Strategic Management Journal
Strat. Mgmt. J., **37**: 2204–2221 (2016)
Published online EarlyView in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/smj.2570
Received 5 October 2014; Final revision received 18 February 2016

THE IMPACT OF CONTEXT AND MODEL CHOICE ON THE DETERMINANTS OF STRATEGIC ALLIANCE FORMATION: EVIDENCE FROM A STAGED REPLICATION STUDY

ANINDYA GHOSH,[1,2†] RAM RANGANATHAN,[3*†] and LORI ROSENKOPF[4†]
[1] *Strategy Department, Indian School of Business, Hyderabad, India*
[2] *Department of Entrepreneurship, IESE Business School, University of Navarra, Barcelona, Spain*
[3] *Department of Management, McCombs School of Business, University of Texas, Austin, Texas, U.S.A.*
[4] *Department of Management, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*

Research summary: *Endogenous characteristics of alliance network structure have repeatedly been shown to predict future alliance ties in the strategic management literature. Specifically, the concepts and measures of relational, structural, and positional embeddedness (per Gulati and Gargiulo, 1999), as well as interdependence, are foundational for many studies. We explore these determinants of alliance formation by replicating the baseline analyses of Ahuja, Polidoro, and Mitchell's, 2009* SMJ *article. We examine the impact of empirical choices with respect to time period, underlying data generating model, and industry by isolating each effect in turn. We demonstrate that while geographic similarity and product-market similarity each robustly predict the interdependence effect, the effects of both technological similarity as well as the embeddedness predictors are sensitive to context and/or method.*

Managerial summary: *Our examination of alliance formation in the chemical and semiconductor industries during the 1990s demonstrates how new alliances may be predicted by both the technical, geographic, and product-market fit of potential partners as well as characteristics of each partner's previous network participation. Comparing our results to an earlier study, we find that geographic and product-market similarity predict alliance formation across both industries and time frames while prior ties between partners predict alliance formation only when these industries are less mature. Other network participation indicators generate nuanced effects, which underscore the importance of quasi-replications of alliance formation across industries and time periods in building evidence-based management theories.*

INTRODUCTION

The approach of predicting dyadic alliance ties between firms by examining alliance network structure is deeply entrenched in the strategic management literature (e.g., Ahuja, 2000; Ahuja *et al.*, 2009; Ahuja, Soda, and Zaheer, 2012; Chung, Singh, and Lee, 2000; Garcia-Pont and Nohria, 2002; Gulati, 1999; Li and Rowley, 2002; Rosenkopf, Metiu, and George, 2001; Rothaermel and Boeker, 2008; Wang and Zajac, 2007). All these studies manifest "an endogenous dynamic" (Gulati and Gargiulo, 1999: 1453) in that the pattern of prior alliances in the network, termed "network embeddedness," predicts subsequent alliance formation.

Keywords: alliances; networks; endogeneity; replication; ERGM
*Correspondence to: Ram Ranganathan, Management Department The University of Texas at Austin McCombs School of Business 2110 Speedway Stop B6300 Austin, TX 78712-1282. E-mail: ram.ranganathan@mccombs.utexas.edu
† Authors contributed equally.

According to Gulati and Gargiulo (1999), this embeddedness can be relational (a function of the prior history of ties between the firms in a dyad), structural (a function of a small subset of the network, such as transitivity), or positional (a function of the full network, such as joint centrality). Of course, in these studies, alliance formation is also predicted by characteristics that can be assessed independently from the actual presence of the alliance ties themselves (e.g., Chung *et al.*, 2000; Rothaermel and Boeker, 2008; Wang and Zajac, 2007); constructs such as interdependence[1] derive from the characteristics of each firm in the dyad.

Despite the widespread use of this approach, several empirical concerns exist. First, network effects may be sensitive to the industry environment firms inhabit (Rowley, Behrens, and Krackhardt, 2000) as well as the time period within an industry (Madhavan, Koka, and Prescott, 1998). Second, scholars have called attention to the issue of endogeneity in modeling network ties (e.g., Ahuja *et al.*, 2012; Stuart and Sorenson, 2007). Traditional regression approaches to predicting alliance formation treat the existing alliance network as exogenous, unrealistically assuming that firms do not act strategically to achieve their current network positions (Stuart and Sorenson, 2007). Similarly, such approaches also ignore the real interdependencies between alliance choices across firms such as Gimeno's (2004) demonstration that firms that are competitors in the product-market domain are influenced by each other's alliance choices. Finally, composite network statistics used to represent network embeddedness mechanisms, such as joint centrality, can mask multiple underlying network generating processes, as the measure is aggregated across multiple network levels.

To examine the theoretical and empirical implications of these issues, we replicated the baseline analyses from Ahuja, Polidoro, and Mitchell (APM), published in *SMJ* in 2009. APM demonstrated effects of positional embeddedness via combined centrality; strategic interdependence via technical, geographic, and product-market similarity; and relational embeddedness via previous alliance ties, using the preferred network modeling technique available to strategy scholars at that time. APM's econometric specification is largely comparable to the seminal Gulati and Gargiulo (GG) work,[2] but their work is more amenable to replication by its focus on a clearly bounded industry.[3]

Our replication proceeds in three stages. First, we reproduce the APM estimation strategy on publicly available data in their same industry (chemical), only varying the time period (our data span 1991–2000 while their proprietary data span the prior decade). Next, holding the industry and time period constant, we address model specification and estimation concerns using exponential random graph models (ERGMs), a recent methodological advance that allows the explicit modeling of underlying network formation processes and dependencies in the alliance network data. Finally, we examine the role of industry context by repeating the ERGM analysis for alliance data in the semiconductor industry over the same time period.

Our results generate two important implications for future alliance formation research. First, we demonstrate both robustness and sensitivity of results across industry and time period differences. With regard to robustness, we find that geographic and product-market similarity each consistently predicts alliance formation in both our contexts. In contrast, we find that the maturity of an industry can substantially alter the effect of both technological similarity and previous ties, suggesting the need for future research to incorporate such contingencies and boundary conditions. Second, we demonstrate that traditional measures of positional embeddedness, such as joint centrality, conflate multiple network generating processes. By using recent advances in ERGM techniques to decompose these processes empirically, we are able to suggest

[1]Sometimes termed variously as "homophily," "heterophily," "similarity," or "complementarity" in the alliance literature (Ahuja *et al.*, 2012; Rothaermel and Boeker, 2008).

[2]The main difference between the specification of APM and that of GG is that APM does not include a *Structural embeddedness* measure. We examine this issue when we replicate using an alternative method.

[3]While the GG study represents the most comprehensive articulation of endogenous determinants of alliance formations, two interrelated challenges precluded an effective replication. First, their sample was largely proprietary and used a combination of the Cooperative Agreements and Technology Indicators (CATI) database and several other hand-collected sources. Second, their data set combines data from three different industrial settings (automotive products, new materials, and industrial automation), which do not correspond to well-defined standard industry classifications for which other established and more commonly available data sets could be searched via SIC codes. In comparison, while APM's study also used proprietary data, it was set in the global chemical industry, constituting a relatively unequivocal composition of firms.

DOI: 10.1002/smj

more precise network predictors to refine concepts like positional embeddedness by grounding them in actual tie-forming mechanisms such as preferential attachment and transitivity.

REPLICATION STAGE I: ISOLATING TIME DIFFERENCES

Sample, data sources, and measures

We collected data for all alliances between firms in the global chemical industry for the years 1991–2000 using SDC Platinum, the largest cross-industry data set of strategic alliances. As Schilling (2009) notes, the SDC Platinum data set prior to 1990 is quite sparse, so we chose 1991 as the first year of our alliance observations.[4] We chose a 10-year period to closely match APM's breadth (9 years, 1983–1991). To build the industry-based alliance network, we selected only those alliances that had *both* partner firms belonging to the focal industry (and at least two partners in the focal industry for multiparty alliances). We matched the firms participating in these alliances to company background and financial data available in Compustat—United States, Compustat Global Fundamentals, Bureau Van Dijk (BvD), and OSIRIS, and supplemented this with industry-specific data from Chemical and Engineering News (CEN) and DataQuest. We matched these data using Committee on Uniform Security Identification Procedures (CUSIP) identifiers, Global Company Key (gvkey), stock market tickers, and company names.[5] We further winnowed the set of firms to the largest 150 firms in each industry in each year by revenue.[6,7] We collected patent data for these firms from the NBER patent data set using the approach specified in the Bronwyn Hall Patent Name Matching project to match company names to patent assignees (Hall, Jaffe, and Trajtenberg, 2001). Further, like APM, we include those alliances that involved at least two of the aforementioned 150 firms leading to a final replication sample of 202 strategic alliances between 139 firms. This level of alliance intensity is lower than that reported for the APM sample, which had 97 firms engaging in 338 alliances during their earlier time frame.

We replicated APM's baseline model (see their Model 2, p. 953). APM's model includes the variables *Combined centrality* and *Combined centrality squared* to measure "positional embeddedness" (calculated using the geometric mean of the eigenvector centrality scores of the two member firms in a dyad), the variables *Previous alliances* and *Previous alliances squared* to measure "relational embeddedness" (calculated using the number of alliances the two member firms in a dyad had in the past) and the variables *Technical similarity*, *Technical similarity squared* (calculated using similarity of patents), *Geographic similarity* (calculated using similarity of international country presence), and *Product market similarity* (calculated using similarity of industries firms participate in) to measure "interdependence." APM also include dyadic controls for *Size*, *Liquidity*, *Debt-equity*, *Patents* and *R&D*. Our independent variables for the network embeddedness and technical similarity measures were created in identical fashion to APM's variables. For the geographic and product market similarity, as we had less granular data than APM, we computed a direct similarity (homophily) variable based on country location and primary industry participation. Among the controls, we created identical measures to APMs the only exception was that we used a Debt-to-Asset ratio instead of a Debt-to-Equity ratio due to data limitations. Like APM, we created a binary dependent variable set to 1 if the two firms in the dyad formed an alliance in a year, and 0 otherwise. We used a five-year moving window (years "t−5" through "t−1" for alliance formation year "t") to construct our alliance formation network variables.

Table 1 displays the descriptive statistics and correlations for the APM chemical replication sample. Although our time period is later than APM's, our sample is comparable in terms of descriptive statistics (see APM, Table 1, p. 951). For instance, the

[4]Like APM, we collected additional presample data (in our case, from 1986 to 1990) to build a "Prior Ties" measure for our replication, but we did not use this period to measure the dependent variable.

[5]In instances in which a subsidiary firm was involved in the alliance and the subsidiary firm's financial data were not reported separately, we matched the data to the parent firm's financials using the same techniques. To account for industry mergers, we used the acquiring firm or the merged entity's financials where separate financials were not reported for the constituent firms in the year of the alliances.

[6]APM's sample in comparison consists of the largest 97 firms in the chemical industry.

[7]For firms whose financials were reported in an international currency, we converted to USD using foreign exchange rates available from the Federal Reserve Bank data set (accessed through WRDS).

Table 1. Descriptive statistics (chemical)

	Variables	Mean	S.D.	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Alliance formed	0.007	0.08	0	1	1															
2	Joint centrality	0.024	0.07	0	0.96	0.04	1														
3	Joint centrality squared	0.006	0.03	0	0.91	0.02	0.88	1													
4	Common alliance partners	0.100	0.4	0	5	0.03	0.75	0.64	1												
5	Technical sismilarity	1.410	0.52	0	2	0.01	0.14	0.11	0.13	1											
6	Technical similarity squared	2.260	1.23	0	4	0.01	0.16	0.12	0.14	0.97	1										
7	Geographic similarity	0.224	0.42	0	1	0.04	0.1	0.08	0.09	0.08	0.08	1									
8	Product-market similarity	0.187	0.39	0	1	0.04	0.1	0.09	0.07	0	0	0.09	1								
9	Previous alliances	0.017	0.17	0	6	0	0.35	0.46	0.3	0.05	0.06	0.08	0.07	1							
10	Previous alliances squared	0.030	0.62	0	36	0	0.29	0.49	0.29	0.03	0.03	0.06	0.06	0.82	1						
11	Size	0.414	0.29	0	1	0.03	0.02	0.02	0.02	−0.04	−0.03	0.12	0.06	0.04	0.04	1					
12	Performance	0.040	0.05	0	0.62	−0.01	−0.05	−0.05	−0.03	0.03	0.03	−0.08	−0.03	−0.02	−0.02	−0.04	1				
13	Liquidity	0.710	0.19	0.1	1	0.01	0.09	0.07	0.06	−0.01	0	0.09	0.06	0.04	0.03	0.08	−0.09	1			
14	Debt-equity	0.406	0.3	0	1	0	0.08	0.07	0.05	−0.02	−0.01	0.02	0.1	0.05	0.04	0.16	−0.05	0.09	1		
15	Patents	0.197	0.29	0	1	0	0	0	0	−0.1	0.04	−0.03	−0.03	0	0	0.03	0	0.01	0	1	
16	R&D	0.297	0.29	0	1	0.02	0.04	0.04	0.04	−0.04	−0.03	0.06	0.08	0.05	0.05	0.51	−0.03	0.15	0.12	0.06	1

dependent variable *Alliance formed* has a mean of 0.007 in our sample versus 0.01 in APM's sample while the standard deviation is 0.08 versus 0.10.

Results using identical model

Following APM, we model the data as longitudinal panels, creating a record of the dependent and independent variables for each unique dyad in the sample for each year of the replication period. We pool the observations and use probit regression (probit in Stata) with year dummies and robust standard errors clustered on the dyad to estimate the coefficients.[8] To facilitate comparison, we include APM's baseline results in the first column of Table 2.

Model 1 displays our replication results for APM's pooled probit method. Model 2 adds the structural embeddedness measure (*Common alliance partners*) set forth in the seminal GG model but not included in APM.[9] The last column of Table 2 compares APM's baseline model to our replication results from Model 1. We found an identical positive (marginal effect[10] = 0.004 in both APM and our replication) and significant (p-value < 0.001) effect for *Geographic similarity* and a comparable positive (marginal effect = 0.005 in our replication versus 0.010 in APM) and significant (p-value < 0.001) effect for *Product-market similarity*. We also obtained a comparable positive (marginal effect = 0.036 in our replication versus 0.029 in APM) and significant (p-value = 0.004) base effect of positional embeddedness (*Combined centrality*). Although we obtained a comparable coefficient for the second-order term *Combined centrality squared* (marginal effect = -0.04 in our replication versus -0.03 in APM), it was not statistically significant (p-value = 0.165 in our replication versus p-value < 0.01 in APM). For *Technical similarity*, we obtained a weaker effect size (marginal effect = -0.003) relative to APM (marginal effect = -0.043 in APM) and the coefficient was not statistically significant (p-value = 0.545). For

[8]Following APM, we also confirmed that our results were consistent with random-effects models.

[9]Note that the inclusion of structural embeddedness (which is not significant) fully preserves our Model 1 results, and is included to allows us to directly compare the probit models with subsequent ERGMs, which include better specified measures for all the three embeddedness mechanisms.

[10]Calculated as change in the probability of observing an alliance if the variable is increased by one unit, holding all other variables at their respective sample means.

DOI: 10.1002/smj

Table 2. Replication of APM: chemical industry, probit regressions, 1991–2000[a]

		APM replication (pooled probit)	
Models variables	APM original (pooled probit)	1	2
Alliance formation mechanisms			
Positional embeddedness			
Combined centrality	1.60	2.49	2.04
	(0.32)	(0.87)	(1.07)
Combined centrality squared	−1.45	−2.83	−2.74
	(0.51)	(2.04)	(2.09)
Structural embeddedness			
Common alliance partners			0.10
			(0.10)
Strategic interdependence			
Technical similarity	−2.40	−0.21	−0.21
	(1.13)	(0.35)	(0.35)
Technical similarity squared	1.04	0.11	0.10
	(0.39)	(0.15)	(0.15)
Geographic similarity	0.25	0.28	0.28
	(0.07)	(0.08)	(0.08)
Product-market similarity	0.58	0.33	0.33
	(0.13)	(0.08)	(0.08)
Relational embeddedness			
Previous alliances	0.44	0.74	0.70
	(0.06)	(0.77)	(0.78)
Previous alliances squared	−0.03	−1.23	−1.20
	(0.02)	(0.39)	(0.39)
Dyad level controls			
Size	−0.01	0.37	0.36
	(0.09)	(0.14)	(0.14)
Performance	−2.92	0.40	0.43
	(1.02)	(0.63)	(0.63)
Liquidity	0.21	−0.12	−0.12
	(0.12)	(0.19)	(0.19)
Debt-equity	0.00	−0.26	−0.26
	(0.09)	(0.14)	(0.14)
R&D	0.04	0.10	0.10
	(0.09)	(0.13)	(0.13)
Patents	0.17	−0.04	−0.03
	(0.07)	(0.16)	(0.16)
Constant	−2.71	−2.67	−2.66
	(0.88)	(0.30)	(0.30)
Year dummies	Included	Included	Included
Log likelihood	−1810.68	−650.16	−649.89

[a] Robust standard errors in parentheses.

relational embeddedness (*Previous alliances* and *Previous alliances squared*), while the coefficients we obtained are in the same direction as those of APM's, we did not obtain significance for the main effect (p-value = 0.336) and the coefficient for the second-order effect (*Previous alliances squared*) is negative and significant (marginal effect = -0.018 in our replication versus -0.001 in APM). Figure 1 compares the overall effect from the first and second-order terms for relational embeddedness between our replication and APM's original results.

Since the variable *Previous alliances* can take on only positive integer values, we can conclude from the graph that the net effect of *Previous alliances* on alliance formation likelihood is positive for APM, but negative for our replication. This is driven by the magnitude of the negative coefficient for *Previous alliances squared* (-1.23) in our replication relative to the positive coefficient for *Previous alliances* (0.74). In contrast, APM's coefficients for *Previous alliances* and *Previous alliances squared* were 0.44 and -0.03, respectively.

DOI: 10.1002/smj

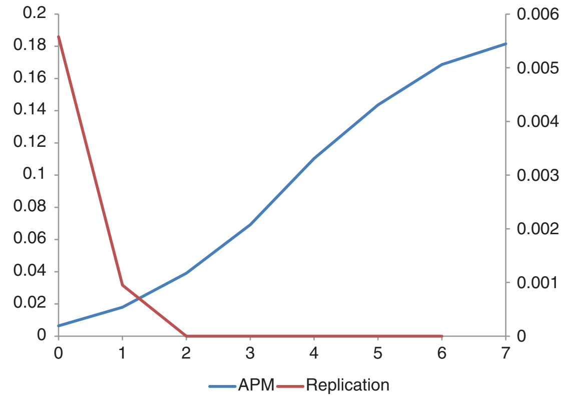


Figure 1. Effect of previous alliances (relational embeddedness) on alliance formation. (*Note*: In the above graph X-axis measures *Previous alliances*, Y-axis (left) measures *Likelihood of alliance formation* for APM's study, Y-axis (right) measures *Likelihood of alliance formation* for our probit replication. All other variables are held at their mean values for the respective samples)

In sum, while our probit replication results in the chemical industry in a later time period are broadly consistent with those of APM, providing confirmatory replications for many of their findings, we are unable to reproduce the effects for interdependence in the technological domain, and we obtain a different effect for relational embeddedness in the alliance network. We will examine how the evolution of the chemical industry may have driven some of the differences we observe in our Discussion section.

Another explanation for replication differences: model choice

While the stage of the industry may account for some differences, we also know that the application of traditional dyadic regression methods such as probit to network data suffers from several shortcomings (Stuart and Sorenson, 2007). First and most well known, these methods assume independence across alliance ties, which many studies, including GG (see p. 1482), acknowledge as an issue (Robins *et al.*, 2007; Stuart, 1998).[11]

Second, predictors derived from one observed network inherently lack an adequate stochastic component because they do not have a corresponding probability distribution function defined at the network level. For example, on observing one distribution of centrality measures across firms from an *actual* alliance network composed of 150 chemical firms, we are unable to assess whether such a distribution is typical across *all possible* structures of 150-firm networks that can emerge, given what we know about the underlying firm, dyadic and network level mechanisms of alliance formation, and after accounting for the possibility of random variation in the network generating process (Holland and Leinhardt, 1970, 1981; Pattison *et al.*, 2000).

Third, the use of endogenous network variables to predict alliance formation further introduces dependencies over time among the observed network variables. The models in Table 2 predict alliance formation using measures that are derived from snapshots of the *same* network—in other words, independent variable measures such as centrality, are derived from a cumulative set of past alliance choices. Despite the theoretical rationale that the current structure of a network shapes its future evolution, the use of such measures in traditional regression models assumes that this structure is independent of the prior network structure.

Finally, it is difficult to isolate different theoretical mechanisms from composite network measures like *Combined centrality*. Centrality is first calculated as a function of the full network for each individual firm, and when the eigenvector centrality score is utilized by researchers, its intent is to proxy for theoretical mechanisms such as power, visibility, or information control (Ahuja *et al.*, 2009; Gulati and Gargiulo, 1999). Yet, at best, this approach conflates multiple underlying mechanisms. For example, are higher centrality scores driven by preferential attachment, where firms with existing alliance ties will attract even more ties (e.g., Powell *et al.*, 2005)? Or by transitivity (Gulati, 1995)? Likely both, and the practice of combining the centrality scores of each firm in the dyad further obscures these underlying mechanisms.[12]

[11]For instance, an alliance between one firm in a focal dyad and a competitor of the other firm in the focal dyad may create competitive pressures that result in a tie forming in the dyad. Dependencies may arise at distant points in the network as well—an alliance between competitors of the two focal firms in a dyad may precipitate an alliance forming between them. In the presence of such interdependencies, traditional methods of predicting alliance formation can lead to incorrect inferences. While various methodologies such as clustering of standard errors or correcting for autocorrelation have been used in prior literature, such statistical methodologies are limited to the firm or to the dyad level and cannot handle complex multilevel tie dependencies that influence alliance formation.

[12]Note that APM (in Model 3) also include dummies distinguishing low-centrality and socially-asymmetric dyads (where one firm

DOI: 10.1002/smj

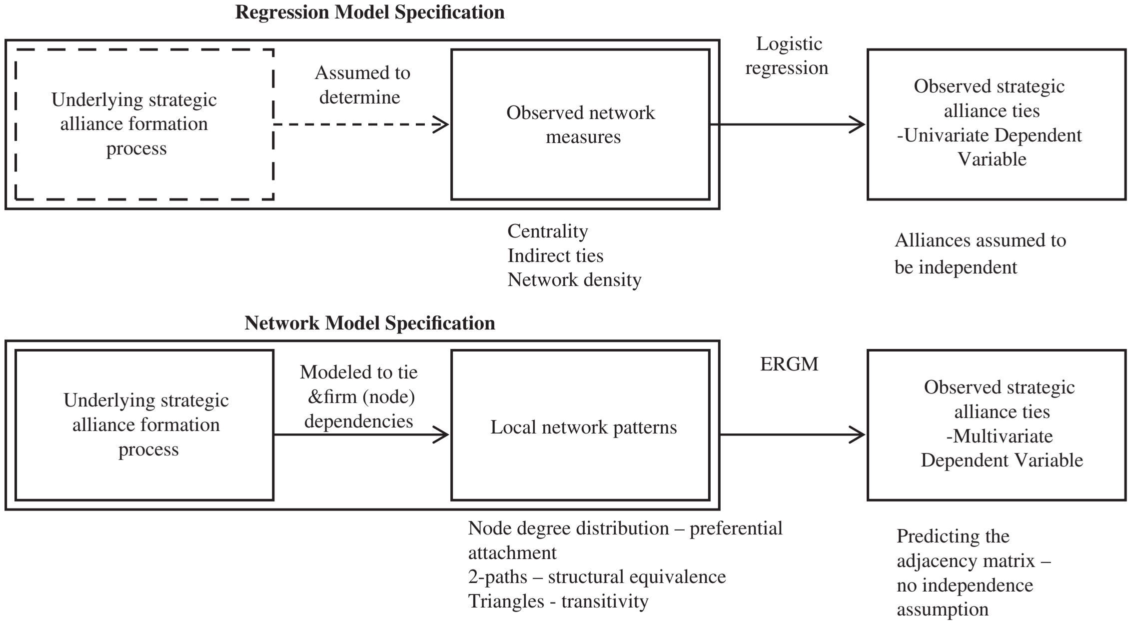


Figure 2. Comparing traditional regression and ERGMs in modeling alliance formation

REPLICATION STAGE II: ISOLATING MODEL DIFFERENCES

To isolate how modeling issues may affect results, we tested a theoretically comparable alliance formation specification using exponential random graph models (ERGMs), a recent advance in social network methodology that overcomes several of these limitations (see Cranmer and Desmarais, 2011; Holland and Leinhardt, 1981; Lusher, Koskinen, and Robins, 2012; Robins *et al.*, 2007; Snijders *et al.*, 2006). Figure 2 illustrates the differences between ERGMs and traditional regression methods. ERGMs[13] are parametric models of networks (c.f. Lusher *et al.*, 2012) defined by identifying relevant local network structural elements that reflect underlying tie formation mechanisms. While traditional regression methods impute endogenous network mechanisms by observing summary network measures over time, ERGMs clearly specify these endogenous processes that generate the observed network structures. ERGMs also shift the level of analysis from the dyad to the network by predicting the entire adjacency matrix of alliances as the dependent variable.[14] This allows the modeling of dependencies between pair-wise combinations in the adjacency matrix as well as dependencies beyond dyads, while also relaxing the stringent assumption of cross-dyadic independence in traditional regression models. Finally, ERGMs are stochastic in nature, treating the observed network as one instantiation within a distribution of possible networks generated by the proposed mechanisms, thus accounting for the possibility of random variation at the network level that traditional methods cannot.

Measuring alliance formation mechanisms in ERGMs

ERGMs allow us to separate the conflated measures used in traditional regression models through precise local network structural specifications that each capture mutually exclusive and collectively exhaustive tie formation mechanisms. Our first task was therefore to select the specific local network structural elements that best correspond to the traditional hypothesized alliance formation mechanisms, summarized in Figure 3. We discuss each of

had high centrality while the other had low). These results, while significant when utilized in place of combined centrality, did not achieve significance when combined centrality was simultaneously included in the model. Similarly, GG also utilized centrality ratio as an explanatory variables, but it was not significant.

[13] Also known as *p** models.

[14] ERGMs generate a joint prediction for all the n*(n – 1)/2 dyads for an n-node network.

DOI: 10.1002/smj

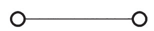
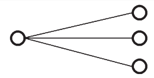
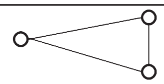
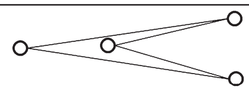
Structural Elements – Tie Dependence		
Local Alliance network structure	***Endogenous Social Process***	***Regression model term name***
	Network density	*edges* – the edges in the observed network
	Positional embeddedness, Preferential attachment	*gwdegree* – geometrically weighted degree distribution in the observed network
	Structural embeddedness (Transitivity/triad closure)	*gwesp* – geometrically weighted edge-wise shared partners
	Structural embeddedness (Structural Equivalence)	*gwdsp* - geometrically weighted dyad-wise shared partners

Figure 3. Local alliance network structural elements in ERGMs

these elements conceptually here and provide additional mathematical details corresponding to their estimation in the Appendix S1.

The most basic network component, *edges*, is analogous to an intercept term in a traditional regression, and captures the base propensity of any alliance tie to form as a function of the count of the alliance ties in the network. The remaining components unpack positional and structural embeddedness mechanisms and are nested hierarchically.

For positional embeddedness, APM as well as GG suggest that a firm's network position allows it to benefit from information about alliance opportunities in the network beyond its immediate partners and also provides a signal of ability and prestige in the collaboration context. APM expected this effect to yield a diminishing advantage with increasing embeddedness; hypothesizing an inverted U-shaped effect and measuring it using *Combined Centrality*, the geometric mean of the eigenvector centralities of the two firms in the dyad. Since this approach overlooks dependencies at different levels and conflates multiple mechanisms, we instead model this mechanism in ERGMs by fitting the degree distribution local network structure which solely models a degree-based preferential attachment mechanism.[15] The network statistic for this structure, called *gwdegree* (geometrically weighted degree distribution), is modeled using a probability distribution function derived from the curved exponential family.[16] Put simply, as a firm's degree (number of alliance ties) increases, this statistic decreases exponentially (Hunter, 2007; Hunter and Handcock, 2006). Thus, with a positive and significant coefficient, the log-odds of an alliance tie increases for all dyadic combinations, but the increase is of smaller magnitude for dyads whose constituent firms already have higher degrees. By its exclusive focus on the degree of each firm[17] and the specification of a probability distribution function that permits inference of the degree based attachment mechanism (Handcock, 2003), this approach removes the conflation of mechanisms inherent in the composite joint centrality measure.

For structural embeddedness, GG hypothesize a transitivity mechanism where firms are more likely to enter into an alliance when they share common partners. Note that the measure for this transitivity mechanism in our probit replication (Table 2, Model 2)—a simple count *of Common alliance partners* between two firms—is highly correlated with the *Combined centrality* measure for positional embeddedness ($r = 0.75$). We overcome this conflation with ERGMs by using triangle configurations[18] corresponding to the theoretical mechanism of transitivity, estimated using *gwesp*

[15]This is measured by the presence of k-stars, that is a central node connected to k others, in the network—for detailed treatment see Hunter (2007) and Hunter and Handcock (2006).

[16]Equation (A1) in the Appendix provides the mathematical basis of the *gwdegree* statistic.

[17]A firm's degree is a network statistic that is likely to be a more relevant local measure that affects alliance formation, than a global network level centrality measure. For instance, it is easily conceivable that firms select other firms as partners if they have many existing alliances (high degree) or vice versa. It is less obvious that they carry out eigenvector centrality calculations while exercising such choices.

[18]Also referred to as triad closure or triangles.

(geometrically weighted edge shared partners), which fits the distribution of the number of triangles also using a curved exponential distribution.

Another limitation of using the traditional count measure is that we are unable to capture the influence of nested substructures (e.g., equivalence considerations within substructures consisting of two, four, or more firms as depicted in Figure 3). ERGMs allow us to model these nested elements using shared partner distribution structures[19] that are estimated using a measure called *gwdsp* (geometrically weighted dyad shared partners)—a statistic of the distribution of shared partner firms by unconnected firms. Thus, the simultaneous use of *gwesp* and *gwdsp* allows us to make stronger inferences of transitivity (Robins, Pattison, and Wang, 2009).[20]

Finally, ERGMs also enable the modeling of node-level and tie-level attributes that additionally motivate alliance formation. We modeled relational embeddedness as a tie attribute by a count of prior alliances for each dyad. We also modeled the interdependence mechanisms by creating tie characteristics that captured similarity on the firms' geography, product-market, and technology vector (based on patent classes) attributes. In contrast, we included all the control measures (*Size, Performance, Debt Ratio, Liquidity*, and *Solvency*) from the APM baseline model as nodal attributes. Here, ERGMs differ slightly from our probit models as they internally calculate the effect for these nodal attributes based on the dyadic sum rather than ratios.

ERGM results

Estimation using ERGMs involves the use of a software module that supports a corresponding implementation of the network structures. The measures corresponding to the ERGM local network structural elements as well other firm and dyad level covariates are specified using a probability function and estimated through maximum likelihood estimation in R (Handcock *et al.*, 2008).[21] We assessed model fits using the Akaike's Information Criterion (AIC) (Akaike, 1998) and log-likelihood statistics, and we employed graphical tests of goodness of fit (Goodreau, Kitts, and Morris, 2009) to visualize the match between the predicted and observed networks.[22]

Table 3 reports our ERGM results and compares them to our probit replication results from Model 2 in Table 2.[23] With ERGMs, a positive coefficient indicates the higher likelihood of presence of a local network element than one would expect by chance, conditional on the rest of the network, whereas a negative coefficient indicates a lower probability of the presence of the structure than expected (Lusher *et al.*, 2012).

A comparison of our ERGM to the probit replication results demonstrates the robustness of the geographic (marginal effect of 0.0006 versus 0.004 in probit) and product market similarity (marginal effect of 0.001 versus 0.004 in probit) predictors, and reveals three major differences. First, whereas *Technical similarity* was insignificant in the prior case, it is positive and significant in the ERGM (p-value 0.023, and marginal effect of 0.0007 versus -0.008). Thus, after modeling the structure of the

[19]While triangle configurations are structural mechanisms that correspond to indirect ties, shared partner distributions model the idea that structural equivalence and multiple connectivity lead to clustered regions in the network.

[20]The benefit of using these geometrically weighted statistics is that they allow us to parsimoniously describe the network data by reducing the number of parameters. For example, the degree fitting term only uses two parameters instead of using (n – 1) parameters where "n" is the highest degree observed in the network. Similarly instead of fitting multiple triangles the *gwesp* statistic uses only two parameters.

[21]The estimation uses the MCMC-MLE procedure in the "ERGM" package, a part of the statnet suite of package for R. Substantively, this approach first involves generating a large number of possible networks that might be observed based on these local network conditions, and then asking whether the network of ties in the focal sample is a likely member of the this family. It is important to reiterate that the generated comparison group of networks is not deterministic but stochastic. Expressed probabilistically, we are ultimately able to determine the probability of observing the sampled network, given the input specifications of different local network elements.

[22]Each plot compares the observed data to one hundred randomly generated simulated networks obtained from the fitted models. This provides a visual sense of the model fit in terms of some key properties of the network such as the degree distribution. All three structural network statistics fit a curved exponential family model that requires the estimation of the decay parameter "α". This is achieved through an iterative process of fitting models for different values and choosing the one that provides the lowest AIC value. Our models report the decay parameter for the best fit model. Additional models and plots are available on request from the authors.

[23]The models shown here are those that provide the best fit (both in terms of the AIC and log-likelihood and visual goodness of fit), therefore we report our results based on it below. We do not report a model with the *gwdsp* term because the ERGM with this term would not converge which is an indication that such structures are not prevalent in the network. This finding also supports our assertion that less localized positional measures obscure rather than clarify mechanisms.

DOI: 10.1002/smj

Table 3. Comparison of probit and ERGM results for chemical industry, 1991–2000[a]

Models variables	Probit replication (Table 2, Model 2)	ERGMs 1	ERGMs 2
Alliance formation mechanisms			
Positional embeddedness			
Preferential attachment (combined centrality or *gwdegree)*	2.04	−1.33	−0.83
	(1.07)	(0.21)	(0.25)
Structural embeddedness			
Transitivity (common alliance partners or *gwesp*)	0.10		0.37
	(0.10)		(0.07)
Strategic interdependence			
Technical similarity	−0.21	0.55	0.50
	(0.35)	(0.23)	(0.23)
Geographic similarity	0.28	0.50	0.46
	(0.08)	(0.15)	(0.15)
Product-market similarity	0.33	0.95	0.90
	(0.08)	(0.14)	(0.14)
Relational embeddedness			
Previous alliances	0.70	0.62	0.47
	(0.78)	(0.38)	(0.35)
Dyad level controls			
Size	0.36	−0.02	−0.02
	(0.14)	(0.03)	(0.03)
Performance	0.43	−0.81	−0.67
	(0.63)	(1.83)	(1.89)
Liquidity	−0.12	−0.09	−0.06
	(0.19)	(0.24)	(0.24)
Debt-equity	−0.26	0.33	0.22
	(0.14)	(0.59)	(0.57)
R&D	0.10	0.53	0.52
	(0.13)	(0.24)	(0.22)
Patents	−0.03	−0.03	−0.02
	(0.16)	(0.02)	(0.02)
Edges	NA	293.50	287.41
		(70.13)	(68.38)
Year	Yes	Yes	Yes
Log-likelihood	−649.89	−937.02	−928.81
AIC	NA	1900.04	1885.62
BIC	NA	1998.63	1991.8

[a] Robust standard errors in parentheses.

network more accurately and accounting for interdependencies in our ERGM, we see that similarity between two firms in the technology domain increases their alliance propensity. Second, while the *Common alliance partners* measure for structural embeddedness in our probit replication was not significant, the *gwesp* measure in our ERGM is positive and significant (p-value < 0.00001). It is important to note that while the other variables of interest did not substantively change with the inclusion of *Common alliance partners* in the probit replication, we did observe a drop in significance for *Combined centrality* (Table 2—compare Models 1 and 2), as these two measures are correlated. Thus, by using a better specified measure in our ERGM, we find support for GG's original hypothesis that structural embeddedness has a positive effect on alliance formation.[24]

[24] This transitivity effect can be interpreted by considering how the probability of the alliance changes when a pair of connected firms increases its number of shared partners by one, *ceteris paribus*.

$\log(p_{after}/p_{before}) = 0.376 \times (1 - e^{\wedge}(-1.3))k = 0.37 \times 0.72\ k$

In other words, it is easiest to complete a triangle when none exists ($k = 0$); such a change leads to an increase of 0.37 on the log-probability scale beyond the effects predicted by other model terms. However, this increase diminishes for each unit increase in k. Thus, completing a two-triangle when a triangle already exists only results in an additional increase of 0.37×0.72; completing

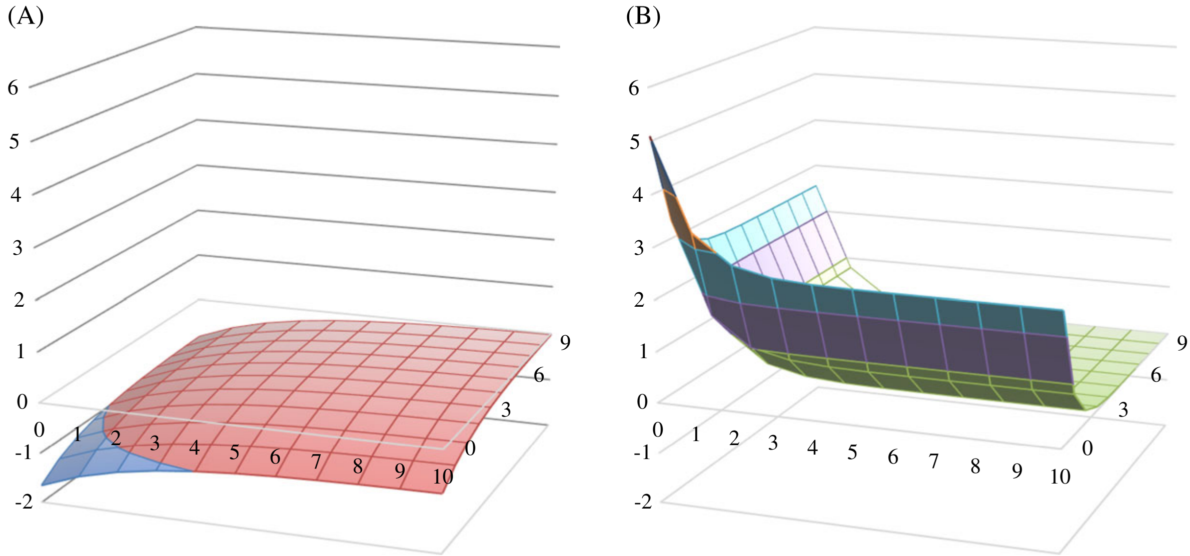


Figure 4. (A) Plot of *gwdegree* in chemical and (B) plot of *gwdegree* in semiconductor. (In the above graphs, the vertical axis measures the change in the log-odds of an alliance forming between two firms and the two axes on the horizontal plane measure the degrees [number of ties] of those two firms.)

Third, the coefficient for the positional embeddedness term *gwdegree* (preferential attachment) is negative and significant (p-value < 0.00001). Recall that in ERGM, a negative coefficient for this term means that firms with higher degrees (high number of alliance partners) have a higher likelihood of entering into further collaborative ties compared to firms with lower degrees, but the increase in this likelihood diminishes as degree increases, as displayed in Figure 4(A).[25] In the graph, the vertical axis is a measure of the change in the log-odds of a tie forming between two firms "i" and "j" if their respective degrees increased from [D_i, D_j] to [D_{i+1}, D_{j+1}]. The axes on the horizontal planes are the degrees of each of the two firms in the dyad. While our probit replication did not obtain an effect for the squared term of *Combined centrality*, the concave curvilinear effect depicted in the graph does suggests evidence for the positional embeddedness mechanism originally posited by APM. As the figure illustrates, the slope is positive, indicating higher odds of alliance formation between the two firms as their combined degree score increases.

All the other alliance formation mechanisms estimated in our ERGM are comparable to our probit replication. We continue to find positive and significant effects for *Geographic Similarity* (p-value < 0.0001) and *Product-market similarity* (p-value 0.000078), and are unable to find significance for the relational embeddedness (*Previous alliances*) measure (p-value 0.13373, marginal effect = 0.23).

REPLICATION STAGE III: ISOLATING INDUSTRY DIFFERENCES

The semiconductor industry as a contrasting context

To isolate and analyze the potential effect of industry context on our results, we replicated our ERGM analyses using semiconductor industry data from the same time frame. The semiconductor industry, with a distributed locus of technological development across firms, a modular set of products, and continuous innovation pressures, offers a rich contrast to the more mature, process-based chemical industry that is less susceptible to

a three-triangle when a two triangle already exists only gives 0.37×0.53 and so on. Thus, as two firms that are already in an alliance, share more and more partners, the propensity to find additional shared alliance partners decreases.

[25] The model estimates the decay parameter to be 1.1, and the coefficient obtained is −0.83. The change statistics for two nodes with degrees i and j is $(1 - e^{(-0.1.1)})i + (1 - e^{(-0.1.1)})j = 0.67i + 0.67j$. So with a coefficient of -0.83, the log-odds of an alliance decrease for all degree values of i and j, but this decrease would be of smaller magnitude when i and j have higher degree.

rapid price-performance improvements based on innovation (Rosenkopf and Schilling, 2007). We followed an identical sampling strategy to the chemical replication (1991–2000, largest 150 firms, within-industry alliances) from the same data sources. Our final sample consisted of 321 strategic alliances.[26]

ERGM results in semiconductor

We developed an equivalent ERGM for the semiconductor industry alliance sample. Table 4 shows the results of the semiconductor ERGM side by side with the results from the chemical ERGM.

Several results are consistent; the effects for structural embeddedness (Transitivity—*gwesp* [p-value < 0.00001]), *Geographic Similarity* (p-value = 0.000126) and *Product-Market Similarity* (p-value < 0.00001) continue to persist across the two industries.[27] Yet others differ. *Technical similarity,* which was significant and positive in chemical, is not significant in semiconductor (p-value = 0.553, marginal effect = 0.02). In contrast, we do find a positive effect for relational embeddedness in semiconductor (*Prior Ties*, p-value = 0.000126, marginal effect = 0.75) which was insignificant in chemical. Perhaps the most striking difference is that the effect for positional embeddedness is reversed in semiconductor (p-value < 0.00001). The corresponding measure in our semiconductor ERGM is positive and significant—recall that a positive coefficient for this term suggests an "anti-preferential attachment" mechanism which is somewhat comparable to a negative effect for a combined centrality measure in traditional regressions. Thus, this demonstrates a tendency for alliance formation in low-degree firms, a mechanism that runs counter to that proposed by the positional embeddedness theory.[28] Figure 4(B) shows the graphical plot of this mechanism and relative effect size. When compared to the corresponding plot for the chemical industry ERGM (Figure 4(A)), the slope of the curve is negative and steeper in the semiconductor ERGM, indicating an effect for positional embeddedness that is stronger and opposite to what we observed in the chemical industry ERGM. Among the control variables, in addition to the strong positive effect for *R&D* (p-value = 0.0084) (as was the case with chemical), there are also positive effects for *Patent Count* (p-value < 0.00001) and *Size* (p-value = 0.0127).

Table 4. Comparison between chemical and semiconductor ERGM results, 1991–2000[a]

Models variables	ERGM for chemical (Table 3, Model 2)	ERGM for semiconductor
Alliance formation mechanisms		
Positional embeddedness		
Preferential attachment (*gwdegree*)	−0.83	2.55
	(0.25)	(0.60)
Structural embeddedness		
Transitivity (*gwesp*)	0.37	0.56
	(0.07)	(0.11)
Strategic interdependence		
Technical similarity	0.50	0.06
	(0.23)	(0.25)
Geographic similarity	0.46	0.56
	(0.15)	(0.15)
Product-market similarity	0.90	0.73
	(0.14)	(0.17)
Relational embeddedness		
Previous alliances	0.47	0.69
	(0.35)	(0.25)
Dyad level controls		
Size	−0.02	0.08
	(0.03)	(0.03)
Performance	−0.67	−0.36
	(1.89)	(0.27)
Liquidity	−0.06	0.14
	(0.24)	(0.20)
Debt-equity	0.22	−0.14
	(0.57)	(1.01)
R&D	0.52	0.11
	(0.22)	(0.04)
Patents	−0.02	0.15
	(0.02)	(0.03)
Edges	287.41	369.50
	(68.38)	(76.72)
Year	Yes	Yes
Log likelihood	−928.81	−841.12
AIC	1885.62	1712.24
BIC	1991.8	1813.53

[a] Robust standard errors in parentheses.

[26] The means and standard deviations for the measures in the semiconductor industry are quite similar to those of the chemical industry displayed in Table 1 (additional correlation tables for the semiconductor industry available on request).

[27] The log-odds of two firms to ally given they operate in the same sectors within the industry increases by 0.70, which is a decrease in the marginal effect from 0.0013 to 0.001. For two firms from the same country, the log-odds are higher by 0.55, which correspond to an increase in the marginal effect from 0.0006 to 0.0007.

[28] In contrast, our examination of alliance formation in the semiconductor context using APM's traditional probit methods failed to yield significant effects for *Geographic* and *Product Market Similarity*. However, they yielded significant results for positional embeddedness (*Combined centrality*) and *Technical similarity.*

DISCUSSION

Our efforts to replicate the baseline model of alliance formation from APM's *Strategic Management Journal* article, and subsequently, isolate effects of changes in method and context, demonstrate the robustness of predictors like geographic and product-market similarity, while illustrating the nuances of the remaining predictors arising from different industry settings, time periods, and empirical methods. Table 5 summarizes our results for ease of comparison and integration. The rows in Table 5 indicate the mutually exclusive theoretical mechanisms identified in alliance formation literature, broadly categorized as endogenous structural network drivers and dyad-specific strategic interdependence factors. Each column of Table 5 represents a replication across successive shifts in the dimensions of time, choice of method and industry context away from the focal APM study. We first discuss the effects of research context shifts (in time period and industry) and then the effect of shifting to ERGMs. In each case, we develop implications both for empirical research and for theory.

Implications of shifts in research context

Davis and Marquis (2005) argued that theoretical mechanisms underlying firms' behavior and relationships between these mechanisms may be reshaped by shifts within industries as well as transitions in the broader social and economic environments in which they are embedded. As seen by comparing Columns A and B (time shift) as well as Columns C and D (industry shift), our replications demonstrated that contextual choices of time period and industry can affect results.

While certain effects persist across the two time periods and two industries we compared (see Rows 5 and 6 for geographic and product-market similarity, as well as Row 3 for structural embeddedness), others vary with the context in which researchers situate their empirical investigation. In particular, the effect of technical similarity varies dramatically across our replications (see Row 4). Focusing first on the time shift (Columns A and B), while APM derived an inverted-U effect, our probit replication in the later time period did not obtain a significant effect. Focusing next on the industry shift (Columns C and D), our ERGMs also yielded different results—a positive effect of technical similarity in chemical, and a nonsignificant

Table 5. Comparison of predicted mechanisms across replications isolating one dimension of change

Replication dimension shift =>theoretical mechanisms		(A) Original APM probit result ('83-'91, Table 2)	(B) Time: '91-'00 (chemical, probit, Table 2)	(C) Model: ERGM ('91-'00, chemical, Table 3)	(D) Industry: semiconductor ('91-'00, ERGM, Table 4)
Network mechanisms	1. Positional embeddedness	Positive + curvilinear	Positive + curvilinear	Positive + curvilinear	Negative + curvilinear
	2. Relational embeddedness	Positive + curvilinear	Negative + curvilinear	Not significant (p = 0.26)	Positive
	3. Structural embeddedness			Positive	Positive
Strategic interdependence	4. Technical	Negative + curvilinear	Not significant (p = 0.55)	Positive	Not significant (p = 0.55)
	5. Geographic	Positive	Positive	Positive	Positive
	6. Product-market	Positive	Positive	Positive	Positive

Curvilinear (second-order) effect size comparable to APM, but not statistically significant (*p*-value = 0.165).
Only second-order effect is statistically significant. First-order effect is not significant (*p*-value = 0.336).

effect in semiconductor. Each of these contrasts may be consistent with prior research about industry and technology life cycles, in that as technical uncertainty reduces, firms will migrate toward more exploitation-based alliances (Lavie and Rosenkopf, 2006). In other words, firms in less mature industries where technical uncertainty is higher (chemical in the 1980s relative to chemical in the 1990s; or semiconductor in the 1990s relative to chemical in the 1990s) are more likely to seek novelty through recombining dissimilar technologies (exploration alliances) rather than through consolidation within the same technological domain (exploitation alliances).

While these contrasts may be instructive, the substantial discrepancies in the effects for *Technology similarity* in our replications raise other empirical implications about variation in the measure across industry contexts. For example, following prior research, APM used a continuous measure obtained through the use of Euclidean distance between firms using patent classes. This measure carries the strong implicit assumption that all classes are equidistant even though distance between classes likely varies both between and within industry.[29] Therefore, future research should utilize measures that account for the nature of the technology classes within a given industry. Empirical techniques that conceptualize the technology space using citation networks and then apply clustering algorithms to identify similar technologies may be useful to obtain better fit between the concept and measure going forward (Sytch, Tatarynowicz, and Gulati, 2012).

We also note differences in results across time and industry for the positional and relational embeddedness predictors (see Rows 1 and 2). Since the use of ERGMs fundamentally changes our measures, theorizing, and interpretations, in this section we will only examine the time period shift in chemical (probit analyses) and leave interpretation of the industry shift (ERGM analyses) for the subsequent section. Specifically, for positional embeddedness, while APM obtained a positive, curvilinear result for combined centrality for their data, only the main effect obtained significance in the chemical industry in our later time period data (see Row 1, Columns A and B). In addition, relational embeddedness no longer predicts alliance formation in our data (Row 2, Columns A and B).

A plausible explanation for these shifts is the evolution of the chemical industry which focused, globalized and consolidated during this time period. Some of the biggest restructuring in the chemical industry since the 1920s occurred during the 1990s (Alperowicz, 2014). Industry concentration increased during the 1991–2000 period with 2,866 mergers, divestitures, and asset sales—in comparison there were only 931 such corporate events in the prior decade.[30] These transitions may have shifted the drivers of alliance formation. The need to tap international markets may have spurred alliances with new local firms,[31] and consequently, the importance of trust (as evidenced through repeated ties) in renewing existing partnerships may have diminished. Increases in ties for leading firms coupled with consolidation would lessen the diminishing impact of centrality originally observed by APM. Consolidation may also have resulted in the acquisitions of existing alliance partners (e.g., Yang, Lin, and Peng, 2011), thus making relational embeddedness a predictor of future acquisitions rather than alliances. Finally, in a maturing industry, firms may have already accrued sufficient information on the performance of partners from previous partnerships, thus making the mere fact of having had a previous alliance a poor predictor of future partnerships (e.g., Holloway and Parmigiani, 2014).[32]

All of these findings reinforce the need to elucidate scope conditions for findings generated in any research context. Researchers should be explicit about comparing their models and contexts against predecessor papers, seeking to isolate changes to one issue and contrasting results. Through

[29]For example, a patent for logic circuitry (e.g., class 326: Electronic digital logic circuitry) will be much more dissimilar (and conceptually more distant) to a patent in materials (e.g., class 505: Superconductor technology: apparatus, material, process), than to a patent for photolithography steps (e.g., class 716: Computer-aided design and analysis of circuits and semiconductor masks).

[30]Source: SDC Platinum, where both acquirer and target firm are in the chemical industry.

[31]This was also the period during which several Asian chemical firms from India and China came into prominence and investment opportunities opened up in Eastern Europe.

[32]Our thesis about industry shifts in this decade in the chemical industry was further substantiated when we re-ran our probit regressions dropping successive years from the beginning of the sample (i.e., 1992–2000, 1993–2000 …)—we lose effects for variables such as Combined Centrality immediately. In contrast, when we ran the regressions dropping years from the end of the period (i.e., 1991–1999, 1991–1998 …), we retain effects till our sample shrinks by ~50 percent. Thus, the industry shift and evolution of alliance drivers likely became more pronounced as time progressed.

DOI: 10.1002/smj

incremental, staged variations, such as we have demonstrated here, researchers can investigate and develop mid-range theories with contingencies rooted in underlying characteristics, such as uncertainty and concentration, of particular industries at particular times.[33]

Implications of model choice

The use of ERGMs, a methodological advance not available earlier to network researchers, to address shortcomings inherent in dyadic regression analyses also enabled us to disentangle traditional network embeddedness measures while allowing for better specification of network formation mechanisms. While combined centrality has become a well-established measure to connote positional embeddedness in our literature, we have argued that it conflates different mechanisms such as transitivity, preferential attachment, and homophily. In contrast, ERGMs more clearly identify and separate the positional and structural characteristics by modeling preferential attachment (represented by *gwdegree*) and transitivity mechanisms (represented by the nested local network structure components *gwesp*, and *gwdsp*) while simultaneously modeling other mechanisms such as interdependence.

Empirically, we observed the effect of shifting from a probit model to an ERGM when our analyses are isolated to the 1991–2000 chemical industry and we vary the modeling technique (compare Table 5, Columns B and C). For network embeddedness predictors, the full specification of all three measures in our ERGM generates a positive result for structural embeddedness (Row 3) as well as a positive curvilinear result for positional embeddedness (Row 1). Thus, separating the traditional *Combined centrality* construct into its constituent parts allows for more precise testing of specific mechanisms. We demonstrate that that the structural mechanisms spurring alliance formation are nodal (degree/visibility) and triangular (transitivity/common ties) in the 1991–2000 chemical industry. Broader network constructs like *gwdsp* do not predict alliance formation. This outcome is consonant with recent work suggesting that arguments premised on multistep information flows between organizations are often unrealistic (Ghosh and Rosenkopf, 2015).

Our most important theoretical implications focus on positional and relational embeddedness. Using the better-specified ERGMs allows us to separate these mechanisms effectively, which vary with industry context (Table 5, Rows 1 and 2, Columns C and D). For positional embeddedness, alliance formation for firms in the chemical industry exhibits preferential attachment, as the coefficient of *gwdegree* is negative; in contrast, the coefficient is positive for the semiconductor industry, suggesting that low degree firms have higher alliance propensities.[34]

How can we explain these divergent effects when a cursory examination of the alliance networks in both industries both present comparable centrality distributions and visible hubs? We must simultaneously consider the differing effects of relational embeddedness, which is positive in semiconductor but not significant in chemical. Semiconductor firms that have low degrees (existing number of alliance partners) have a higher propensity to ally while in chemical high-degree firms capitalize on their positional advantage. Taken together, semiconductor firms reproduce preexisting relationships while seeking novel partners, whereas chemical firms seek partnerships with the most well-established firms with less regard for prior relationships.

These differing industry results likely stem from variation in alliance formation drivers across industries. The semiconductor industry, characterized by rapid technological change and distribution of technological knowledge, manifests higher technological uncertainty and lower concentration. Here, entrepreneurial (read low degree) firms may be able to forge alliances based on the potential value of their technological capabilities, such as the design capabilities of fabless semiconductor firms that came to the fore during our study's time frame. Even

[33]While the practice of pooling multiple industries to obtain generalizability is common, we caution against this approach. In post-hoc analyses, we found that pooling our data across both industries obscured context-specific findings (results available from authors on request).

[34]In additional analyses (available on request), we also compared ERGM results for the semiconductor industry with probit regressions in the same industry (i.e., varying only the method while holding the context and the time period constant). While effects for *Geographic similarity* and *Previous alliances* were significant and consistent, and that for *Technical similarity* was insignificant and consistent between the two models, effects for the other constructs differed dramatically. This further underscores the risks of relying solely on traditional regression methods to test alliance formation mechanisms.

highly embedded incumbent firms in such contexts may frequently need to partner with peripheral or unconnected firms to tap into technologies that could become disruptive or to establish standards of products and processes they develop. In contrast, the chemical setting, characterized by consolidation and global market expansion, manifests lower technological uncertainty and higher concentration, favoring well-established (read high degree) firms.

Of course, while we have identified the benefits of using ERGM techniques, such an approach is not without its limitations. For instance, ERGM is known to be unstable to missing data and model misspecification issues (commonly referred to as "model degeneracy" problems). ERGM statistical packages are also not widely available across platforms (to the best of our knowledge, "R" is the only statistical programming toolkit that provides reasonably documented ERGM routines), and are undergoing frequent changes, thus requiring ERGM users to invest in additional learning. Finally, while ERGM provides a robust method to estimate whether specific tie generation mechanisms underlie an observed network at a point in time, modeling evolutionary processes where these mechanisms may shift over time is less straightforward. However, in balance, as we demonstrate in this article, the benefits of ERGM outweigh both its limitations and those of traditional network modeling methods.

CONCLUSION

Our effort to replicate the baseline analyses of Ahuja *et al.* (2009) explored the impact of shifts in time period, modeling approach, and industry context on the well-established predictors of alliance formation in the strategy literature. As a result, we identified two critical implications for future research in this domain. First, and perhaps not unexpectedly, context does matter; any choice of industry and time period represents a particular stage of an industry lifecycle, and the findings are likely contingent on underlying characteristics of the industry's technology and organization. Future research must acknowledge these issues and seek comparability across studies to assess contingencies effectively.

Second, our use of the newer ERGM technique, not available earlier to network researchers, enabled us to unpack traditional network measures and demonstrate how to improve operationalizations of the network embeddedness constructs. While the original explication of positional, structural, and relational embeddedness set forth the combined centrality measure as the appropriate representation of positional embeddedness, our analyses demonstrated that combined centrality conflates multiple mechanisms. The mutually exclusive and collectively exhaustive predictors available in ERGMs allowed us to use standard predictors of prior ties, common ties, and degree distribution without violating the assumption of independence across data and measures in traditional dyadic regressions. Such an approach, by retaining multiple levels of analysis in the network regression rather than consolidating all measures to the dyadic level, allowed us to identify divergent mechanisms of alliance formation across contexts.

Accordingly, we argue that the ERGM approach should become an essential part of the standard toolkit for future research on alliance formation, and that alliance researchers should seek to replicate extant findings using this method. In so doing, studies become more comparable not just via triangulation across a variety of industries and time periods, but also because the design of ERGMs avoids customized and equivocal composite network measures. Such comparability benefits our field by enabling the development of mid-range theory about underlying constructs such as uncertainty and concentration, which shape alliance network structure across a wide swath of industry and temporal contexts.

REFERENCES

Ahuja G. 2000. The duality of collaboration: inducements and opportunities in the formation of interfirm linkages. *Strategic Management Journal* **21**(3): 317–343.

Ahuja G, Polidoro F, Mitchell W. 2009. Structural homophily or social asymmetry? The formation of alliances by poorly embedded firms. *Strategic Management Journal* **30**(9): 941–958.

Ahuja G, Soda G, Zaheer A. 2012. The genesis and dynamics of organizational networks. *Organization Science* **23**(2): 434–448.

Akaike H. 1998. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer New York.

Alperowicz N. 2014. 1990s: industry restructures, draws apart from oil and pharma. Available at: http://www.chemweek.com/lab/1990s-Industry-restructures-draws-apart-from-oil-and-pharma_63787.html 2015) (accessed 05 Sep 2014).

Chung SA, Singh H, Lee K. 2000. Complementarity, status similarity and social capital as drivers of alliance formation. *Strategic Management Journal* **21**(1): 1–22.

Cranmer SJ, Desmarais BA. 2011. Inferential network analysis with exponential random graph models. *Political Analysis* **19**(1): 66–86.

Davis GF, Marquis C. 2005. Prospects for organization theory in the early twenty-first century: institutional fields and mechanisms. *Organization Science* **16**(4): 332–343.

Garcia-Pont C, Nohria N. 2002. Local versus global mimetism: the dynamics of alliance formation in the automobile industry. *Strategic Management Journal* **23**(4): 307–321.

Ghosh A, Rosenkopf L. 2015. PERSPECTIVE—shrouded in structure: challenges and opportunities for a friction-based view of network research. *Organization Science* **26**(2): 622–631.

Gimeno J. 2004. Competition within and between networks: the contingent effect of competitive embeddedness on alliance formation. *Academy of Management Journal* **47**(6): 820–842.

Goodreau SM, Kitts JA, Morris M. 2009. Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography* **46**(1): 103–125.

Gulati R. 1995. Social structure and alliance formation patterns: a longitudinal analysis. *Administrative Science Quarterly* **40**(4): 619–652.

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

Gulati R, Gargiulo M. 1999. Where do interorganizational networks come from? *American Journal of Sociology* **104**(5): 1439–1493.

Hall BH, Jaffe AB, Trajtenberg M. 2001. The NBER patent citation data file: lessons, insights and methodological tools. National Bureau of Economic Research. No. w8498.

Handcock MS. 2003. Statistical models for social networks: inference and degeneracy. *Dynamic Social Network Modeling Analysis* **126**: 229–252.

Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. 2008. Statnet: software tools for the representation, visualization, analysis and simulation of network data. *Journal of Statistical Software* **24**(1): 1548.

Holland PW, Leinhardt S. 1970. A method for detecting structure in sociometric data. *American Journal of Sociology* **76**(3): 492–513.

Holland PW, Leinhardt S. 1981. An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association* **76**(373): 33–50.

Holloway S, Parmigiani A. 2014. Friends and profits don't mix: the performance implications of repeated partnerships. *Academy of Management Journal* **59**(2): 460–478.

Hunter DR. 2007. Curved exponential family models for social networks. *Social Networks* **29**(2): 216–230.

Hunter DR, Handcock MS. 2006. Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics* **15**(3): 565–583.

Lavie D, Rosenkopf L. 2006. Balancing exploration and exploitation in alliance formation. *Academy of Management Journal* **49**(6): 797–818.

Li SX, Rowley TJ. 2002. Inertia and evaluation mechanisms in interorganizational partner selection: syndicate formation among US investment banks. *Academy of Management Journal* **45**(6): 1104–1119.

Lusher D, Koskinen J, Robins G. 2012. *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Cambridge University Press.

Madhavan R, Koka BR, Prescott JE. 1998. Networks in transition: how industry events (re) shape interfirm relationships. *Strategic Management Journal* **19**(5): 439–459.

Pattison P, Wasserman S, Robins G, Kanfer AM. 2000. Statistical evaluation of algebraic constraints for social networks. *Journal of Mathematical Psychology* **44**(4): 536–568.

Powell WW, White DR, Koput KW, Owen-Smith J. 2005. Network dynamics and field evolution: the growth of interorganizational collaboration in the life sciences. *American Journal of Sociology* **110**(4): 1132–1205.

Robins G, Pattison P, Kalish Y, Lusher D. 2007. An introduction to exponential random graph (p*) models for social networks. *Social Networks* **29**(2): 173–191.

Robins G, Pattison P, Wang P. 2009. Closure, connectivity and degree distributions: exponential random graph (p*) models for directed social networks. *Social Networks* **31**(2): 105–117.

Rosenkopf L, Metiu A, George VP. 2001. From the bottom up? Technical committee activity and alliance formation. *Administrative Science Quarterly* **46**(4): 748–772.

Rosenkopf L, Schilling M. 2007. Comparing alliance network structure across industries: observations and explanations. *Strategic Entrepreneurship Journal* **1**(3–4): 191–209.

Rothaermel FT, Boeker W. 2008. Old technology meets new technology: complementarities, similarities, and alliance formation. *Strategic Management Journal* **29**(1): 47–77.

Rowley T, Behrens D, Krackhardt D. 2000. Redundant governance structures: an analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic Management Journal* **21**(3): 369–386.

Schilling MA. 2009. Understanding the alliance data. *Strategic Management Journal* **30**(3): 233–260.

Snijders TA, Pattison PE, Robins GL, Handcock MS. 2006. New specifications for exponential random graph models. *Sociological Methodology* **36**(1): 99–153.

Stuart TE. 1998. Network positions and propensities to collaborate: an investigation of strategic alliance formation in a high-technology industry. *Administrative Science Quarterly* **43**(3): 668–698.

Stuart TE, Sorenson O. 2007. Strategic networks and entrepreneurial ventures. *Strategic Entrepreneurship Journal* **1**(3–4): 211–227.

Sytch M, Tatarynowicz A, Gulati R. 2012. Toward a theory of extended contact: the incentives and opportunities

for bridging across network communities. *Organization Science* **23**(6): 1658–1681.
Wang L, Zajac EJ. 2007. Alliance or acquisition? A dyadic perspective on interfirm resource combinations. *Strategic Management Journal* **28**(13): 1291–1317.
Yang H, Lin ZJ, Peng MW. 2011. Behind acquisitions of alliance partners: exploratory learning and network embeddedness. *Academy of Management Journal* **54**(5): 1069–1080.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix S1. ERGM statistics, specification and estimation.

DOI: 10.1002/smj