves selecting an appropriate set of positions.

The second step asks people about their connection with each position. Specific contacts are sometimes known, but often not. Nan Lin has been a leading advocate for positional measures of social capital, and offers an example survey item in which positions are defined by an assortment of occupations from high to low socioeconomic status (Lin, 2001:18): “Here is a list of jobs (show card). Would you please tell me if you happen to know someone (on a first-name basis) having each job?” If the respondent knows more than one contact in a category, he or she is asked to “think of the one person whom you have known the longest (or the person who comes to mind first)." When a respondent answers "yes," there are follow-up questions asking how long he or she has known the contact, the nature of the relationship with the contact, and so on. Often-used positional measures of social capital are the heterogeneity of contacts (number of occupations is akin to number of bridges assuming that contacts in different occupations are non-redundant, see Erickson, 1996; 2001, for two productive applications) and "upper reachability," which is the highest status in which the respondent has a personal contact.

This is not the place to offer a critique of positional measures, though a rigorous comparison of positional and network measures would be welcome. What can be said by way of summary critique is that positional measures have at least two virtues: An obvious one is that they are inexpensive: it is easy and quick for a survey respondent to provide the data. Second, they generate results. The primary disadvantage of positional measures is not a defect so much as a risk: positional measures are leveraged against the accuracy of the first step, the delineation of positions. For example, scholars outside the United States follow Lin's lead in using positional measures of social capital based on translated American occupational categories. Such use poses no problems as long as the American categories correspond to structurally equivalent contacts in the application country. However, if there is structural variation within a category (e.g., lawyers whose clients are major corporations might have access to resources different from those to which personal injury lawyers have access, or professors at a nationally prominent university differ in some ways from professors at a community college), then the assumption that contacts are redundant within positions is violated and the inference from positional contact to social capital is unclear. A strength of Walker, Kogut, and Shan's (1997) analysis is that they study structural equivalence to identify positions in terms of which their study population is stratified before computing positional measures.

It might seem that positional measures are hopelessly flawed by their lack of data on the relations between contacts. Positional measures cannot distinguish the columns in Figure B2; a sparse network is the same as a clique, and both are the same as a hierarchical network. However, turn the situation around and consider Figure B2 in light of positional distinctions. If a manager cited the three contacts at the top of the middle column and they were all three from the same segment of a company, then the manager indeed would have no social capital as is implied by the network constraint scores. But what if each contact worked in a different function, or a different division, or in a different company? Then the dense network among them would reinforce the strength of their bridge relationships with one another and the manager would be, in contrast to the high constraint score, rich in social capital. Recall the successful manager in Figure 2.2 whose closed network coordinated leaders in three divisions of the company.