Journal of Social Structure

JoSS Article: Volume 12

Choosing the ‘β’ Parameter When Using the Bonacich Power Measure

Simon Rodan
Department of Organization and Management, College of Business
San José State University, San Jose, CA 95192-0070
simon.rodan@sjsu.edu

Abstract

Bonacich (1987) suggested a family of centrality measures that provide a useful way of modeling questions of power and network constraint. However, the literature offers little guidance regarding the choice of β, the parameter which alters the way the measure accounts for the effect of having powerful contacts in ones network. In this paper I explore the way the choice of the β parameter affects the power indices the Bonacich measure generates. I consider three network properties which might affect the way the choice of β influences the Bonacich power indices. I find that in high density networks with few internal ‘chains’ and few pendants, the choice of β is largely immaterial. Conversely, in sparse networks, those with a high proportion of pendant nodes, or those with many chains, the value of β has a substantial effect on the power indices the measure generates. Next I consider whether power indices produced by interior values of β might be represented as a linear combination of “pure” vectors, those generated with values of β at either end of the parameter range and β = 0. I find that in the vast majority of cases a linear combination of “pure” vectors power is equivalent to using indices produced by interior values of β, making the choice of β largely moot. Finally, in the unlikely case that this disaggregation is inappropriate, I discuss the question of determining an appropriate value of β empirically.

Key Words

Bonacich power; social network analysis

Acknowledgements

I want to thank Steve Borgatti, Martin Kilduff, David Krackhardt, Martin Gargiulo, James Moody and three anonymous reviewers for their comments and suggestions.

Introduction

The original motivation for this paper came from a manuscript I read some months ago in which the authors noted that they were measuring power using the Bonacich measure and that they had used the default value of β without any explanation as to why. This struck me as a little odd since presumably the choice of β matters, particularly as the mechanisms the measure is supposed to capture are quite different for negative values of β than those for positive values. However, this gave me pause to wonder whether my assumption was correct and ask: how much does the choice of β really does matter in practice?

Bonacich (1987) suggested a formula that produced not one but a 'family' of centrality measures, depending on the chosen value of β. The term centrality is perhaps misleading in that although the measure can generate results similar (and even identical) to other centrality measures (specifically eigenvector- and degree centrality) its unique contribution is in its application to the question of network constraint and power. The measure reconciles Cook, Emerson, Gillmore and Yagamashi's (1983) theoretical predictions regarding the exploitation of dependency and notions of status with centrality in information flows (Freeman, 1977). Eigenvector centrality presupposes that it is useful to have powerful contacts, the assumption being that powerful contacts are willing to help when asked or that one benefits simply by association. Cook *et al.* (1983) on the other hand, looked at networks in which relations are in tension, where competition, bargaining, struggle, and negotiation predominate, networks in which powerful contacts are constraining and disadvantageous. The Bonacich measure can produce results by varying the value of a parameter, β, that mimic exactly those of Eigenvector centrality or that match Cook Emerson *et al.*'s prediction for power dependency relations.

The measure generates a value for each node, which values depend on the structure of the network and the value of β. The β parameter might be thought of as characterizing the climate or culture of the network. A positive value means the anyone who knows powerful others is made more powerful as a result, while a negative value makes those connected to powerful nodes weaker. It is suggestive of the nature of action; positive β values correspond to situations where status and 'reflected glory' matter. It also captures the notion than when one asks a well-connected influential contact for assistance, they are willing to help. A negative β, in contrast, seems to fit a more competitive "dog-eat-dog" world (Burt, 1992; White, 1992:10). Competition dominates; being linked to powerful people is constraining.

The formula by which the Bonacich measure derives the values of power for each node is:

$$c_i(\alpha,\beta) = \sum_j (\alpha - \beta c_j) R_{i,j} \qquad \textbf{1.}$$

This is Equation 3 from Bonacich's 1987 paper, where *c* is the derived nodal attribute (a centrality or power score which I will refer to here as a *power index*), R is an adjacency matrix, and α is a scaling factor. Solutions are suggested for values of β in the range ($-\kappa$,$+\kappa$) where $\kappa = 1/\upsilon^1$ and υ is the largest Eigenvalue in the solution of $\lambda x = Rx$. With β=0 the measure is equivalent to degree centrality; at β = κ it is equivalent to eigenvector centrality. Setting β = $-\kappa$ yields values consistent with Cook, Emerson *et al.*'s prediction regarding power (Cook, Emerson *et al.*, 1983; Emerson, 1962). However, beyond this the literature offers little guidance regarding how to choose an appropriate value of β. If indeed this is a family, all of whose members are equally useful, when is it appropriate to choose members other than β = κ, β = $-\kappa$ or 0? I begin with the antecedent question "does the choice of β really matter?" Looking at the way the measure's output responds to different values of β in the context of some relatively simple networks, I explore the circumstances in which the choice of β matters most.

(When) does the choice of β matter?

Consider first, a very simple network In the triad $j \leftrightarrow i \leftrightarrow k$, where j and k are separated by a structural hole, using a β of κ gives i a power index of 0.707 (or $1/\sqrt{2}$) while j and k each have 0.5. This models a situation in which i has two sources of potential aid, while j and k have only one. Nevertheless, the power indices for j and k are positive suggesting that they do benefit from knowing i. Changing β to $-\kappa$ drops j and k's power indices to -0.5 while i's is unchanged. Here j and k are worse off for knowing i, whose power works against them rather than for them, just as Emerson (1962) described. The vector of power indices generated with $\beta = \kappa$ is quite different from that generated when $\beta = -\kappa$.

To assess the extent to which the choice of β matters for the power indices, P_β, I define

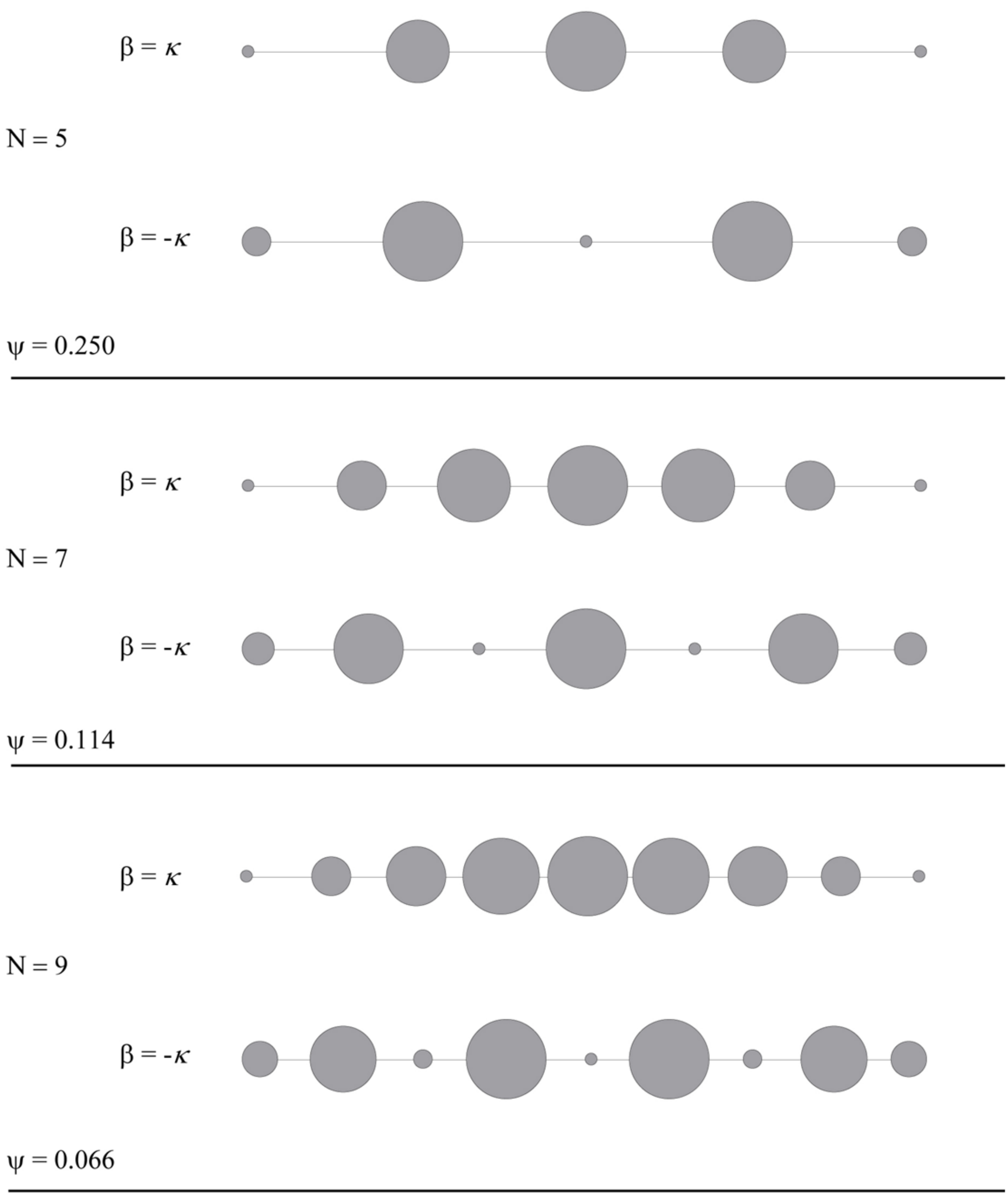
$$\psi = \mathrm{corr}(P_{-\kappa}, P_{+\kappa}) \qquad \textbf{2.}$$

where $P_{-\kappa}$ is P_β with β set to $-\kappa$ and $P_{+\kappa}$ is P_β with β set to $+\kappa$. When ψ is very close to 1, $P_{-\kappa}$ and $P_{+\kappa}$ are almost identical. Investigation showed that a correlation between any two vectors generated with interior values, ($\beta < |\kappa|$) was strictly greater than ψ. Thus if ψ is close to 1, the correlation between any two sets of power indices will be at least as high or higher and $P_{-\kappa} \approx P_{+\kappa} \approx P_\beta$ for any value of β.

Figure 1 shows how power is distributed in longer chains; the first, Cook, Emerson *et al.*'s Network 1c, has 5 nodes. As the chains lengthen, ψ declines. In fact, the relationship between chain length and ψ is quite precise: ψ is exactly proportional to 1 / N for chains with even numbers of nodes and to $1 / N^2$ for chains with odd numbers of nodes. In other words, the longer the chain the more important is the choice of β. Another feature of the network shown in Figure 1 is that with β set to $-\kappa$, powerful positions alternate along the chain beginning with the ends of the chain which are weak by virtue of their dependence on a unique contact. The two pendant nodes, one at each end 'seed' the network by providing their contacts with power and thus pendant are in a sense the well-spring of power variation in the chain. Networks without dependent actors will be much more evenly balanced in terms of power and control than ones with pendant nodes who are structurally in a weak position than thus provide power to those on whom they depend.

Figure 1

Power indices in different length chains for $\beta = +\kappa$ and $\beta = -\kappa$

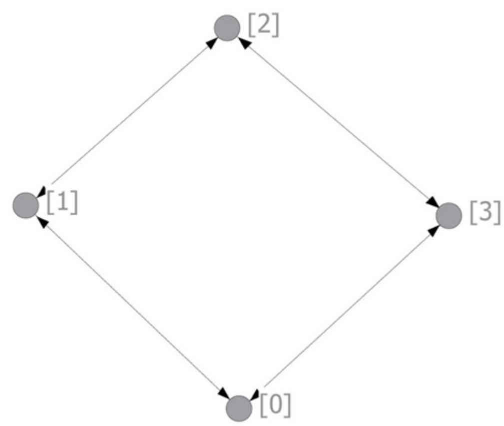


The size of the nodes represents their relative power; node sizes are illustrative and not to scale.

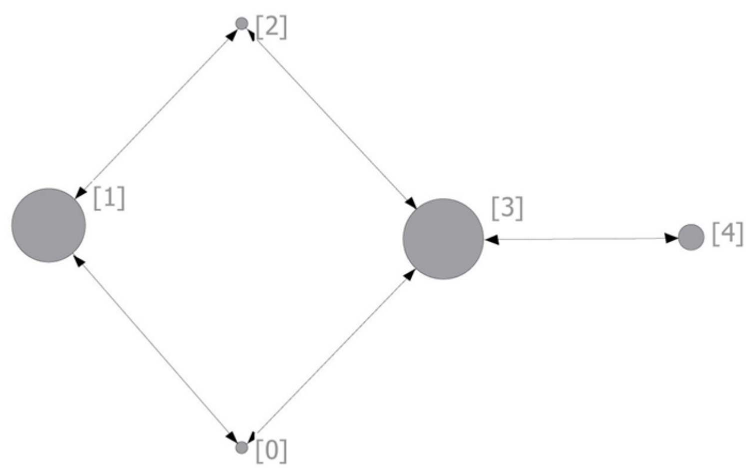
This is illustrated in **Figure 2**, which shows two similar networks and the relative power of the nodes, generated using β = -κ. The four nodes alone, (0-3, upper panel) are all equal in power. The addition of node 4 (lower panel) with a tie to 3 makes 3 powerful because he can exploit 4's dependence; 3's power makes 0 and 2 weak, and their weakness in turn empowers 1. Larger rings exhibit the same principle with power equality being disrupted by the addition of a single pendant node, even when the pendant is many arcs away. When power is modeled as competitive and network relations as adversarial, power and powerlessness alternate as one moves away from the source (node 4).

Figure 2

The effect of pendant nodes



ψ = 1



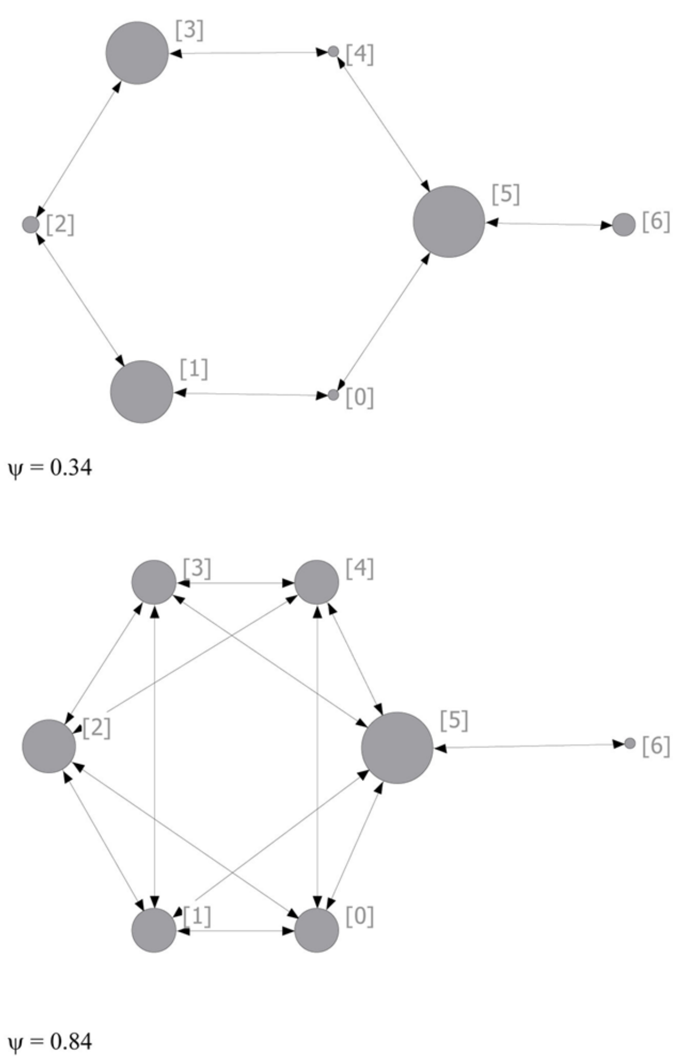
ψ = 0.4

The size of the nodes represents their relative power; node sizes are illustrative and not to scale.

This alternating power is suggestive of wave propagation, in particular standing waves such as those found in string or wind instruments (e.g., Backus, 1969) or church bells (Perrin, Charnely, and DePont, 1983; Rossing and Perrin, 1987). Extending the wave analogy, paths of differing lengths to the same point will cause the waves to interfere, constructively or destructively. In the case of constructive interference, small perturbations compound with positive reinforcement, while in the case of destructive interference, perturbations cancel each other out resulting in diminution or complete cancellation of the signal. This can be seen in **Figure 3**. The upper network shows the alternation of power ($\beta = -\kappa$) resulting from node 6's dependence on node 5. However, in the lower panel, each node is now connected not only to its erstwhile adjacent node but to the node adjacent to that which was previously its immediate neighbor. For example, node 3 is now connected not only to nodes 2 and 4 but to 1 and 5 as well. The pendant node aside, there is a much more even distribution of power among nodes in the lower panel.

Figure 3

The impact of 'bypass' ties

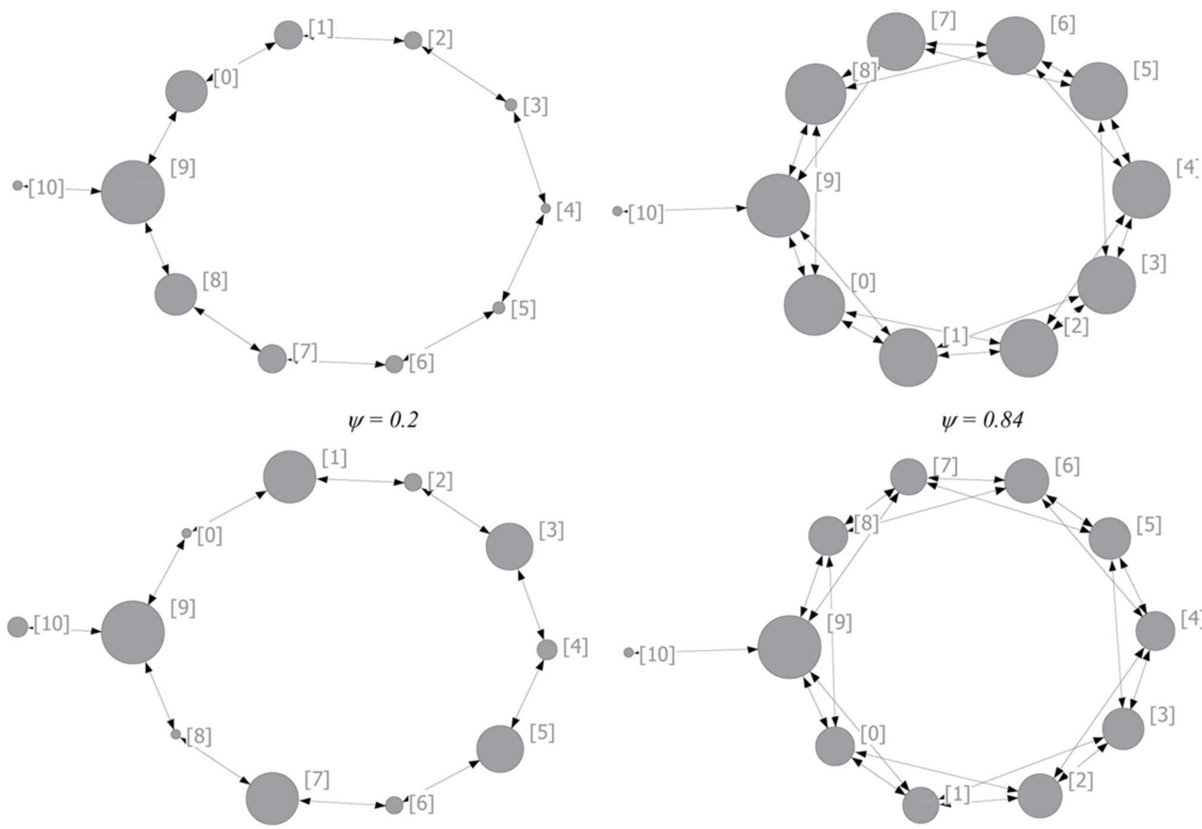


The size of the nodes represents their relative power; node sizes are illustrative and not to scale.

Multiple paths of differing lengths between nodes will diminish differences in power. The additional ties in the lower panel provide a ‘bypass’ for the transmission of the ‘signal’ from node 6 so that as it propagates away from node 5, node 3 receives a ‘negative’ signal directly from 5 and a positive signal from node 4 which has just inverted the negative signal it received from 5. The more dense the network the greater the number of paths of different lengths exist between any two nodes. Density should reduce differences in power (and similarly the propagation of status) through the network which is likely to be more evenly distributed as network density rises (**Figure 4**).

Figure 4

The Effect of Bypass Ties (Density) on the Transmission of Status (κ = 1, top) and Power (κ = -1, bottom) in Larger Rings



The above discussion provides the rationale for the three network properties that are the subject of the next part of this study; average distance between nodes, the density of the network and some measure of pendant nodes. These three network properties are likely to make the choice of β more or less important. In networks that are dense, have few chains or bridges (that is they have low average distance between nodes) with and have few pendants, the choice of β is unlikely to matter; conversely in sparse networks that have many bridges, long chains and numerous pendants, it will be critical.

Extending the investigation to less stylized structures

So far, only very simple and regular structures have been considered. However, the networks we tend to study are often larger, usually far less regular and far more complex. To better understand the way a researcher’s choice of β affects the power indices in practice, I applied the measure using different values of β to four of the empirical network datasets that come with the UCINET software (Borgatti, Everett, and Freeman, 2002); the society women from "Deep South", international trade flows, the network of relationships around Cosimo de Medici in Renaissance Florence, and the bank wiring room network from the Hawthorn studies. In the last three the multiplex networks were combined into a single binary graph by adding the superimposed ties.[2]

Table 1 shows the correlations between vectors of power indices generated using different values of β. What is striking about the top three tables is how little difference there is between the power indices, even when comparing those generated with $\beta = \kappa(P_{+\kappa})$ and $\beta = -\kappa\ (P_{-\kappa})$. For example, in both the Deep South data and the international trade flow data, the two sets of indices are correlated at over 0.98. This seems to call into question the usefulness of the Bonacich measure; if in real networks it produces almost identical results regardless of the value of β, and only generates different results for highly stylized and arguably unrealistic networks, is it a useful measure?

Table 1 *Correlation between vectors of power indices at different levels of β*

Deep South — N = 18, Density = 0.908, Av dist = 1.03

	Mean	Stdev	1	2	3	4
1 beta = -1.000000	0.601	0.076				
2 beta = -0.500000	0.605	0.073	0.999			
3 beta = 0.000000	0.609	0.071	0.998	0.999		
4 beta = 0.500000	0.613	0.069	0.995	0.998	0.999	
5 beta = 1.000000	0.617	0.067	0.990	0.994	0.997	0.999

International Trade — N = 24, Density = 0.810, Av dist = 1.14

	Mean	Stdev	1	2	3	4
1 beta = -1.000000	0.547	0.142				
2 beta = -0.500000	0.562	0.134	0.999			
3 beta = 0.000000	0.575	0.127	0.996	0.999		
4 beta = 0.500000	0.587	0.121	0.991	0.996	0.999	
5 beta = 1.000000	0.597	0.116	0.983	0.990	0.995	0.999

Bank Wiring Room — N = 14, Density = 0.599, Av dist = 1.32

	Mean	Stdev	1	2	3	4
1 beta = -1.000000	0.443	0.156				
2 beta = -0.500000	0.474	0.152	0.992			
3 beta = 0.000000	0.499	0.151	0.974	0.994		
4 beta = 0.500000	0.519	0.153	0.948	0.980	0.995	
5 beta = 1.000000	0.536	0.158	0.915	0.956	0.981	0.995

Renaissance Florence — N = 16, Density = 0.225, Av dist = 1.71

	Mean	Stdev	1	2	3	4
1 beta = -1.000000	0.192	0.264				
2 beta = -0.500000	0.327	0.231	0.965			
3 beta = 0.000000	0.422	0.223	0.899	0.981		
4 beta = 0.500000	0.504	0.230	0.820	0.936	0.985	
5 beta = 1.000000	0.609	0.282	0.686	0.820	0.897	0.951

The last network, however, provides an indication that it is. Padgett's networks from Renaissance Florence show a markedly lower correlation (0.686) between the power index vector generated with $\beta = -\kappa$ and that produced at $+\kappa$. Two things distinguish this last network from the other three; first the average distance between nodes for this network is higher than the others and second, it is considerably sparser. With longer chains and unique paths along which power oscillations may develop, very different power indices are generated when negative values of β are used than those produced using a positive value of β.

Chains, density and pendants

To explore the proposition that chains, density and pendants play an important role in determining whether the choice of β matters, the Bonacich measure was applied to a number of randomly generated graphs. Four network generating algorithms were used. One creates ties between nodes at random. The second creates a single component, rewires this randomly a few times, and then adds new ties. The third creates a number of clusters of a certain density (using the second algorithm just described) and then rewires the intra-cluster ties to join one cluster to another until the ratio of tie density between clusters to that within clusters reaches a specified level. These networks are similar in topology to the Watts "Cavemen" graphs (Watts, 1999). The algorithm generates networks with higher variance in closeness and eigenvector centrality than the random graph generator. The last algorithm creates networks based on network 1e from Cook, Emerson *et al.* (1983) to which some ties are added to increase the network density. At low densities, this last algorithm generates networks with a high proportion of pendant nodes. Each network parameter—such as size, density, number of re-wirings, number of added nodes, and the number of additional ties—was set stochastically.

Three network measures were calculated. A chain-link node was defined as any node with just two alters. The proportion of chain-link nodes was this number divided by the network size. The proportion of pendant nodes was the number of pendants divided by network size. Density was calculated in the usual way. A visual inspection of the data from the simulations showed a positive relationship between ψ (Equation 2) and density. The relationship between ψ and the proportion of chain-link nodes between nodes was strongly negative. The relationship between ψ and the proportion of pendants was less clear, with substantial variability in ψ for low proportions of pendant nodes and low variability for high proportions. Despite this heteroscedasticity, there was a fairly clear negative trend, with ψ declining on average as the proportion of pendants increased.

Table 2

Summary statistics

	Mean	Stdev	1	2	3	4
1 ψ	0.573	0.309				
2 Size	19.110	6.482	-0.178 †			
3 Proportion of chain-link nodes	0.144	0.057	-0.431 ***	0.121		
4 Density	0.246	0.174	0.681 ***	-0.427 ***	-0.353 ***	
5 Proportion of pendant nodes	0.217	0.254	-0.561 ***	0.055	-0.168 †	-0.398 ***

Pair-wise correlations between ψ, network size, density, proportion of chain-link nodes and the proportion of pendant nodes are shown in **Table 2**. ψ is negatively related to the proportion of chain-link nodes and the proportion of pendant nodes and positively related to density. To ascertain that all three network properties have an independent effect on the value of ψ, I regressed ψ against network size, density, the proportion of chain-link nodes and the proportion of pendant nodes. Because of the heteroscedasticity in the relationship between ψ and the proportion of pendant nodes, the Stata *rreg* procedure was used for the estimation.

Table 3

Regression of ψ with density, proportion of chain-link nodes and proportion of pendant nodes

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.7129 ***	0.781 ***	0.1219	0.8373 ***	0.5546 ***
	(0.087)	(0.078)	(0.089)	(0.071)	(0.076)
Size	-0.007 †	-0.0057	0.0058 †	-0.0047	0.0032
	(0.004)	(0.004)	(0.003)	(0.003)	(0.002)
Proportion of chain-link nodes		-0.7486 ***			-0.7212 ***
		(0.13)			(0.087)
Density			1.3464 ***		0.6417 ***
			(0.149)		(0.124)
Proportion of pendant nodes				-0.7102 ***	-0.5741 ***
				(0.101)	(0.073)
Adj R Sq	0.0123	0.168	0.443	0.341	0.626

N =100, * $p < 0.05$, * $p < 0.01$, *** $p < 0.001$

After a model (Model 1) in which network size is entered alone, Model 2 shows the effect of proportion of chain-link nodes alone, Model 3 the effect of density and Model 4 that for the proportion of pendants. As expected, ψ declines as the proportion of chain-link nodes and the proportion of pendants increase, and increases with density. Size, for all intents and purposes, has no effect on ψ. Model 5 show all variables entered together. The coefficients for density and the proportion of pendant nodes are reduced suggesting that some of the variance accounted for by pendants when used alone is now taken up by density and vice versa. This is to be expected since the two are correlated (-0.4); nevertheless, independent effects for all three network properties remain.

Choosing a better β—interior values

Thus far, I have shown only that for dense networks, networks with few chains, or networks with a small proportion of pendant nodes, $P_{+\kappa}$ and $P_{-\kappa}$ are highly correlated and thus the value of β for all practical purposes is immaterial. On the other hand, if one has a sparse network or one with long chains or one with a high proportion of pendants, then the choice of β will make a considerable difference. In such cases, how should one go about choosing a value for β? A first cut seems fairly straight forward: if one assumes power is being used to exert control over others, a context characterized by competition and bartering, then a negative value is called for. On the other hand if one thinks power is used cooperatively, then use a positive value. However, if this truly is a “family” of measures when might one want to use an interior value of β, such as $-\kappa$ / 3.14 ? Unfortunately, nothing in the literature helps answer this question. One answer might be that interior values represent different mixes of cooperative and competitive behavior. For example, if actors behaved cooperatively 25% of the time and competitively the rest, one might choose $\beta = -\kappa / 2$.[3] Indeed, if β is related to the mix of 'pure' strategies, competition and cooperation, one might think of β not so much as a parameter to be chosen theoretically ex-ante, but rather as a variable carrying information about the network that needs to be uncovered empirically. I return to the question of empirically estimating β in the final part of this paper.

If one starts by considering the two mechanisms that the Bonacich measure captures, adversarial and cooperative power, as 'pure' strategies then perhaps the effects that intermediary values of β best represent might equally well be captured as a linear combination of two sets of power indices, one with $\kappa = -\beta$ representing struggle and competition and one with $\kappa = \beta$ to capture cooperation and status. If intermediary values of β are representative of different mixes of the two pure strategies, mixed strategies might be disaggregated into two pure components.

Suppose, by way of example, one is interested in understanding how network structure is related to salary, and one hypothesis is that power derived from ones position in the network may play a role. This may be because actors in the network bargain with each other and behave competitively but status and cooperation might also be present. Assuming that $P_{-\kappa}$ characterizes competitive power and $P_{+\kappa}$ power deriving from status and cooperation, an empirical approach to determining which was present would be to regress salary against $P_{-\kappa}$ and $P_{+\kappa}$. The first would pick up competition and power exploited in an adversarial manner, the second status and influence deployed cooperatively. With such an approach, one wouldn't need a family of measures but just two; one at $\beta = -\kappa$ to tap competition and the adversarial use of power, the other $\beta = +\kappa$ (or eigenvector centrality) to capture status and cooperation. For this to be a valid modeling approach, intermediate values of β must be able to be represented as a weighted average of the two 'pure strategies', $+\kappa$ and $-\kappa$.

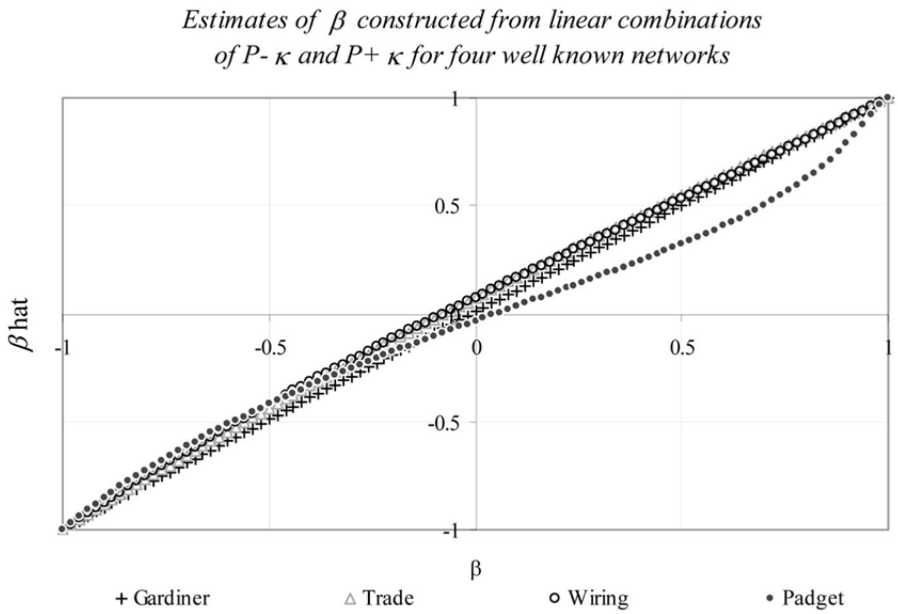
To test this I turned first to the four UCINET supplied networks. For each, I generated a vector of power indices, P_β, using randomly chosen values of β. Next I created two 'pure strategy' power index vectors, $P_{-\kappa}$ for competition, and $P_{+\kappa}$ for cooperation and status by association, the two ends of the β parameter range. I then regressed P_β against $P_{-\kappa}$ and $P_{+\kappa}$, first constraining the proportions of each vector and setting the intercept to 0 (Equation 2) creating a weighted average of the two pure strategies, and then without constraint (Equation 3), a more general linear combination.

$$P_\beta = \alpha_1 P_{-\kappa} + (1-\alpha_1) P_{+\kappa} + \varepsilon \qquad \textbf{3.}$$

$$P_\beta = \alpha_0 + \alpha_1 P_{-\kappa} + \alpha_2 P_{+\kappa} + \varepsilon \qquad \textbf{4.}$$

If P_β is accurately represented as a linear combination of $P_{-\kappa}$ and $P_{+\kappa}$ then values of α_1 and β should be related by $\beta = 1-\alpha_1*2$ in the constrained case and $\beta = \alpha_2-\alpha_1$ in the unconstrained case. The two estimates of β, constrained and unconstrained, are highly correlated (0.99). **Figure 5** shows the unconstrained estimates of β plotted against the actual values.[4]

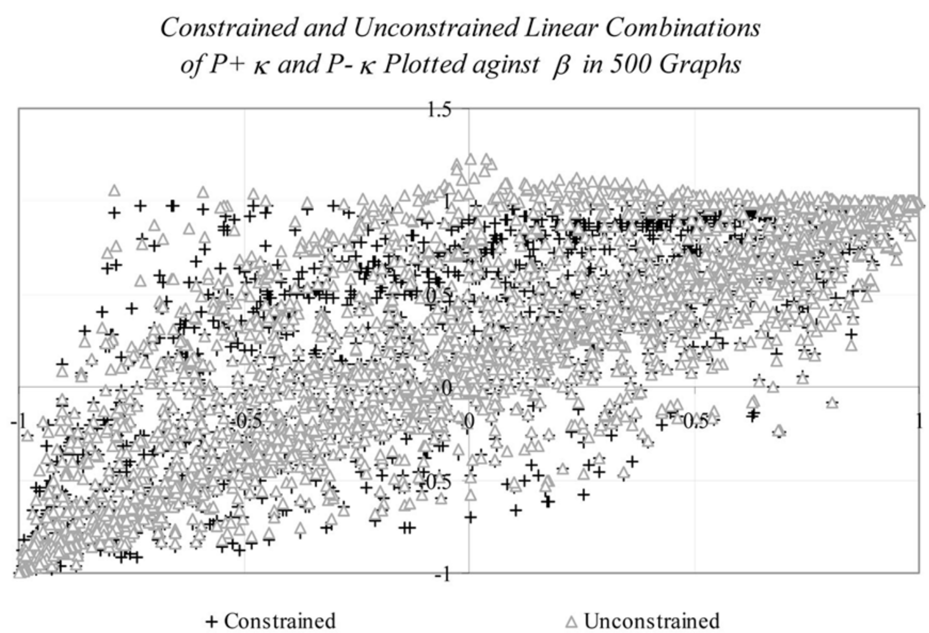
Figure 5



While for the Deep South network, β is well represented as a linear combination of $P_{-\kappa}$ and $P_{+\kappa}$ the international trade and the bank wiring room networks produce estimates that are slightly high while in the case of renaissance Florence interior values of β between about $-\kappa / 8$ and κ are slightly inflated and too low for $-\kappa$ to $-\kappa / 8$.

The problem is still more apparent in **Figure 6**, which repeats the same process this time on a large number of synthetic graphs. These graphs were constructed from the same four network generating algorithms described earlier under 'Chains, Density and Pendants.' The figure shows clearly that in a large number of cases, neither a weighted average of $P_{-\kappa}$ and $P_{+\kappa}$ nor an unconstrained best fitting linear combination of the two pure vectors is equivalent or even particularly close to interior values of β.

Figure 6

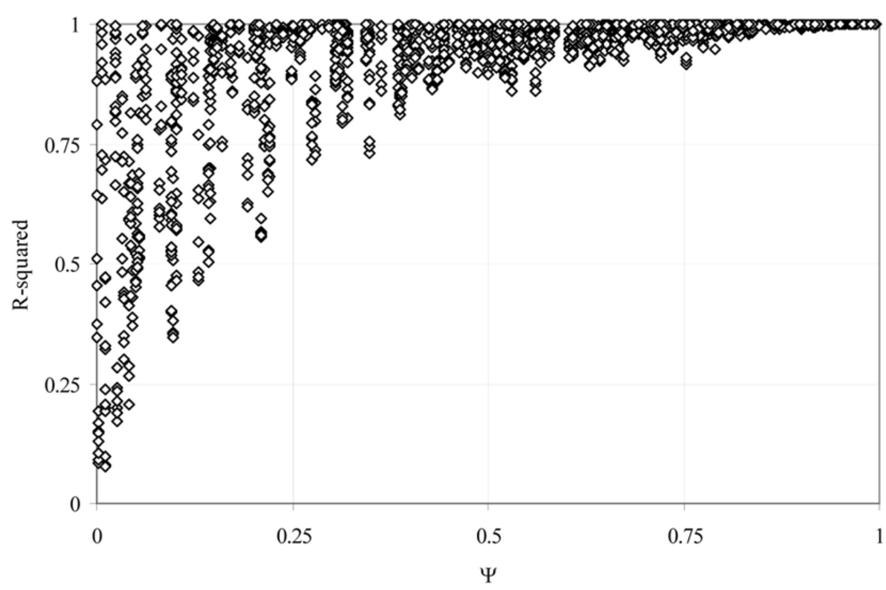


However, while the proportions of $P_{-\kappa}$ and $P_{+\kappa}$ cannot always be used to reliably infer a corresponding value of β, this does not mean that P_β will provide a measure which taps variance that $P_{-\kappa}$ and $P_{+\kappa}$ together could not do equally well. That is, while a point in **Figure 6** may not lie on the leading diagonal and the proportions of $P_{-\kappa}$ and $P_{+\kappa}$ don't provide a way of inferring β, that is not to say P_β cannot be accurately constructed from combination of $P_{-\kappa}$ and $P_{+\kappa}$. To answer this one must look at the proportion of variance in P_β explained by $P_{-\kappa}$ and $P_{+\kappa}$, in other words the R-squared values of the unconstrained regressions. If the R-squared is 1, there is no variance in an empirical phenomenon that P_β will capture that would not equally well be modeled using $P_{-\kappa}$ and $P_{+\kappa}$

For most of the synthetic networks networks, weighted values of $P_{+\kappa}$ and $P_{-\kappa}$ afforded an almost perfect substitute for interior values of β with R-squareds of over 99%. This suggests that generally, using these two vectors in a regression will capture exactly the same variance as using a single power index vector with an interior value; in which case, there is no particular reason to agonize over the choice of an interior value, and the question "What's the right value of β to use?" is moot.

However, in about 37% of the generated networks the variance in P_β explained by $P_{+\kappa}$ and $P_{-\kappa}$ was less than 0.9, and in 5% of cases, less than 0.6. It is therefore not inconceivable that, depending on the network topology, assuming $P_{+\kappa}$ and $P_{-\kappa}$ will account for every member of the Bonacich family of measures would be a mistake. The networks for which $P_{-\kappa}$ and $P_{+\kappa}$ are least correlated, those where the effects of competitive power are most distinct from status and cooperative influence, are also those for which the weighted $P_{-\kappa}$ and $P_{+\kappa}$ disaggregation often works least well.

Figure 7 *R-squared values from regressions of P_β on $P_{+\kappa}$ and $P_{-\kappa}$ plotted against ψ, the correlation of $P_{+\kappa}$ and $P_{-\kappa}$*



In **Figure 7**, the lowest R-squared values from the regressions of P_β against $P_{-\kappa}$ and $P_{+\kappa}$ appear on the left side of the figure where the correlation between $P_{-\kappa}$ and $P_{+\kappa}$ is fairly low, in other words where the effect of the different mechanisms is most pronounced.

Because at its mid-point, the Bonacich formulation produces a measure of degree centrality, this fits the bill as a third ‘pure’ vector. In networks for which power indices created with interior values of β are not well represented as combinations of competitive power ($P_{-\kappa}$) and eigenvector centrality ($P_{+\kappa}$), might they be better captured as combinations of three pure cases, $P_{-\kappa}$, $P_{+\kappa}$ and degree centrality (P_0)? The next step was to see whether a three vector solution produced a better fit in the cases in which two pure vectors did not.

Figure 8

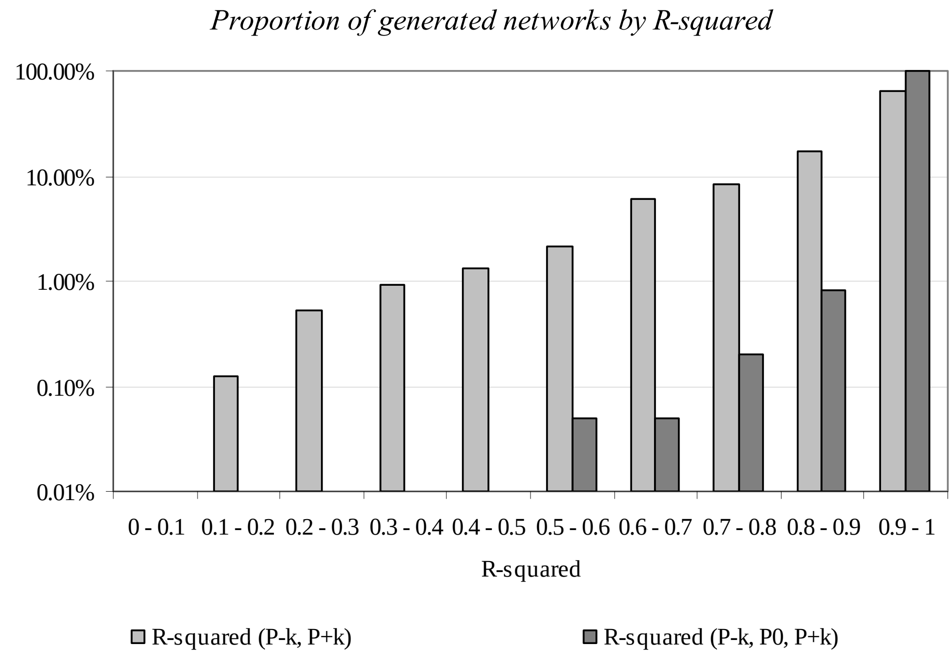
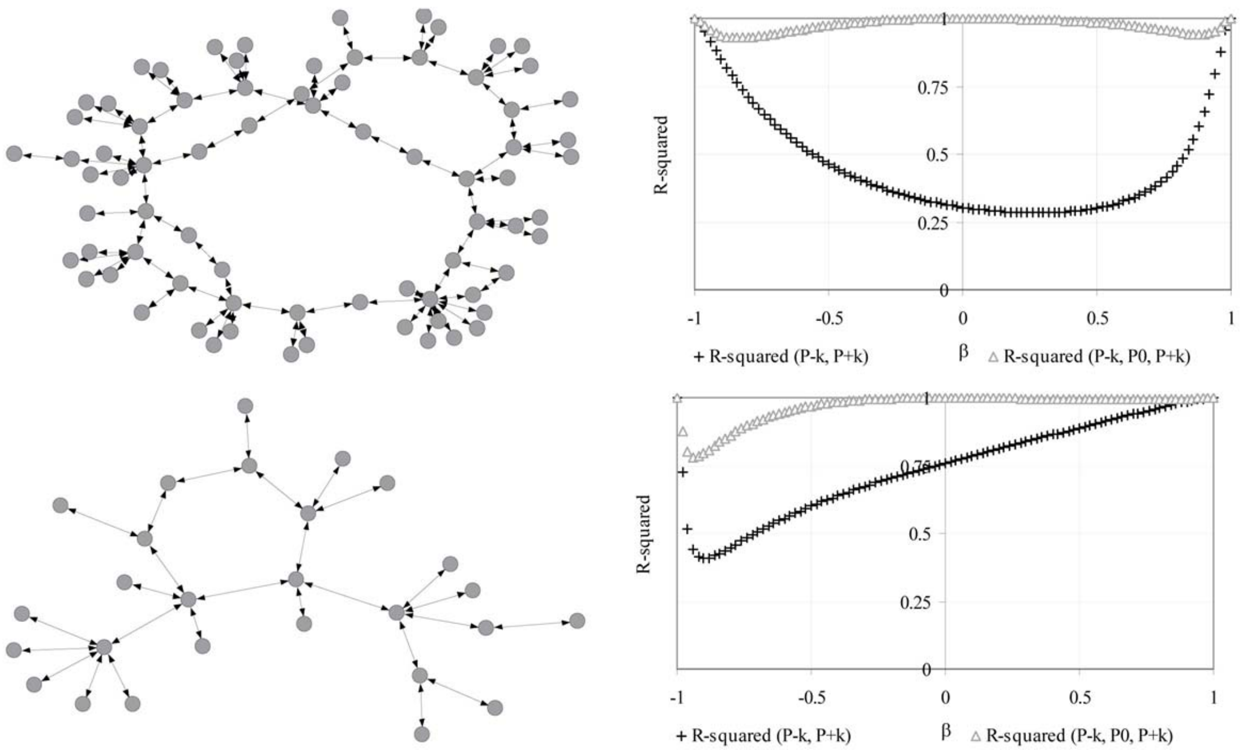


Figure 8 shows the distribution of levels of the explanatory power for regressions of P_β with $P_{-\kappa}$ and $P_{+\kappa}$ and P_β with $P_{-\kappa}$, P_0 and $P_{+\kappa}$. Using all three vectors reduces the instances of poor fit fairly dramatically, with 99% of cases showing explained variance between 0.9 and 1. Thus, in most cases, using $P_{-\kappa}$, P_0 and $P_{+\kappa}$ would essentially be no different from picking a particular interior value of β. The approach has the decided advantage of offering an intuitive interpretation, and making the ex-ante choice or empirical determination of β unnecessary. This suggests a simple modeling strategy; if $P_{-\kappa}$ and $P_{+\kappa}$ don’t adequately capture power indices at interior values of β, $P_{-\kappa}$, P_0 and $P_{+\kappa}$ almost certainly will. Instead of including in a model one vector of power indices (P_β), one would include two or three ‘pure mechanism’ vectors.

For any given network assessing whether two or three pure mechanism vectors are needed is a fairly straight forward task; one calculates the three vectors $P_{+\kappa}$, P_0 and $P_{-\kappa}$ and then P_β for a range of intermediary values. The P_β vectors are then regressed against $P_{+\kappa}$ and $P_{-\kappa}$ or $P_{+\kappa}$, P_0 and $P_{-\kappa}$. An inspection of the R-squared values at different values of β will be an indication as to whether to use $P_{-\kappa}$ and $P_{+\kappa}$ or $P_{-\kappa}$, P_0 and $P_{+\kappa}$. While simple in concept, in practice this is quite laborious, so I created a small command-line utility to do this which I describe in the appendix. Briefly, for a given network the utility generates data from which the plots in **Figure 9** were created.

Figure 9

Two networks and their respective plots of R-squared values from regressions of P_β against $P_{-\kappa}$ and $P_{+\kappa}$, and $P_{+\kappa}$ P_0, and $P_{+\kappa}$ plotted against β for a single network



The figure shows two carefully chosen networks and the degree to which two or three pure strategy vectors completely mimic a single vector generated with an interior value of β. The networks were selected to illustrate cases in which the disaggregation just described would be inappropriate. On the left are the networks, on the right the plots generated by the diagnostic utility. The upper network shows very poor performance for a two vector approach with over 69% of the range falling below 50% explained variance. The three vector approach is a considerable improvement with the lowest level of explained variance being 92%. Here a three vector approach seems warranted. The lower network presents a problem however because if the network is predominantly competitive, neither the two or three pure vector approaches work very well. The two vector solution falls below 50% explained variance over 24% of the range between $\beta = -\kappa$ and $\beta = \kappa$, while the three vector solution explained less than 90% of the variance over 27% of the range. For this network neither the two nor the three pure vector substitution would capture the same variance as some members of the Bonacich family at intermediary β values. Here one might choose to use P_β with an appropriately chosen value of β rather than settling for the approximation represented by the combination of the three pure strategy vectors. Thus we are forced to return to the problem of the choice of β that a two or three pure vector approach would have enabled us to circumvent.

β as an empirical variable

When estimating the effects of structure on some outcome in networks in which the value of β matters, and for which using $P_{-\kappa}$ and $P_{+\kappa}$ or even $P_{-\kappa}$, P_0 and $P_{+\kappa}$ is problematic, rather than choosing β ex-ante, one might instead search empirically for the particular value of β that is most closely associated with the outcome variable of interest. Using such a search based approach, the choice of β is now framed as an empirical question, namely: “what value of β generates a model of transmission through the network that best explains some observed distribution of characteristics among the nodes?” Finding a value of β which best characterizes some property of interest at the nodes is easily achieved by generating a series of power index vectors and correlating each with this particular property. The value of β that generates P_β which correlates most closely with the dependant variable tells us something, albeit in a rather general sense, about the characteristics of actors in the network in aggregate.

To generate such a set of power indices, I created another small command line utility[5] that generates a sequence of power index vectors for a range of values of β between $-\kappa$ and $+\kappa$. Again, the idea is straightforward but the computational task is onerous; the utility makes the search for an optimal β fairly simple. The utility can also generate the correlations between a range of power index vectors and a vector of nodal properties and write out the P_β vector calculated at the optimal value of β, that is the value yields a P_β that most closely correlates with the input vector for the network being modeled. The command format and an example are shown in the appendix.

However, there is still a problem in that one is now left with the question of interpretation; what exactly does the found value of β, one that is not a linear combination of $P_{+\kappa}$, P_0 and $P_{-\kappa}$, really mean? When β is well represented by some mix of competition ($P_{-\kappa}$) cooperation or status ($P_{+\kappa}$), and possibly prestige (degree centrality, P_0) interpretation is straightforward; but in such cases a better approach would be to use $P_{-\kappa}$, P_0 and $P_{+\kappa}$. Since an empirical search for an optimal β is needed only when P_β is not a linear combination of $P_{-\kappa}$, P_0 and $P_{+\kappa}$ what the discovered value of β tells us is far from clear; there is now no obvious theoretical mechanism with which it can be associated. This leads me to conclude that intermediate values of β should perhaps best be avoided, first since the cases in which P_β cannot be represented by $P_{-\kappa}$, P_0 and $P_{+\kappa}$ are likely to be fairly rare, and second, when this is the case, the impetration of a value of β found through an empirical search is not conducive to straightforward interpretation.

Conclusion

In answer to the question, “Does the choice of β matter?” my conclusion is in most cases it doesn’t. For some networks, particularly those that are dense, have small average distances between pairs of nodes, or have no pendant nodes, the power indices generated by the Bonacich measure will be identical or at least very highly correlated and the choice of β is moot. For networks which P_β vectors are not highly correlated, the choice of β can still be avoided by using either the pair of pure vectors $P_{-\kappa}$ and $P_{+\kappa}$ or if these together are inadequate $P_{-\kappa}$, P_0 and $P_{+\kappa}$. In what I imagine are likely to be fairly rare cases, networks give rise to power indices created with interior values of β that are not amenable to disaggregation into three pure components. Here one has the option of empirically choosing an appropriate value of β, though it is not clear what such a vector represents in terms of the modeling of real social phenomena.

---

[1] In the description that follows I will use $\kappa = 0.999/\upsilon$ rather than $1/\upsilon$ since the matrix $(I-1/\upsilon\ R)$ would be singular. $(I-\kappa R)$ however is not. Moreover, Borgatti has observed (private communication) that there are many regimes each defined by the reciprocals of the eigenvalue for each eigenvector. Since typically the largest eigenvalue and its corresponding eigenvector, are used (for example UCINET), in the discussion that follows my investigation refers only to the regime defined by the reciprocals of this principal eigenvector.

[2] A logical ‘OR’ operation.

[3] A quarter of the way from $-\kappa$ to κ.

[4] A similar exercise was undertaken using three vectors, $P-\kappa$, P_0 and $P+\kappa$ (power, degree centrality and eigenvector centrality). The results were very similar although the quadratic form of the relationships between the regression coefficients and the actual value of β made reconstruction of β problematic.

[5] A Linux and a Windows version of the utility can be downloaded from http://strategy.sjsu.edu/rodan_s/research/bin/BFMG.tar.gz [April 2011]

References

Backus, J. (1969). *The Acoustical Foundations of Music*. New York, NY: W.W. Norton.

Bonacich, P. (1987). "Power and Centrality: A Family of Measures Power and Centrality: A Family of Measures." *The American Journal of Sociology* 92, 5: 1170-1182.

Borgatti, S.P., M.G. Everett, and L.C. Freeman (2002). UCINET for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Burt, R.S. (1992). *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Cook, K.S., R.M. Emerson, M.R. Gillmore, and T. Yagamashi (1983). "The Distribution of Power in Exchange Networks: Theory and Experimental Results." *American Journal of Sociology* 89, 2: 275-305.

Emerson, R.M. (1962). "Power Dependency Relations." *American Sociological Review* 27: 31-40.

Freeman, L.C. (1977). "A Set of Measures of Centrality Based on Betweenness." *Sociometry* 40, 1: 35-41.

Perrin, R., T. Charnely, J. DePont (1983). "Normal Modes of the Modern English Church Bell." *Journal of Sound and Vibration* 91, 1: 29-49.

Rossing, T.D., and R. Perrin (1987). "Vibration of Bells." *Applied Accoustics* 20: 41-70

Watts, D.J. (1999). "Networks, Dynamics, and the Small-World Phenomenon." *The American Journal of Sociology* 105, 2: 493-527.

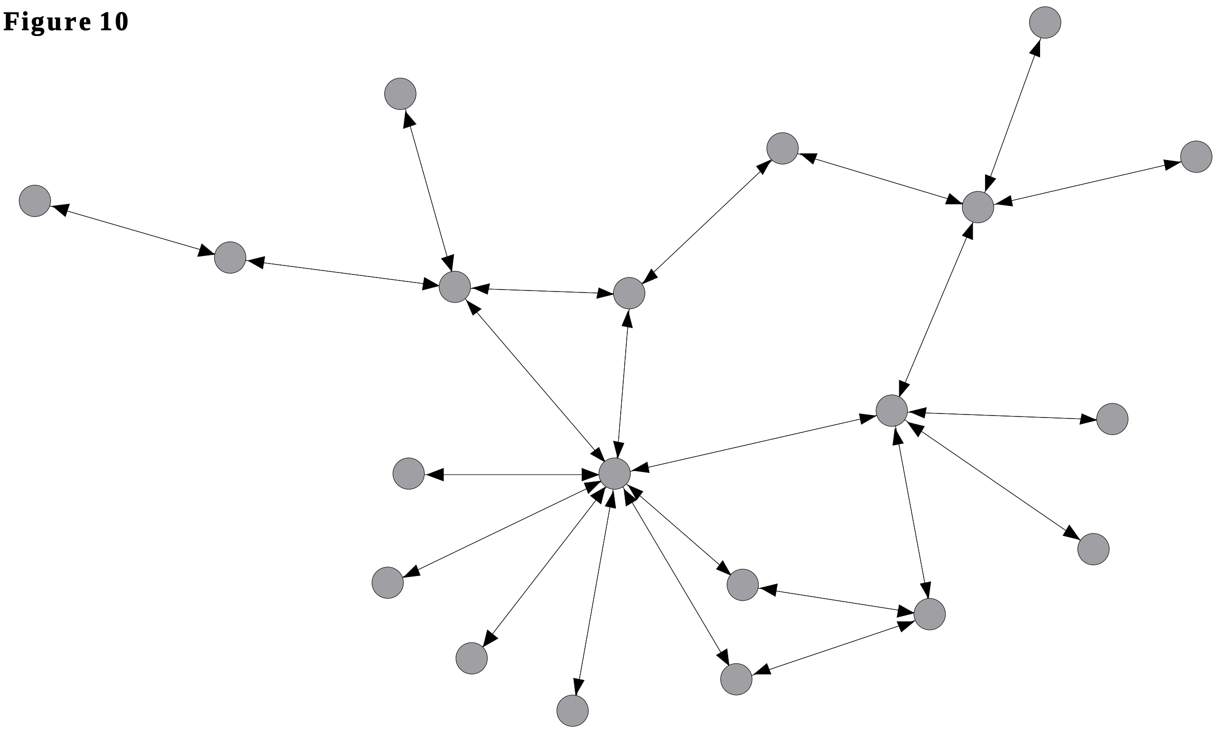
White, H.C. (1992). *Identity and Control*. Princeton, NJ: Princeton University Press.

Appendix

Diagnostic Utility

This tool provides an indication of whether the power indices the Bonacich measure produces for a given network can be adequately represented by a combination of $P_{-\kappa}$ and $P_{+\kappa}$ or $P_{-\kappa}$, P_0 and $P_{+\kappa}$. In the example below I use the network in **Figure 10**.

Figure 10



This is its DL file adjacency matrix

```
DL
NR=20
NC=20
FORMAT=FULLMATRIX
DIAGONAL=PRESENT

DATA:
0 1 1 0 0 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0
1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0
1 0 0 0 1 0 1 1 0 1 0 0 0 0 1 0 0 1 1 1
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1
0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
```

Once saved as, for example, “test.dl”, the command to invoke the program is:

C:>[path to program\]bonacich_disgnostic_tool.exe [path to input file]

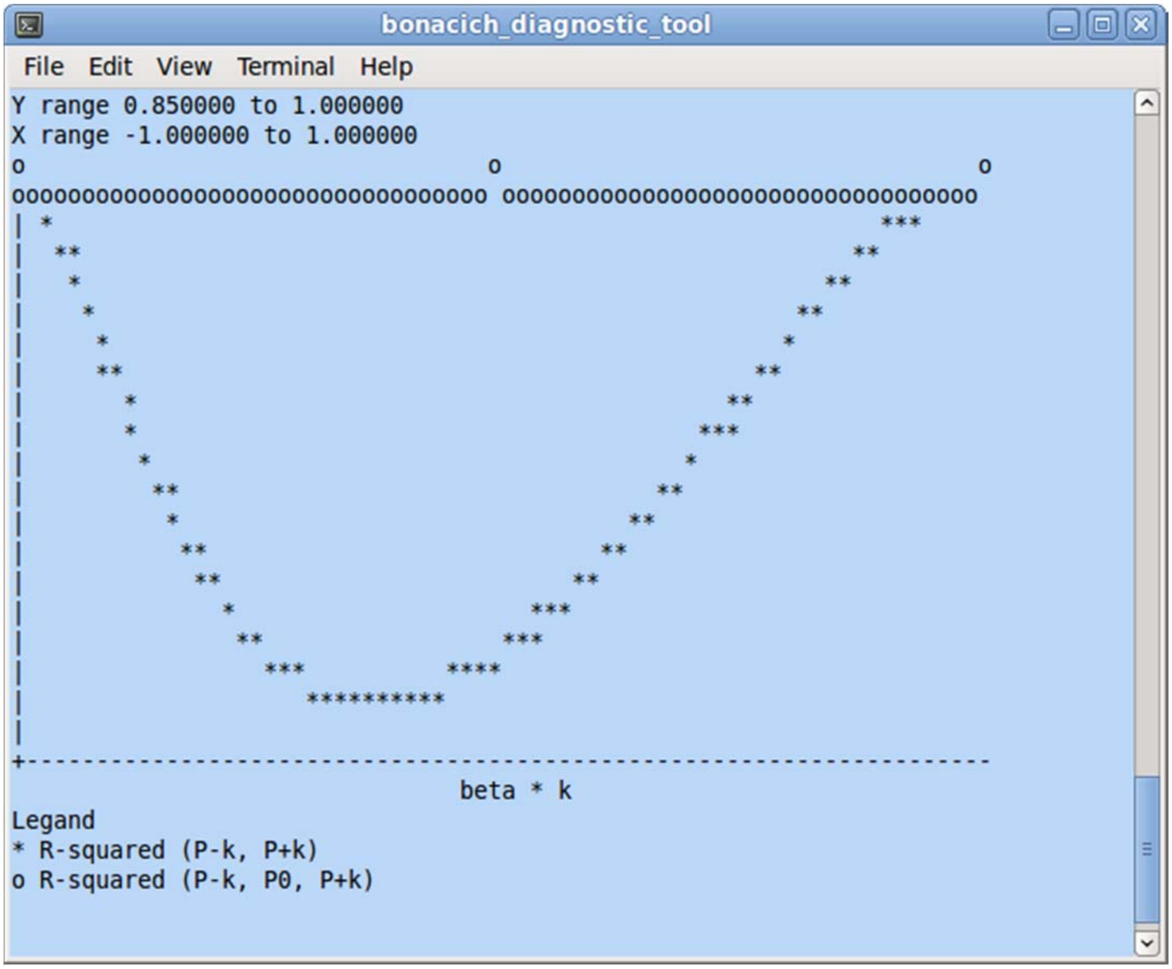
The path to the input file containing the network data can be specified on the command line; if it is not, the program will prompt you for the file path after a brief explanation of the program and the output it produces. The input file must be a square binary adjacency matrix in DL format.

The program will produce some on-screen diagnostic information. First it reports the degree of correlation between the three pure strategy vectors:

```
Correlation between P-k and P+k = 0.883327
Correlation between P-k and P0 = 0.959981
Correlation between P0 and P+k = 0.975144
```

It also produces a fairly crude character based graph of the R-squared values and β shown in **Figure 11**.

Figure 11



Finally, it writes a new file to the same directory in which the input file resides, with the diagnostic data from which the graphs in **Figure 9** were generated.

In this example, the two vector representation is fairly weak, with the R-squared dropping to close to 0.85 at a β value of about -0.2. In contrast, the three vector $P_{-\kappa}$, P_0 and $P_{+\kappa}$ approach seems fine for this network.

Multiple P_β Generator

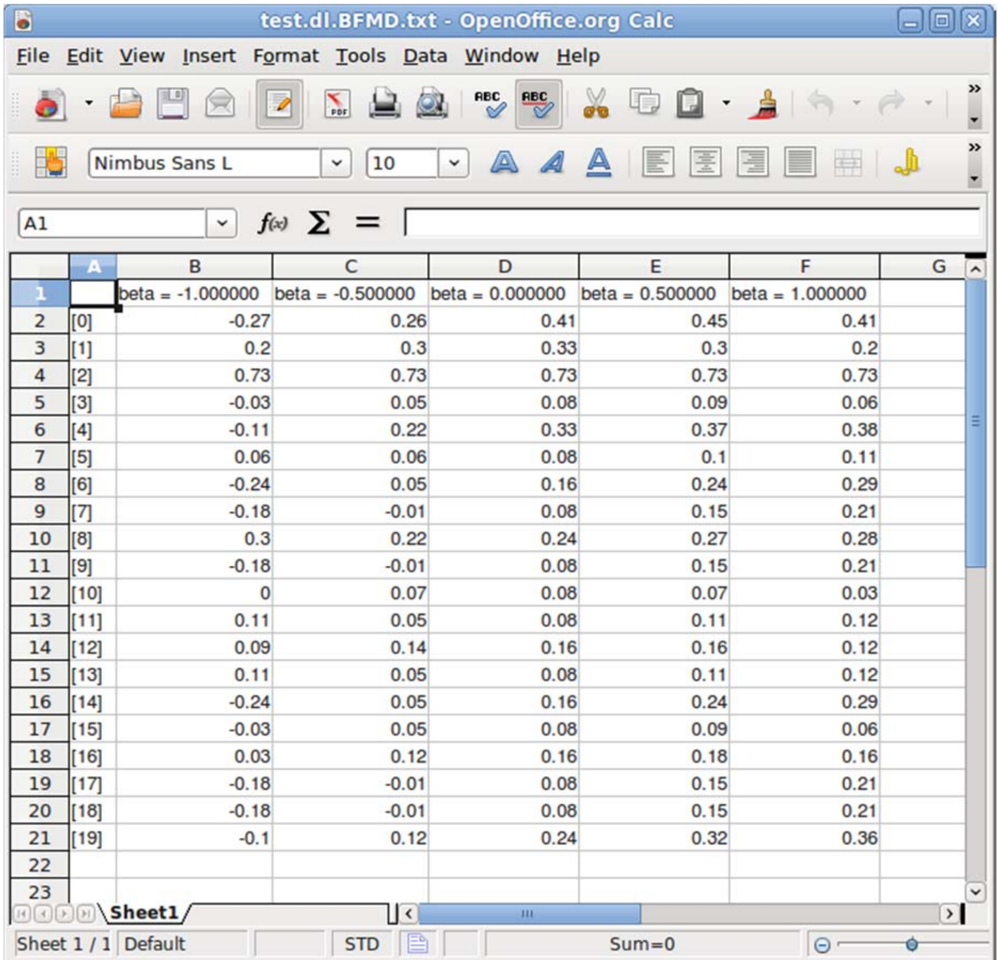
This utility generates multiple power index vectors for a verity of β values and optionally can correlate each of these with any number of dependent variables.

Using the file, "test.dl", and issuing the following at the command prompt:

C:>[path to program\]bonacich_family_generator.exe –net [path\]test.dl

will generate two files, one a tab-delimited text file the other a dl format file. Both contain the same data. The .txt file output will look like **Figure 12**, shown here after being imported into a spreadsheet.

Figure 12



test.dl.BFMD.txt - OpenOffice.org Calc

	A	B	C	D	E	F
1		beta = -1.000000	beta = -0.500000	beta = 0.000000	beta = 0.500000	beta = 1.000000
2	[0]	-0.27	0.26	0.41	0.45	0.41
3	[1]	0.2	0.3	0.33	0.3	0.2
4	[2]	0.73	0.73	0.73	0.73	0.73
5	[3]	-0.03	0.05	0.08	0.09	0.06
6	[4]	-0.11	0.22	0.33	0.37	0.38
7	[5]	0.06	0.06	0.08	0.1	0.11
8	[6]	-0.24	0.05	0.16	0.24	0.29
9	[7]	-0.18	-0.01	0.08	0.15	0.21
10	[8]	0.3	0.22	0.24	0.27	0.28
11	[9]	-0.18	-0.01	0.08	0.15	0.21
12	[10]	0	0.07	0.08	0.07	0.03
13	[11]	0.11	0.05	0.08	0.11	0.12
14	[12]	0.09	0.14	0.16	0.16	0.12
15	[13]	0.11	0.05	0.08	0.11	0.12
16	[14]	-0.24	0.05	0.16	0.24	0.29
17	[15]	-0.03	0.05	0.08	0.09	0.06
18	[16]	0.03	0.12	0.16	0.18	0.16
19	[17]	-0.18	-0.01	0.08	0.15	0.21
20	[18]	-0.18	-0.01	0.08	0.15	0.21
21	[19]	-0.1	0.12	0.24	0.32	0.36

Each row has the power indices for a node at different values of β, each column has the power indices for a particular value of β.

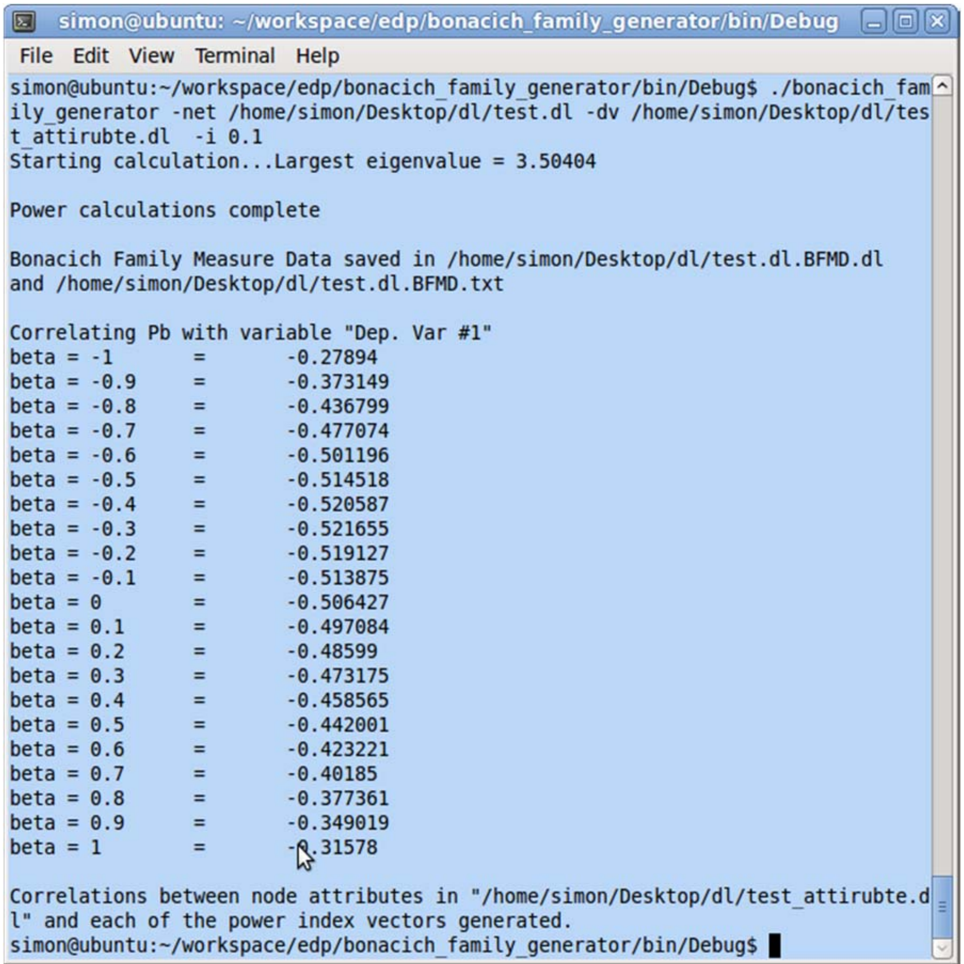
Adding the optional argument –i n where n is a decimal in the range (0,1] will adjust the interval between adjacent values of β. The default is 0.5 which generates power indices at five values of β as shown in **Figure 12**.

The –dv [path\]<name of attribute file.txt> argument can be used to correlate the power indices at each β value with a node attribute variable or variables.

The attribute file should be a matrix with a row for each node and a column for each dependant variable to be tested. This file can be either a DL format (which the utility identifies by the extension '.dl') or a simple text file (extension '.txt'). If the file is a simple text file it may contain only the data, or it may have column headings; but if it does, each row must also have a row label.

For example using the network in 'test.dl' and an attribute file containing the following values {1, 3, 2, 4, 3, 5... 20} the program prints the correlations to the screen as shown in **Figure 13**.

Figure 13



In this example a β value of -0.3 produces power indices that correlate most strongly with the dependant variable provided.

The source code for both programs is available on request.