J. Phys. A: Math. Theor. **41** (2008) 224023 (12pp) doi:10.1088/1751-8113/41/22/224023

A social network's changing statistical properties and the quality of human innovation

Brian Uzzi[1,2,3]

[1] Kellogg School of Management, Northwestern University, Evanston, IL, USA
[2] Northwestern Institute on Complexity (NICO), Evanston, IL, USA
[3] Haas School of Business, University of California at Berkeley, CA, USA

E-mail: uzzi@northwestern.edu

Received 10 December 2007, in final form 26 March 2008
Published 21 May 2008
Online at stacks.iop.org/JPhysA/41/224023

Abstract
We examined the entire network of creative artists that made Broadway musicals, in the post-War period, a collaboration network of international acclaim and influence, with an eye to investigating how the network's structural features condition the relationship between individual artistic talent and the success of their musicals. Our findings show that some of the evolving topographical qualities of degree distributions, path lengths and assortativity are relatively stable with time even as collaboration patterns shift, which suggests their changes are only minimally associated with the ebb and flux of the success of new productions. In contrast, the clustering coefficient changed substantially over time and we found that it had a nonlinear association with the production of financially and artistically successful shows. When the clustering coefficient ratio is low or high, the financial and artistic success of the industry is low, while an intermediate level of clustering is associated with successful shows. We supported these findings with sociological theory on the relationship between social structure and collaboration and with tests of statistical inference. Our discussion focuses on connecting the statistical properties of social networks to their performance and the performance of the actors embedded within them.

PACS number: 89.75.−k

"...the stage is 'The Mirror of Nature,' and the actors are 'The Abstract, and brief Chronicles of the Time;' – and pray what can a man of sense study better?" (The Critic 1779, in Sheridan 1962).

What drives outstanding human achievement? Many have believed that great minds work in isolation and find singular inspiration. Yet, recent research has shown that this stereotype only rarely fits the reality. Following the careers of all the notable scientists, artists and philosophers since recorded time in the Eastern and Western civilizations, Collins (1999) found that the most creative geniuses, from Freud and Darwin to Beethoven and Curie, were embedded in networks of other scientists, researchers and artists who shared ideas through competition and collaboration. This trend is apparently intensifying in contemporary science. Teams on average now produce the most highly cited work, inverting the advantage that individuals once possessed (Wuchty *et al* 2007). At the same time, methodological developments have made the mysteries of large network structures reckonable, enabling new analyses of how network topology may affect human dynamics (Granovetter 1973, Barabási *et al* 2000, Burt 2004, Guimera *et al* 2005, Colizza *et al* 2006, Caldarelli 2007).

In this paper, we build on previous results to analyze the statistical properties of a large collaboration network over a long time period, with a focus on how change in the network's topology affects the performance of the system (Uzzi and Spiro 2005). We examined whether different topologies could shape the organization of the creative talent, amplifying or stultifying its innovativeness. Our analysis investigates the statistical properties of the collaboration network of all the creative artists that made Broadway musicals from 1945 to 1989. Our results indicate that the topography of the network may substantially affect the performance of the actors within it, suggesting that the arrangements of talent, not just the presence or absence of talent, underlies successful creative enterprises.

Broadway musical creative artist network, 1945–1990

The Broadway musical industry (BMI) network of creative artists includes the artists responsible for the creation of Broadway musicals, an internationally recognized performance art that blends music, lyrics, dance, stage design and story into a single seamless artistic production. It takes its name from the locale in the New York City along Broadway Street where it developed in the late 1890s. Typically, six specialists—choreographer, librettist, composer, lyricist, producer and director—team up to create a musical. In this network, artists are nodes and collaborations between two artists are undirected edges. Our analysis examines all Broadway musicals made from 1945 to 1989. After 1945, an influx of new talent created a distinct post-1945 period. Since virtually all productions are done by a team of approximately six or seven artists, our network is a bipartite network with artists clustered within teams of about the same size. Artists who worked on the same show are the members of a fully linked team-clique (e.g., the team that made 'Evita'). Links form between artists on different teams when an artist works on multiple teams.

Our data include all 474 musicals of new material released between 1945 and 1989 as well as 49 musicals that closed in preproduction during the 1945–1989 period. These data allow us to capture all the professional links among artists and not just those due to finished products, which avoids network sampling bias. We have information on a musical's opening date, theatre, creative artist team, financial success and critical success. These data are recorded in *Playbills* (Simas 1987, Green 1996). By adding new shows to the network and removing inactive artists, the network structure can change with time. Based on interviews with contemporary artists, we determined that artists 'drop out' of this network after seven years of no new productions. (Failure to drop out inactive artists gives a biased view of the real network by maintaining impossible links such as a link between Andrew Lloyd Weber (*b.* Mar 22, 1948) and George Cohan (*b.* Jul 4, 1878—*d.* Nov 5, 1942). Consequently, we added nodes and ties each year with the founding of new productions and deleted nodes and their ties that were inactive for seven years. The seven-year decay rule was confirmed statistically. There is $\leqslant 5\%$ chance that an artist would make another show once they have been inactive for seven years. Our first year of observation was constructed using the artists who had been active within the prior seven years and each year after that new artists were added to it and inactive artists dropped from it. To link the network's structure to performance,

Table 1. Network statistics.

Network	n	M	k_{mean}	L	CC	r	α
Broadway	4 756	*28 180*	*10.84*	*1.37*	*0.40*	*0.21*	1.09
Rappers	5 533	57 972	20.95	3.90	0.18	0.06	3.50
Movie actors	449 913	25 516 482	113.43	3.48	0.20	0.21	2.30
Board directors	7 663	55 392	14.44	4.60	0.59	0.28	–
Jazz musicians	1 275	38 326	60.30	2.79	0.33	0.05	–
Brazilian pop music	5 834	507 005	173.80	2.30	–	–	2.57

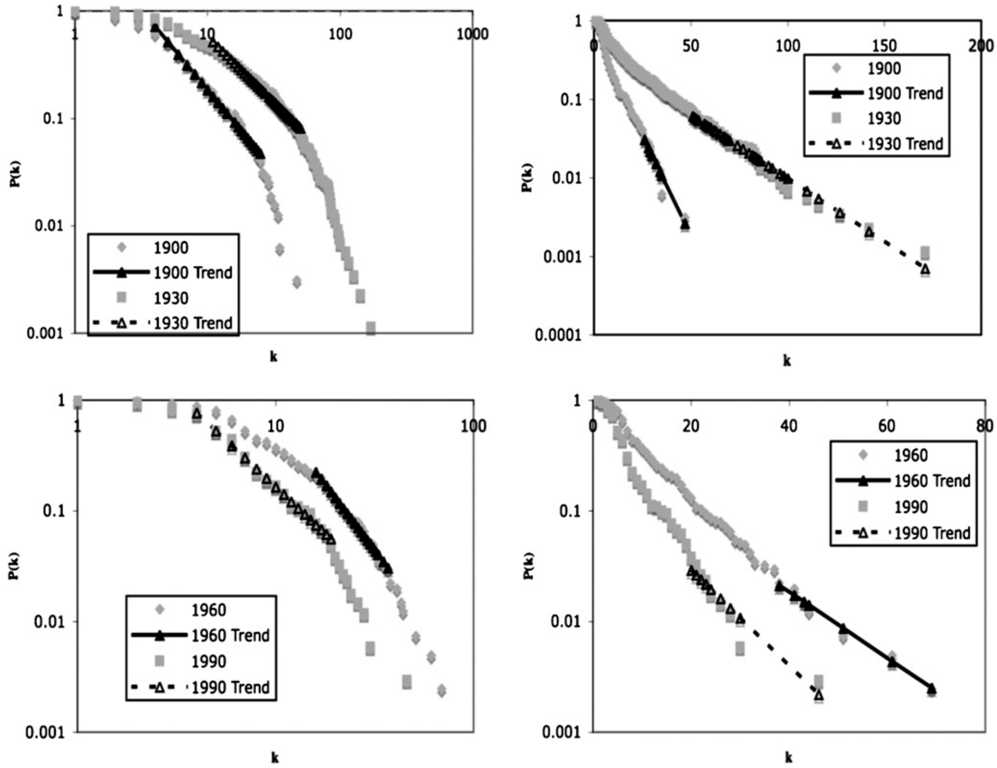
Note: Network statistics for various artists' networks reported in Smith (2006). n is the number of nodes in the network, M is the number of undirected edges, k is the mean degree per node, L is the average path length, CC is the clustering coefficient, r is the degree–degree correlation coefficient between artists and α is the power-law scaling exponent. All statistics represent the values of the last year of observation, which includes all the nodes and links previously existing in the network. For the Broadway musical network, the yearly means are 464, 10.84, 1.35, 2537.33, 0.35, 0.09 and 1.55 for n, M, k, L, CC, r and α, respectively.

we measured success using the two standard industry metrics: whether the show made or lost money and critic's reviews. Table 1 compares the BMI network with other bipartite networks reported in the literature.

Network analysis

We analyzed four global properties shown to distinguish the types of networks: degree distributions, assortativity, clustering and path length (Watts 2004, da Fontoura Costa *et al* 2007, Zhang *et al* 2007). Scale-free distributions reflect the influence a few actors have over many other actors (Wasserman and Faust 1994). Figure 1 shows the degree distribution plots for 1930 and 1960. Plots for other time periods were similar. To test the true underlying distribution, we used Clauset *et al*'s (2007) method, which implements the Kolmogorov–Smirnov (KS) statistic to compare different distributions and reduce the misleading inferences that have followed from using only visual inspection (Amaral and Guimera 2006, Clauset *et al* 2007). The KS statistic looks at the maximum difference between the actual distribution and the values predicted by the distribution. We checked all the possible combinations of distributions for the power-law, exponential and log-linear distribution across the full degree distribution and for subsets of points in the degree distribution. Figure 1 shows that a power-law distribution best fits the degree distribution in the middle of the distribution and an exponential distribution best fits the degree distribution at the tail. The 'hump' at the head of the distribution arises because team size across productions is nearly invariant, making the odds of persons with fewer than six contacts rare. The results show that the super-connected nodes are few in number and have roughly the *same* $P(k)$. This indicates that in this network there may be an upper limit on the number of contacts a super connector can possess due to finite resources for making and managing ties.

Related to the degree distribution, the degree–degree correlation measures the network's assortative and disassortative mixing (Newman and Park 2003). When the assortativity is high, it means that well-connected nodes are connected to each other. In a collaboration network, it suggests that the most successful artists (those with lots of ties) are working on shows together. Smith (2006) reported that rappers had surprisingly low assortative mixing ($rho = 0.06$ and 0.05, respectively) but that movie actors and board of directors had moderate assortative mixing values ($rho = 0.20$ and 0.27, respectively).



Kolmogorov–Smirnov tests of distributional fits

Year	Range		Kolmogorov–Smirnov statistic			
	x-min	x-max	Power law w/cut-off	Power Law	Exponential	Log Normal
1900	4	25	0.10	0.48	0.16	0.52
1930	11	50	0.06	0.57	0.10	0.52
1960	16	37	0.02	0.07	0.04	0.47
1990	4	19	0.05	0.47	0.32	0.51

Figure 1. Degree distributions, Broadway musical industry creative artists.

Figure 2 shows the degree–degree correlations for the BMI for artists and teams along with the level due to the bipartite structure. Consistent with the fact that most teams in this network are about the same size, there is little expected assortativity due solely to the bipartite structure. The artist-to-artist degree–degree correlations start off quite low at 0.10 but climb up to about 0.20, a value close to that of movie actors and board directors. In contrast, the team-to-team degree correlations show a relatively high average correlation over our time frame of $rho = 0.35$ with a slight downward trend. One explanation for these trends is that assortative mixing depends on the creation of 'stars'. Once stars emerge (in our case the post-1945 stars), assortative mixing can increase at an increasing rate as stars find and work with each other. Also, the relatively moderate values of 0.20 suggest that assortative mixing levels may be limited when the unique creative styles of superstars may be incompatible. This may explain the common observation on Broadway that artists from different 'camps', despite their talent, never work together. Finally, the relatively high assortative mixing level

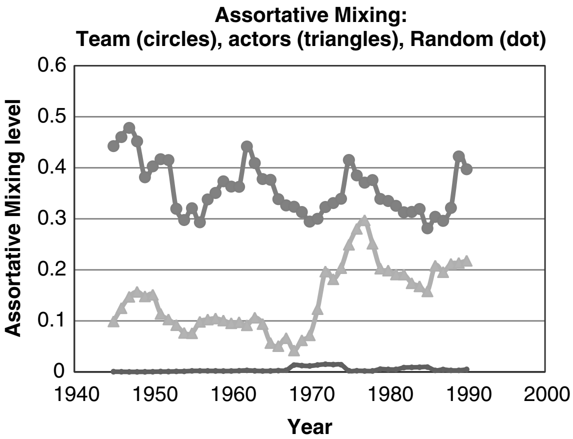


Figure 2. Degree–degree correlations.

of 0.35 for teams suggests that in bipartite networks, on average, connected teams have more comparable degrees than do disconnected teams, which suggests that at the team level there is a more requisite sorting of connections than at the individual level.

Two other important global network properties are the clustering coefficient and average path length. The clustering coefficient was operationalized as the total number of closed triads in the network relative to the number of possible triads, where a closed triad represents the smallest unit of complete clustering (i.e., clustering = 1.0) because each person in the triad (A, B and C) is connected to each other. We used the formula for the weighted clustering coefficient, CC_a, as

$$CC_a = \frac{3 \times \text{number of triangles on the graph}}{\text{number of connected triples of vertices}}. \tag{1}$$

As a bipartite network, the standard clustering coefficient statistics are biased. This is because the fully linked cliques that make up each production team artificially overstate the level of the true clustering, and understate the true path length, making the network appear to have more clustering and a shorter path length than it actually has when compared to the standard random graph. To overcome this problem, we used Newman *et al* (2001) bipartite statistics to measure the topography of the network. To remove the within-team clustering, the bipartite random cluster coefficient maintains the two-degree distributions in the network: the number of individuals per group and the number of groups per individual. The probability that an individual is in j groups is p_j. The probability that a group has k individuals is q_k. These probabilities are used to construct the functions for these distributions in formula (2). Formula (3) is a 'generating' function that keeps track of these distributions for all the actors and teams in the network. Together they are used to calculate the number of neighbors that an individual has in the unipartite projection of the network. Finally, formula (4) is used to calculate a bipartite random cluster coefficient. In formula (4), M is the total number of groups in the network and N is the total number of individuals in the network (Newman *et al* 2001):

$$f_0(x) = \sum_j p_j x^j, \qquad g_0(x) = \sum_k q_k x^k, \tag{2}$$

$$G_0(x) = f_0(g_0'(x)/g_0'(1)), \tag{3}$$

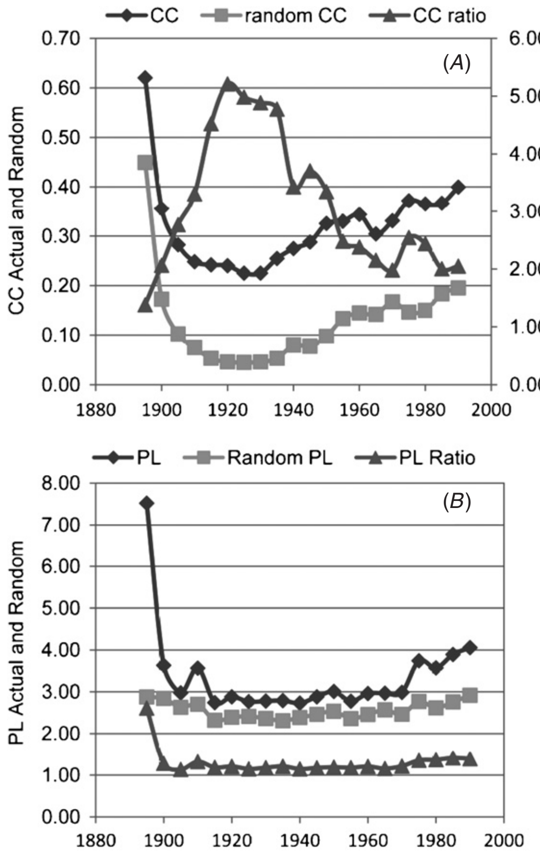


Figure 3. Clustering and path length over time.

$$_bCC_r = \frac{M}{N}\frac{g_0'''(1)}{G_0''(1)}. \tag{4}$$

The average path length for the network is calculated by finding the shortest path from each individual to every other and then taking the average of these scores. The random unipartite path length *understates* differences between the actual and random path lengths in a bipartite network. To adjust for the differences, paths must be traced in two directions: from the perspective of the actor, and the team that the actor is in by taking the first derivative of the functions defined in formula (2), evaluated as 1. The PL *ratio* is equal to (PL *actual*)/(PL *random*); the greater this ratio is, the longer are the average paths between actors then would be expected in a random network of the same size and degree distributions:

$$_b\mathrm{PL}_r = \ln(n)/\ln[f_0'(1)\cdot g_0'(1)]. \tag{5}$$

Figure 3(a) shows that the CC ratio falls and then rises. This network had a very high clustering among artists right after 1945 that consistently dropped, up to about 1970, and then began to climb again up to the present period. In contrast, the PL ratio remained relatively flat over the entire time period.

Looking across our four measures of topography, it is noteworthy that the degree distribution, assortativity and path length changes only minimally despite our decay function and the fairly tumultuous shocks induced by the advent of TV in the 1950s, Rock n' Roll in the 1960s, mass marketed movies in the 1970s and AIDS—all of which affected the industry's talent pool and product demand. From a methodological perspective of the time-series analysis of social networks, the steady-state nature of these features suggests that cross-sectional point estimates of these features provide reliable indicators of these features over short time frames. In addition, it may suggest that these topographic features are slow changing relative to the entry and exit of talent, making them only marginally associated with the yearly changes in the performance of the system. In contrast, the clustering coefficient of the network changes greatly over time, a finding that is consistent with Watts (2004) and da Fontoura Costa *et al* (2007) conclusions that the main discriminating global feature of a social network is the clustering coefficient. Because we are interested in how the network structure and performance are associated, we focused on how changes in clustering vary with the propensity of the artists in the BMI to produce hit musicals.

Time-series clustering and human dynamics

Building on extensive sociological fieldwork regarding how collaboration relates to creative success (see Becker 1982), we speculated that changes in the clustering coefficient (CC) ratio may be associated with the connectivity and cohesiveness of ties between artists across the whole network, which in turn affects the creativity of individual artists by governing their level of access to the creative material and talent that is unevenly distributed in the hands and minds of artists throughout the network.

Extensive sociological fieldwork regarding how collaboration relates to creative success has found that artists improve their chances of producing hit products to the degree that they can access diverse pools of creative material and talent. The more artists can effectively access diverse pools of talent and creative material circulating throughout their global network, the more opportunities they are likely to have to experiment with and create hit products out of the new combinations of existing material (Becker 1982, Collins 1999, Burt 2004). The CC ratio is a structural measure of the degree to which actor's ties are clustered—collaborators of collaborators are collaborators. Furthermore, in bipartite networks, the CC ratio can be further decomposed into 'within-team ties' and 'between-team ties' (Newman *et al* 2001). A 'within-team tie' is a link among nodes on the same team. Within-team ties give a proxy for how cliquish pools of creative material may be. Conversely, a between-team tie is a link between separate production teams. It arises when two separate teams share at least one common member. Between-team ties give a proxy for how much mixing there might be in the global network among diverse pools of talent and creative material within the separate cliques.

Because all members of the same team form a fully linked clique, each within-team clustering coefficient is equal to 1.0. This also means that any random reshuffling of within-team ties on average produces a CC ratio (i.e., the real clustering random clustering) of 1.0. Consequently, a CC ratio that is greater than 1.0 in a bipartite network is due to there being more between-team ties than expected purely by chance. Consistent with the reasoning, we found that the CC ratio correlated with the frequency of between-team ties at $rho = 0.922$ ($p < 0.001$).

While an increase in the level of clustering suggests that teams become more connected to each other through actors that work on multiple productions with different people, it does not tell us about the quality of these bridging ties. Are the ties that bind separate teams made up of first time links or are they made up of repeated ties? According to sociological theory, first time and repeated ties have different associations with successful collaboration (Becker 1982). A repeated relationship suggests that the actors have established an effective working relationship for collaborating and that they have a cohesive relationship based on trust, which lowers partnering costs and increases the chances that the artists will be willing to share the risk of innovation (Uzzi 1997). Similarly, repeated ties are associated with third party ties in common (Watts 2004), which further improves coordination (Coleman 1988). Examining these changes empirically, we found a moderate and statistically significant correlation between the CC ratio and the presence of repeated and third party ties ($rho = 0.27$; $p < 0.01$ and $rho = 0.25$; $p < 0.06$).

Sociologically, these results may suggest that as the CC ratio rises, the links between teams increase in frequency, enabling creative material within teams to spread to, and benefit, other teams in the global network. At the same time, as the CC ratio rises, within-team ties become disproportionately made up of repeated and third party ties as artists increasingly repeat relations with teammates they have worked with in the past or their collaborators have worked with in the past. Thus, as the real clustering increases the network, it becomes more connected and connected by persons who know each other well, facilitating access to diverse creative material circulating around the network in separate teams.

However, as clustering rises, social research also suggests that these same processes can create liabilities for collaboration. Too much triadic closure can stifle effective commercial and artistic collaboration if the social aspects of exchange supplant the practical imperatives. Feelings of obligation and friendship (or the settling of scores) may be so great that the social dynamics undercut the potential gains from collaboration. This can lead clusters to reduce the recruitment of new talent in order to preserve a space for 'friends' (Portes and Sensenbrenner's (1993: 1339). McPherson *et al* (2001) argued that interconnected networks can promote recruitment by homophily, reproducing rather that extending the conventional wisdom with diverse novelty. Finally, intensive interconnectedness makes it more likely that actors on the same teams will access similar rather than different pools of creative material, making it less likely that any artist will produce novel combinations of new material (Uzzi 1997).

This reasoning suggests that an increased structural connectivity in the global network over some threshold may produce detriments, not benefits. When the CC ratio is low, innovation remains isolated in the separate teams because there are few between-team links that support the transfer of creative material and risk taking. When the clustering coefficient ratio is at medium level, an increased network connectivity enables novel ideas within teams to spread to other teams and be supported by repeated and third party ties that promote an efficient and effective partnering. However, when the CC ratio is high, connectivity tends to homogenize creative material within the industry by making the same material available to everyone. At the same time, repeated and third party in-common ties promote closure, decreasing artists' abilities to break out of conventional ideas that have lost their market appeal.

To empirically examine the possible statistical association between the CC ratio and the success of the artists in the BMI network, we measured (i) the percentage of financially successful shows and (ii) the percentage of critically acclaimed shows. The percentage of financially successful shows measures the fraction of shows released that year that turned a net profit. The percentage of critically acclaimed shows measures the fraction of shows released that year that received critical acclaim based on the artistic merit of the show. About 25% of the shows that receive artistic acclaim are financial flops (Uzzi and Spiro 2005). Financial data were published in *Variety*. Critical acclaim data was coded from critics' print reviews on a five-point scale, pan to rave (Suskin 1990).

Similar to natural and biological systems, the behavior of social systems can be affected by variables other than the network's structure. It is therefore important to determine the conditional effect of clustering after parsing out the effects of other factors that affect success.

Previous research has identified several variables (Reddy *et al* 1998). The more talented a creative artist is, the more likely she may be able to turn raw artistic materials into something with special appeal. A proxy for talent is experience since high-talent artists work over and over again, while low-talent artists tend to work only once. Consequently, because the degree–degree distribution of teams indicates whether highly talented people are tending to work with each other, we used this as a measure of the level of talent in the network. We also used a number of artists on the teams releasing new shows who had hit shows in the prior years. Both variables produced the same effects and were correlated, so we report only the former. The rate of hit shows each year also depends on the level of disposable income that consumers have to spend on entertainment. Thus, we controlled for the percentage change in the national gross domestic product adjusted for inflation. Ticket prices affect the financial performance of shows and expectations for artistic quality. We included a variable for the inflation-adjusted ticket price for each year. Direct competition from other shows affects how many shows are likely to have above average performance. To control for this relationship, we added a variable that measured the percentage of musicals playing from the previous year. Finally, we included a fixed effect variable for year to control for time trends and unmeasured variables that remain constant within the year, such as network size, total k, consumer tastes, etc. Network characteristics were measured one year prior to the outcome variables. The control variables were coterminous.

Table 2 presents the OLS regressions of the above models with linear, log-linear and curvilinear specifications of the CC ratio (Tobit regressions, which directly model any potential problems with the bounded dependent variables, produced the same results). The models suggest that the association between the CC ratio and performance is curvilinear ($p < 0.05$). Additionally, they suggest that financial success is affected by economic growth but artistic success is not, which is indicative of the fact that consumer attendance lowers when the economy is down but that critic's reviews should be unaffected.

Figure 4 graphically presents the estimated U-shaped association between the CC ratio and the financial and artistic performance of the system (regression lines show the predicted values with CIs at $p < 0.05$). Based on the preceding sociological fieldwork and theory, a likely account of these findings is that as the CC ratio increases, the separate production teams that make up the small world become increasingly connected to each other and connected by artists who have repeated relationships or third parties in common. This greater level of connectivity and cohesion may boost the performance of the agents within it by increasing their access to diverse and novel creative material circulating in all parts of the small world. However, as this connectivity and cohesion increases beyond a threshold, its benefits decrease, and eventually turn negative. Very high levels of connectivity and cohesion may lead to a homogenization and imitation of the same ideas by the different teams in the network, lowering the opportunity for individual teams to distinguish themselves with an exceptional show material.

To check these findings, we did an out-of-sample test. Table 3 lists the number of shows that opened each year, the hit percentage, and the yearly CC ratio for the 1920–1930 season (Jones 2003: 360–1). While complete data on these shows are unavailable, they can help refute or support the association reported above. During 1945–1989, the flop rate was about 19% for low-CC ratio values (1.0–1.5) and 22% for the high values of clustering (3.0–3.5). Table 3 shows that during 1920–1930, the CC ratio averaged 4.8, nearly 40% higher than the highest value in the 1945–1989 period. This high CC ratio suggests that the flop rate should have been high in the 1920–1930 season and higher than the flop rate from 1945–1989. Table 3 shows that the 1920–1930 flop rate was nearly 90% on average, despite a roaring economy and an industry flush with some of the best artistic talent it has even known (Uzzi and Spiro 2005). Moreover, figure 4(inset) shows that the association between the CC ratio and the hit

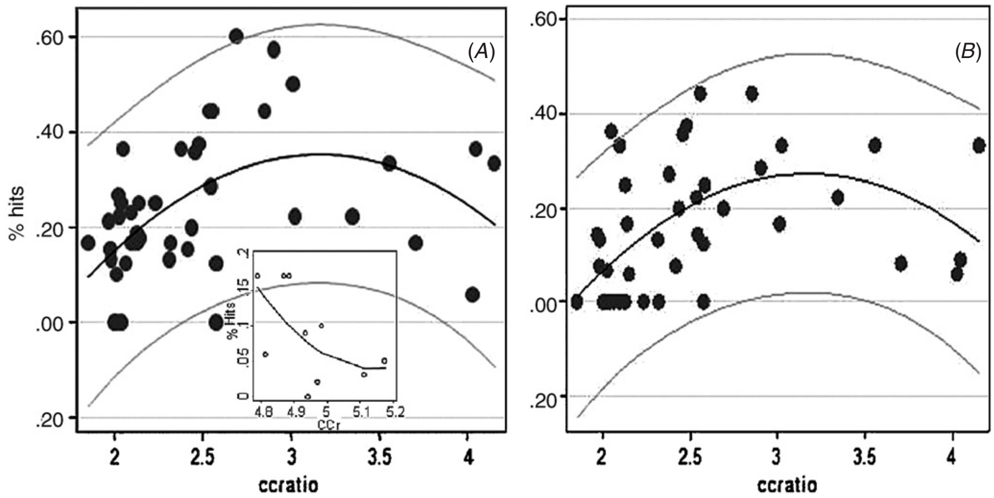


Figure 4. Yearly percentage of financial hit (A) and the percentage of artistic hit (B) shows, 1945–1990.

Table 2. OLS regressions of the association between the CC ratio and the system performance.

	% Financial hit shows			% Artistic hit shows		
Variable	Linear	Log-linear	Curvilinear	Linear	Log-linear	Curvilinear
CC ratio	−0.0665		0.703**	−0.105		0.586**
	(−0.098)		(−0.34)	(−0.077)		(−0.26)
CC ratioˆ2			−0.124**			−0.112***
			(−0.053)			(−0.041)
CC ratio (log)		0.0248			−0.0965	
		(−0.28)			(−0.22)	
Path length ratio	0.588	0.305	0.244	0.547	0.296	0.239
	(−0.59)	(−0.62)	(−0.58)	(−0.47)	(−0.49)	(−0.45)
Yr trend fixed effect	−0.0119*	−0.00736	−0.00842	−0.0131**	−0.00897	−0.0100*
	(−0.0068)	(−0.007)	(−0.0066)	(−0.0054)	(−0.0056)	(−0.0051)
Assortativity	−0.419	−0.628	−0.308	−0.435	−0.642	−0.335
	(−0.59)	(−0.58)	(−0.56)	(−0.46)	(−0.46)	(−0.43)
% Change GDP	1.793**	1.563*	1.543*	0.787	0.576	0.563
	(−0.82)	(−0.83)	(−0.78)	(−0.65)	(−0.66)	(−0.6)
Ticket price	0.0055	0.0029	0.005	0.00176	−0.000802	0.00126
(inflation-adjusted)	(−0.0053)	(−0.0053)	(−0.0051)	(−0.0042)	(−0.0042)	(−0.0039)
Shows playing from	0.0003	−0.00217	−0.0024	−0.00465	−0.00691	−0.00711
previous year	(−0.01)	(−0.01)	(−0.0096)	(−0.008)	(−0.0082)	(−0.0074)
Constant	23.07*	14.49	15.59	25.71**	17.78	19.00*
	(−13)	(−13.3)	(−12.6)	(−10.2)	(−10.6)	(−9.71)
Observations	45	45	45	45	45	45
R-squared	0.28	0.27	0.37	0.52	0.49	0.61

Note: S.E. in parentheses: $p < 0.05$**, $p < 0.10$* (two-tailed).
Bayesian Information Criterion strongly supports the curvilinear models (BIC = 17.17).

rate is strongly negative ($rho = 0.55$; $p < 0.05$, one-tailed) for the 1920–1930 period, a finding consistent with our analysis.

Table 3. CC ratio and financial hit rate, 1920–1930 out of sample prediction (see the inset in figure 4).

	Year										
	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930
Shows	41	42	35	41	40	42	42	49	51	42	34
% Hits	5	5	3	2	0	10	17	6	17	9	17
CC ratio	5.211	5.175	5.112	4.97	4.941	4.983	4.788	4.815	4.868	4.933	4.886

Discussion

We built on recent methodological advances for the measurement of bipartite networks and sociological theory on social interaction to examine how the change in the topography of a social network is associated with the performance. Our findings show that some of the topographical qualities of this network are relatively stable (i.e., degree distributions, assortative mixing and average path length), which suggests that the cross-section analyses of these quantities offer reasonable point estimates of their levels when time differences are not great. In contrast, we found that the clustering coefficient ratio of the network changes substantially over time. Moreover, using the change in the clustering coefficient ratio as a proxy for the connectivity and cohesion of artists within the network, we found that changes in the clustering coefficient were statistically and substantively associated with changes in the performance of the system. Our regression analyses revealed that the clustering coefficient ratio has an inverted U-shaped association with the financial and artistic success of the Broadway. These findings also raise new questions about causality in networks between structure and agency. We hope that future work will begin to make more connections between structure and performance, a trend that already seems to be producing productive knowledge in terms of human productivity (Barabási 2005) and the citation impact of scientific teams (Guimera *et al* 2005).

References

Amaral L A N and Guimera R 2006 *Nature Phys.* **2** 75
Barabási A-L 2005 *Nature* **207** 435
Barabási A-L, Albert R and Jeong H 2000 *Physica* A **281** 69
Becker H 1982 *Art Worlds* (Berkeley, CA: University of California)
Burt R S 2004 *Am. J. Sociol.* **110** 349
Caldarelli G 2007 *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford: Oxford University Press)
Clauset A, Shalizi C R and Newman M 2007 *Preprint* 1062v1 1
Coleman J S 1988 *Am. J. Sociol.* **94** S95–120
Collins R 1999 *The Sociology of Philosophies* (MA: Harvard University Press)
Colizza V, Barrat A, Barthélemy M and Vespignani A 2006 *Proc. Natl Acad. Sci.* **103** 2015
da Fontoura Costa L, Fernandes R and Andrade S 2007 *New J. Phys.* **9** 1
Granovetter M 1973 *Am. J. Sociol.* **78** 1360
Green S 1996 *Broadway Musicals: Show by Show* (Milwaukee, WI: Hal Leonard Corp.)
Guimera R, Uzzi B, Spiro J and Amaral L A N 2005 *Science* **308** 697
Jones J B 2003 *Our Musicals, Our Selves* (Hanover, NH: Brandeis University Press)
McPherson M, Lynn S-L and Cook J M 2001 *Ann. Rev. Sociol.* **27** 415
Newman M E J and Park J 2003 *Phys. Rev.* E **68** 1
Newman M E J, Strogatz S H and Watts D J 2001 *Phys. Rev.* E **64** 1
Portes A and Sensenbrenner J 1993 *Am. J. Sociol.* **98** 1320
Reddy S K, Swaminathan V and Motley C M 1998 *J. Mark. Res.* **35** 370

Simas R 1987 *The Musicals No One Came to See* (New York: Garland Publishing Inc.)
Smith R D 2006 *J. Stat. Mech.* **02** 1
Suskin S 1990 *Opening Night on Broadway: A Critical Quotebook of the Golden Era of the Musical Theatre, Oklahoma! (1943) to Fiddler on the Roof (1964)* (New York, NY: Schirmer)
Uzzi B 1997 *Adm. Sci. Q.* **42** 35–67
Uzzi B and Spiro J 2005 *Am. J. Sociol.* **111** 447
Wasserman S and Faust K 1994 *Social Network Analysis* (Cambridge: Cambridge University Press)
Watts D J 2004 *Ann. Rev. Sociol.* **30** 243
Wuchty S, Jones B F and Uzzi B 2007 *Science* **316** 1036
Zhang Z, Zhou S, Wang Z and Shen Z 2007 *J. Phys. A: Math. Theor.* **40** 11863