The Role of Social Capital in the U.S. Economy:
Evidence from Employee-level Professional Connections from LinkedIn

Shelley Li (University of Southern California)
Frank Nagle (Harvard Business School)
Aner Zhou (University of Southern California)

Abstract

Using a comprehensive dataset on employee-level professional connections, we construct a firm-level network and calculate connectivity measures for a sample of more than 11,000 publicly traded U.S. firms from 2004 to 2018. We provide descriptive evidence about firm-level connectivity and show that most connectivity measures are positively and significantly correlated with important firm characteristics such as firm size, net income, R&D expense, and intangible assets. Heat maps of pairwise connections between S&P 100 firms show interesting patterns of connection between the largest firms: higher levels of connectivity in the Finance and Information industries, within the same firm, and amongst firms within the same industry or geography. Industry-level network diagrams show that Information, Finance, and Manufacturing are highly connected with each other and with the rest of the economy, and that Manufacturing is the most central industry in the network of large companies (followed by Retail and Information). Our network measures are strongly correlated with yet distinct from existing measures of economic ties between industries. Using the firm-level connectivity measures as a proxy for employee social capital and including it as another input into a firm's production function, we find positive and significant productivity effects of firm connectivity. Our study is the first to use large-scale, individual-level connection data to study firm-level connections, contributing to a more in-depth understanding of firms' strategic positions in the economy and the productivity effect of employee social capital.

Introduction

Connections between firms have long been shown to have important relationships with their economic outcomes (e.g., Podolny, 2001; Uzzi, 1997). In particular, prior studies have examined aspects of firm-level networks based on various types of observed economic ties including supplier-customer relationships at the level of the firm (Atalay et al, 2011; Cohen and Frazzini, 2008) or industry (Antras et al, 2012), patent co-authorship (Breschi and Catalini, 2010), joint ventures (Polidoro et al, 2011), or social and professional connections between high-level firm actors including top executives or board members (Chu and Davis, 2016; Gulati and Westphal, 1999). However, such measures of connectivity only represent the tip of the iceberg in terms of the true depth of connections between firms. In actuality, firms' economic activities are carried out by hundreds or thousands of their employees – whose professional connections could affect the information flow and relational capital upon which business transactions rely. In a world where human capital has steadily become an increasingly important productive input (Ben-Porath, 1967; Black and Lynch, 1996; Romer, 1990), the professional connections of an individual are an extension of their innate human capital, and have important implications for the productivity of the firm who employs them. However, due to lack of available data on such connections, the value of social capital has gone unmeasured in a manner similar to that of intangible technological capital which has been shown to have important productivity impacts, despite being difficult to measure (Corrado, Hulten & Sichel, 2009; Keller et al, 2018).

In this paper, using a wide-ranging professional connection dataset of the professional social network LinkedIn, we build a firm-level network using employee-level connections for over 11,000 U.S. public firms and generate various network/connection measures for each firm-year from 2004 to 2018. These measures help us gain a deeper understanding of how firms are connected to each other in the modern economy, the types of professional connections they have, which firms are most central in the economic network, and how network measures are correlated with firm characteristics. Further, we estimate the impact of firms' connectivity measures on their total-factor productivity.

Existing studies on firm-level connections face various limitations. For example, input-output tables from governmental sources like the Bureau of Labor and Statistics (BLS) and the Bureau of Economic Analysis (BEA) are at the industry level, not firm specific. Firm-level connections based on board interlock capture a very narrow view of firms' social capital by restricting connections to the board level, and represent an indirect measure of connections by assuming a connection if two people serve at the same corporate board. Other firm-level connections measures have been built to apply to only a limited set of companies (e.g., co-patenting networks are limited to companies with patents; customer-supplier relationships are limited to a subset of public companies and their major customers). With the rise of online professional social networks like LinkedIn, almost all public companies have employees on these networks who are connected with other professionals. Therefore, our dataset enables us to build a much more comprehensive network of firms based on a much larger set of employees allowing for a more in-depth picture of the connectivity between firms.

Data

LinkedIn is an online professional social network founded in 2003. By the end of 2018, LinkedIn had 624 million users with more than 27 billion connections worldwide. It also has information about 6.9 million companies and 115 million job positions (as reported in user profiles).

To allow for deeper analysis, we chose to focus on employees of publicly traded firms in the United States, for which there is substantial publicly available information. In particular, to generate our sample, we match public companies in the CompuStat dataset to companies reported on the LinkedIn platform using the following method. For each CompuStat company, we found the closest LinkedIn company using the *dedupe* package[1], a library for the Python programming language that does a fuzzy match between text items and provides a confidence score from 0 to 1 that represents the accuracy of the match. We improved the matching algorithm by cleaning company names using the python *cleanco* package and incorporating industry classifications as a second variable to match on. 72% companies were matched via this algorithmic method, and we manually checked a random subset of these results to ensure accuracy. The remaining U.S. public companies and S&P 500 companies with international headquarters were then matched manually.

Our analyses focus on three samples of progressively larger firms: the full set of U.S. public companies, the set of S&P 500 companies, and the set of S&P 100 companies.

Sample 1 - U.S. public companies. We define a public company as a U.S. public company if its headquarter is in the U.S. From 2004 to 2018, there are 5,570 U.S. public companies on the LinkedIn platform with all the necessary financial and employee connections data for our analyses. In 2018, 2,819 U.S. public companies are in our sample with more than 8 million LinkedIn users

[1] https://github.com/dedupeio/dedupe

employed at those companies, and 1.3 billion connections sent and 1.1 billion connections received by those users.[2] These LinkedIn users represent 22% of the total number of company employees (as reported in the public disclosures of the company and captured by Compustat).

Sample 2 – S&P 500 companies. We define a company as an S&P 500 company if it was included in the Standard & Poor's 500 index in that year, which captures 500 large companies that are representative of the industries in the US economy. Our sample includes 814 out of 826 S&P 500 firms that issued financial statements between 2004 and 2018. In 2018, 505 out of 507 S&P 500 firms are included. In 2018, the S&P 500 companies in our sample cover 7 million LinkedIn users with 1 billion connections sent and 1.2 billion connections received. Their LinkedIn users represent 27% of the total number of reported company employees.

Sample 3 – S&P 100 firms. The S&P 100 is a subset of the S&P 500 that include the largest companies (by market capitalization) of the S&P 500. The S&P 100 represents roughly 60% of the market cap of the S&P 500 and roughly 50% of the market cap of all US public firms. Our sample includes all the 163 S&P 100 firms that issued financial statements between 2004 and 2018. In 2018, 100 S&P 100 firms are included. There are 4 million LinkedIn users working for the S&P 100 companies at the end of 2018 with 0.6 billion connections sent and 0.7 billion connections received. Their LinkedIn users represent 28% of the total number of reported company employees.

LinkedIn coverage by industry. Although LinkedIn is the most widely-used professional network in the United States, as alluded to above, not all employees of all companies have a profile on it. Therefore, we consider LinkedIn coverage as the percentage of a firm's employees who are present on LinkedIn and calculated it by dividing the number of positions on LinkedIn for a given firm (as reported by the employee/LinkedIn user) in that year by the total number of employees reported by the firm (captured in Compustat) for the subset of firms with non-missing Compustat employee data. LinkedIn coverage varies significantly across industries, with Information, Finance, and Professional, Scientific, and Technical Services holding the highest coverage in the U.S. public firm samples. As average company sizes increase, overall coverage only increases slightly from 22% in Sample 1 to 28% in Sample 3. However, coverage in certain industries increases

[2] On LinkedIn, connections between users are only formed after one user sends a connection request to another user, and that user accepts the request. Here, we report the number of connection requests a user sends and those that they receive separately to allow for a more granular understanding of how the connections are formed.

significantly with size, such as in Mining, Quarrying, and Oil and Gas Extraction, with its coverage increasing from 36%, to 40% and 54%. Other high-coverage industries also seem to increase the coverage with size (Information industry's coverage increases to 52%, while Real Estate's coverage increases to 76%, as we move from the sample of US public firms to the S&P 100 sample).

Using Sample 1 (all public US firms), we counted the total number of connections each employee (of the sample firms) has, removed "extreme" connectors (i.e. those with more than 10,000 connections, a very small proportion of the dataset[3]), and generate a range of individual-level connection measures.

Results

Firm-level connection measures

Using the dataset described above, we build a corporate network based on employee connections. Two firms are connected if there is at least one individual-level LinkedIn connection between these two firms. The strength of the connection between two firms is measured by the natural log of the number of connections between them[4]. Since what we try to measure is the "importance" of the firm to the network, we focus on eigenvector centrality as the main network measure in this paper. Eigenvector centrality measures the influence of a node in the network (Newman, 2008). Companies with strong connections to more central firms have higher eigenvector centrality. In addition to centrality measures, we also generated other firm-level connection measures to capture the extent and type of firm connections. Table 1 shows the descriptive statistics for these firm-level connectivity measures in 2018. In our sample of U.S. public firms, on average, there are 2,909 employees on LinkedIn, with an average of 314 connections per employee (and a total of almost 850,922 connections per firm). The diversity of an individual's network has previously been shown to impact economic outcomes (Eagle, Macy, and Claxton, 2010). Therefore, we dig into the type of connections individuals have, and find that, on average, 10% of all connections are internal connections (i.e. connections made with employees currently in the same firm), 26% are connections with people who have the same education level, 23% are connections with those who have the same experience level, 67% are connections with those who have the same job

[3] Out of the more than 9 million employees in our sample, only fewer than 3000 were dropped due to "extreme" number of connections.

[4] We use the natural log of the number of connections, rather than the number of connections between firms, because the existence of power law is usually observed in social networks.

function, 12% are connections with those who have gone to the same school, and 19% are connections with those who used to work in the same institution.

--Insert Table 1 --

Based on the eigenvector centrality measure, we generated a list of the top 100 most central US public companies in the economy (as of 2018). The list is in Table 2. It is important to recognize that the dataset is not the complete set of all individuals, at all companies, and all of their connections, as discussed above. However, it is interesting to note the different types of companies and how central they are to the economy as captured in our data. In particular, tech and finance companies dominate the top 20 (which is not surprising given those are the two most well represented industries on LinkedIn), but we see companies from industries including healthcare, consumer goods, retail, and real estate breaking into the top 30. The global clustering coefficient for all US public companies is 0.79, which indicates that the firms are fairly tight knit as the measure goes from zero to one. Examining the global clustering coefficient by industry indicates a small degree of variation across industries, with NAICS 81 (Other service (except public admin)) having the lowest clustering coefficient at 0.70 and NAICS 11 (agriculture, forestry, fishing and hunting) having the highest clustering coefficient at 0.88.

--Insert Table 2 --

Correlation between connection measures and firm characteristics

In Table 3, we show the correlation between the firm-level connectivity measures and important firm characteristics. All numbers in **Bold** (*italic*) represent statistical significance at the 1% (5%) level. Eigenvector centrality and total number of firm connections are positively and significantly correlated with all measures of firm size (revenue, assets, number of employees), net income, and intangible assets. Measures representing centrality and total number of firm connections (in general or of a certain type) are positively and significantly correlated with R&D expense. Comparing measures scaled by firm size (assets), we find that the average number of connections is positively and significantly correlated with R&D intensity. Eigenvector centrality is positively and significantly correlated with ROA (return on assets) (at the 5% significance level).

---Insert Table 3 ---

Pairwise Connections between the largest firms

In Figure 1, we present a heat map representing the 2018 S&P 100 firm-to-firm network. The color of the box represents the number of realized connections initiated from one company (horizontal) going to the other (vertical). Therefore, the colorings are not perfectly symmetric as even if two companies are tightly connected, one of them may be initiating more of the connections than the other. For the boxes where the same company is both on the vertical and horizontal, the color represents the number of internal connections at the company itself. Firms are grouped based on industry (NAICS), and within each industry, ordered based on size. Each row provides the composition of connections made by the firm. The number of connections has been log-transformed to the scale on the right and darker colors represent larger number of connections. The figure suggests: (1) there are a high number of connections within firms and amongst firms in the same industry. (2) The Finance and Insurance industry, and the Information industry show the highest level of connectivity, which is partially driven by the higher percentage of employees in those industries that are on LinkedIn. (3) There is a geographic element to the connections as well. For example, companies that are headquartered in Minnesota (e.g., UnitedHealth Group, Target, 3M, and US Bancorp) are well connected to each other, even though they are in different industries. This adds support to prior literature that has discussed the importance of managerial talent to the growth of such clusters (Shaver, 2018). We generated the same heat map for each year going back to 2004 in Figure A1 to Figure A14 (*Appendix*). Looking at heat maps over the years, we see that the connections became overall denser and denser. Even in the very early years (2004, 2005), we can already see a strong diagonal line (indicating connections with those in the same firm) and a denser concentration in the Information industry.

---Insert Figure 1 --

Industry-level network diagrams

Figures 2, 3, and 4 show the industry-level pairwise connections for S&P 100, S&P 500, and all US public firms at the end of 2018, respectively. The darker the line between two industries, the more connections these two industries have to each other and the size of the node represents the centrality of the industry in the overall network. Information, Finance and Insurance, and Manufacturing are highly connected overall and with each other. Manufacturing is the most central industry in this network (which is partially a result of it being the most heavily represented industry in the S&P 100), while Retail Trade and Information are the second and third most central industry, respectively. It is interesting to note that although companies in the Information industry are very well connected in the network of S&P 100 firms (as seen in Figure 1), the industry as a whole is slightly less central compared to traditional industries like Manufacturing and Retail Trade. We generated the same diagrams for each year going back to 2004 in Figure A15 to Figure A53 (in the Appendix). Similar to the heat maps, we see that the connections became denser over the years. In the very early years (2004, 2005), Manufacturing, Information, Finance and Insurance, were already the most central industries, and their centralities relative to other industries decrease over time.

--*Insert Figure 2* ---

--*Insert Figure 3* ---

--*Insert Figure 4* ---

Comparison with other firm-level network measures

Comparing the firm-level network measures from the sample of all US public firms to the 2018 BEA/BLS input-output tables at the same industry level (which measures sales from one industry to another), we find that the natural log of the number of connections between two two-digit NAICS industries is positively correlated with the natural log of the input-output between the same two industries, with a coefficient of 0.33 and a significance level of 0.00. This indicates that our measure of firm-level connectivity strongly correlates with measures of economic ties between industries, but measures something distinct that is not fully captured by flows of sales between industries.

We further compare our network measures with connection measures calculated with data on boards of directors. In 2018, only 0.2% of connected company pairs based on our sample data (of US public firms) have interlocked board members, suggesting that the firm connectivity measures calculated with board interlock data only apply to a very small portion of U.S. public firms. Among these 0.2% company pairs, the natural log of the number of connections between these two firms is not significantly correlated (at the 5% level) with the natural log of the number of interlocked board members, with a coefficient of -0.03 and a significance level of 0.06. This suggests that our measures based on employee-level professional connections can provide insights not covered by traditional network measures based on top-level actors in firms.

Productivity impact of firm-level connectivity

It has long been known that human capital is a critical input into the productivity of the firm (Romer, 1990). However, our data allows us to more deeply consider the role of employee *social capital* in firm productivity. To examine this relationship, we consider various measures of social capital, as captured in the LinkedIn data metrics discussed above. We use these measures in a total factor productivity (TFP) framework with firm-level employee social capital as an input into the production function in addition to the standard inputs of capital and labor. This is consistent with prior literature that has used TFP to calculate the productivity impact of inputs that are traditionally difficult to measure (Bloom and Van Reenen, 2002; Brynjolfsson and Hitt, 1996; Nagle, 2019). Such analysis starts with a log-transformed Cobb-Douglas equation and adds the variable of interest into the equation, yielding our primary regression equation:

$$\ln(Y_{it}) = \alpha \ln(K_{it}) + \beta \ln(L_{it}) + \gamma SC_{i,t-1} + \varepsilon_{it} \tag{1}$$

In this equation, Y_{it} is the output of firm *i*, in year *t*, as measured by real sales[5]. K_{it} is the amount of capital invested by firm i, in year *t*[6], and L_{it} is the number of employees at the firm. $SC_{i,t-1}$ is the level of social capital as measured by the various measures discussed in Table 1, including eigenvector centrality, total number of connections, average number of employee connections, total number of internal connections, and average similarity. Due to high-levels of correlation between these measures, each is only used one at a time in the regression. We use the lagged values for these connectivity measures, rather than the current year, to capture the fact that social capital primarily comes from existing ties, rather than newly formed ones. However, results with current

---

[5] REVT in Compustat deflated by implicit price deflators for gross domestic product (GDP) from the Bureau of Economic Analysis (BEA).

[6] Calculated as the sum of capital stock when the firm first became publicly traded and new investments in each year until year *t*. Capital stock is measured as the firm's net total property, plant, and equipment (PPENT in Compustat) when the firm became public deflated by the GDP deflator for private nonresidential fixed investment from the BEA for that year. The net investment for each year is the difference between in PPENT from *t-1* to *t*, deflated by the appropriate GDP deflator for private nonresidential fixed investment from the BEA.

year social capital are generally similar. Importantly, although our descriptive statistics above only show firm connectivity at the end of 2018, we use historical data that go as far back as 2004 for the TFP analysis. This allows us to construct a large dataset of firm-years that considers differences in social capital both across firms, and within firms over time (using firm fixed effects). All models include a control for year and NAICS2 industry or firm fixed effects to account for variance in LinkedIn coverage and time trends. For measures measuring the number of connections, we take the natural log of the measure. Although most of the values of *SC* are self-explanatory, *Average Similarity* requires more explanation. This measure is the average value for each firm-year across five specific types of connections, i.e. connections with people at the same education level, experience level, with the same job function, shared the same school, or shared at least one prior workplace. Therefore, in aggregate *Average Similarity* measures how homogeneous the employee connections are for a given firm year. Homogeneity could limit the diversity of ideas these employees are exposed to; but it could also represent stronger ties based on shared backgrounds or experiences. When using *Average Similarity*, we control for the total number of employee connections at the firm to help account for the volume of connections.

Table 4, Column 1 shows coefficients from our baseline regression model which includes industry and year fixed effects. Column 2 uses firm fixed effects. The values shown in the table indicate the coefficients γ from equation 1 for a given social capital measure (i.e. they show coefficients from many regressions, each with a different social capital measure). We can see that eigenvalue centrality and various other firm-level connection measures have positive productivity effects. In general, the productivity effects of these connectivity measures become smaller in magnitude when we control for firm fixed effects and the statistical significance drops below the 10% threshold for all measures except eigenvector centrality and average similarity indicating those are likely the strongest predictors. The positive coefficient on the total number of connections with those who have the same job functions even became statistically significant after including the firm fixed effects. These results show that the amount of employee social capital at the firm has positive productivity effects. Using the values from Table 1, we can interpret the coefficients as meaning that for the average firm in the sample, increasing the total number of connections employees of the firm have by 2,586 leads to an increase of real sales of \$2.1M. Controlling for the total amount of employee social capital, we find that the more connections share similar characteristics (i.e., average similarity), the higher the productivity effect from these types of social capital.

In addition to using the TFP calculation, we also considered the role of social capital in firm valuation as captured by the ratio Tobin's Q. Tobin's Q is the ratio of a firm's market value to its replacement value and is calculated as follows:

$$\text{Tobin's Q} = \frac{(Equity\ Market\ Value + Liabilities\ Market\ Value)}{(Equity\ Book\ Value + Liabilities\ Book\ Value)}$$

In this equation, $Equity\ Market\ Value$ is the product of shares outstanding (Compustat item commonshr) and share price at the end of the year (Compustat item prcc_f). $Liabilities\ Market\ Value$ equals the book value of assets (Compustat item at) minus the book value of equity (Compustat item ceq) and deferred taxes (Compustat item txdb).

Prior research has used Tobin's Q to estimate the effect of firm intangible resources on firm valuation (Villalonga, 2004; Hall, Jaffe, and Trajtenberg, 2005). We use Tobin's Q as the outcome variable in the following regression as a way to better measure the impact of social capital on firm value:

$$\ln(Tobin's\ Q_{it}) = \alpha \ln(Assets_{it}) + \beta \ln(R\&D_{it}/Assets_{it}) + \delta ROA_{it} + \rho Leverage_{it} + \gamma SC_{i,t-1} + \varepsilon_{it} \quad (3)$$

We use the log of Tobin's q as the outcome variable in the regression to adjust for the strong positive skewness in Tobin's Q. SC represents the variables of main interest, the social capital measures we used in previous analyses. We control for firm size (*ln(Assets)*), R&D intensity ($ln(R\&D_{it}/Assets_{it})$), firm profitability ($ROA$) and leverage (*Leverage*).

Columns 3 and 4 of Table 4 show the results of this estimation, and again individual measures of social capital are included one at a time in separate regressions. The results show a similar relationship as that with TFP: in the OLS (column 3), all measures of social capital are positive and significant at the 1% level, while in the firm fixed-effect model (column 4), only eigenvector centrality and average similarity are significant. These results indicate that (1) the more central a firm is in the firm-level network of employee social capital, the higher the firm is valued by the market; (2) the more these employee connections are made with those who share similar backgrounds or experiences, the higher the firm's market valuation.

---*Insert Table 4* ---

Table 5 digs deeper into the eigenvector centrality measure of social capital to better understand what types of relationships are driving the primary finding. In particular, we consider the job level and condense jobs into three buckets as follows. Individuals that categorized as owners, partners, or CXOs are classified as “high”, individuals categorized as senior, manager, director, or vice president are classified as “medium”, and individuals categorized as unpaid, training, or entry positions are classified as low. This classification allows us to better understand whose connections are driving the results shown in Table 4 – people at the top, middle, or base of the corporate hierarchy. Using these classifications, we re-calculate each firm’s eigenvector centrality using only the connections of individuals in a given bucket with other individuals in the same bucket (e.g., we first calculate all firms’ eigenvector centrality as if the company only had the employees that are classified as high, then we do the same thing for medium and low). Table 5 shows these results for both the TFP and Tobin’s Q calculations. Again, each measure of eigenvector centrality is included in a regression by itself, but the results are condensed to save space. Across the board, the results show that it is the medium to medium connections that have the largest impact on productivity and valuation. This adds additional weight to our earlier arguments that using high level measures of firm connectivity (like board interlock, or co-patenting) do not capture the true richness of relationships between companies.

--*Insert Table 5* --

Conclusion

In this paper, we use a novel dataset capturing employee-level professional connections between companies to shine new light on a question that economists and sociologists have wrestled with for some time. We show that there is a great deal of connectivity between US public firms, but that this connectivity varies on numerous dimensions. Further, we show that employee social capital (as measured by connections on LinkedIn) has a measurable impact on firm productivity. Although this dataset provides much more granular insights into the connections between firms than existing data, it is important to note its limitations, particularly that the connections seen on LinkedIn are not a random sample of all people, and over-represent certain industries, white-collar workers, and larger companies (especially given our focus on public companies). Despite these limitations, we believe using employee-level connections as a new measure of firm connectivity allows a more granular analysis than existing measures and will allow scholars to revisit earlier studies of firm networks to explore new avenues through which this connectivity may impact productivity and other social and economic outcomes.

References and Notes

Antràs, P., Chor, D., Fally, T., & Hillberry, R. (2012). Measuring the upstreamness of production and trade flows. *American Economic Review*, *102*(3), 412-16.
Atalay, E., Hortacsu, A., Roberts, J., & Syverson, C. (2011). Network structure of production. *Proceedings of the National Academy of Sciences*, *108*(13), 5199-5202.
Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, *75*(4, Part 1), 352-365.
Black, S. E., & Lynch, L. M. (1996). Human-capital investments and productivity. *The American Economic Review*, *86*(2), 263-267.
Bloom, N., & Van Reenen, J. (2002). Patents, real options and firm performance. *The Economic Journal*, *112*(478), C97-C116.
Breschi, S., & Catalini, C. (2010). Tracing the links between science and technology: An exploratory analysis of scientists' and inventors' networks. *Research Policy*, *39*(1), 14-26.
Brynjolfsson, E., & Hitt, L. (1996). Paradox lost? Firm-level evidence on the returns to information systems spending. *Management Science*, *42*(4), 541-558.
Corrado, C., Hulten, C., & Sichel, D. (2009). Intangible capital and US economic growth. *Review of Income and Wealth*, *55*(3), 661-685.
Chu, J. S., & Davis, G. F. (2016). Who killed the inner circle? The decline of the American corporate interlock network. *American Journal of Sociology*, *122*(3), 714-754.
Cohen, L., & Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance*, *63*(4), 1977-2011.
Eagle, N., Macy, M., & Claxton, R. (2010). Network diversity and economic development. *Science*, *328*(5981), 1029-1031.
Gulati, R., & Westphal, J. D. (1999). Cooperative or controlling? The effects of CEO-board relations and the content of interlocks on the formation of joint ventures. *Administrative Science Quarterly*, *44*(3), 473-506.
Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of economics*, 16-38.
Keller, S., Korkmaz, G., Robbins, C., & Shipp, S. (2018). Opportunities to observe and measure intangible inputs to innovation: Definitions, operationalization, and examples. *Proceedings of the National Academy of Sciences*, *115*(50), 12638-12645.
Nagle, F. (2019). Open source software and firm productivity. *Management Science*, *65*(3), 1191-1215.
Newman, M. E. (2008). The mathematics of networks. *The new palgrave encyclopedia of economics*, *2*(2008), 1-12.
Podolny, J. M. (2001). Networks as the pipes and prisms of the market. *American Journal of Sociology*, *107*(1), 33-60.
Polidoro Jr, F., Ahuja, G., & Mitchell, W. (2011). When the social structure overshadows competitive incentives: The effects of network embeddedness on joint venture dissolution. *Academy of Management Journal*, *54*(1), 203-223.
Romer, Paul M. (1990). "Human capital and growth: Theory and evidence," Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 32(1), pages 251-286, January.
Shaver, J. M. (2018). *Headquarters Economy: Managers, Mobility, and Migration*. Oxford University Press.

Uzzi, B. (1997). Social structure and competition in interfirm networks: The paradox of embeddedness. *Administrative Science Quarterly*, 35-67.
Villalonga, B. (2004). Intangible resources, Tobin'sq, and sustainability of performance differences. *Journal of Economic Behavior & Organization*, *54*(2), 205-230.

DRAFT

Table 1 – Descriptive Statistics for Firm-level Measures (2018)

Measures	No. Obs.	Mean	Std. Dev.	Min	25%	50%	75%	Max
Eigenvector centrality	2,819	0.013	0.013	0.000	0.002	0.009	0.020	0.068
Number of employees on LinkedIn	2,819	2,908.61	11,511.01	1	31	230	1,330	285,561
Total number of employee connections	2,819	850,922.20	3,703,151.0	1	8,946	66,263	392,957	90,029,070
Average number of connections	2,819	313.83	193.62	1	200.60	281.12	388.66	2,404.00
Total number of employee internal connections	2,819	84,723.23	449,854.80	0	151	3,370	28,325	12,296,470
Total number of employee connections with people who have the same education level	2,819	220,671.60	1,008,519.00	0	1,925	15,603	93,942	22,624,620
Total number of employee connections with people who have the same experience level	2,819	195,514.70	901,609.00	0	1,619	13,287	80,87	23,601,360
Total number of connections with people who have the same job functions	2,819	573,651.70	2,595,205.00	0.00	5,789.50	43,229.00	251,905.00	61,984,910.00
Total number of connections with people who have gone to the same school	2,819	104,701.70	483,182.00	0.00	805.00	6,812.00	41,695.00	10,427,440.00
Total number of connections with people who have worked at the same institution.	2,819	162,393.10	814,611.20	0.00	875.50	9,502.00	60,138.50	20,181,020.00
Exp(Average Similarity)	2,819	251,386.50	1,153,355.00	0.00	2,181.50	17,833.00	105,244.40	27,460,640.00
Revenue (in millions)	2,819	5,131.65	20,464.27	0.00	72.62	574.98	2,680.50	511,729.00
Assets (in millions)	2,819	12,563.96	85,882.64	0.04	188.25	1,114.84	5,055.10	2,622,532.00
Number of employees (in thousands)	2,819	12.93	57.05	0.00	0.22	1.48	7.44	2,200.00
ROA	2,819	-0.11	1.09	-28.83	-0.02	0.04	0.10	1.75
R&D Expense	1,512	256.71	1,396.85	0.00	1.64	19.60	80.85	28,837.00
Net Income (in millions)	2,819	413.37	2,224.25	-22,355.00	-8.90	18.98	166.93	59,531.00
R&D/Assets	2,819	0.07	0.25	0.00	0.00	0.00	0.05	8.41
Intangible assets (in millions)	2,787	2,090.26	10,316.64	0.00	1.58	74.07	704.34	310,197.00
Real sales (in millions)	2,819	4,647.39	18,533.12	0.00	65.76	520.72	2,427.55	463,438.72
Tobin's Q	2,819	2.98	17.56	0.32	1.03	1.40	2.39	701.06

Table 2 Top 100 Firms with Highest Eigenvector Centrality in 2018

#	Stock Ticker	Company Name	Eigenvector Centrality
1	AMZN	AMAZON.COM INC	0.068
2	IBM	INTL BUSINESS MACHINES CORP	0.067
3	MSFT	MICROSOFT CORP	0.066
4	ORCL	ORACLE CORP	0.065
5	GOOG	ALPHABET INC	0.065
6	WFC	WELLS FARGO & CO	0.062
7	AAPL	APPLE INC	0.062
8	JPM	JPMORGAN CHASE & CO	0.061
9	CSCO	CISCO SYSTEMS INC	0.060
10	C	CITIGROUP INC	0.059
11	CRM	SALESFORCE.COM INC	0.059
12	T	AT&T INC	0.058
13	BAC	BANK OF AMERICA CORP	0.058
14	ADP	AUTOMATIC DATA PROCESSING	0.057
15	FB	FACEBOOK INC	0.057
16	CTSH	COGNIZANT TECH SOLUTIONS	0.056
17	HPE	HEWLETT PACKARD ENTERPRISE	0.056
18	MS	MORGAN STANLEY	0.056
19	UNH	UNITEDHEALTH GROUP INC	0.055
20	INTC	INTEL CORP	0.054
21	JNJ	JOHNSON & JOHNSON	0.054

22	COF	CAPITAL ONE FINANCIAL CORP	0.054
23	PEP	PEPSICO INC	0.054
24	HON	HONEYWELL INTERNATIONAL INC	0.053
25	WMT	WALMART INC	0.053
26	GS	GOLDMAN SACHS GROUP INC	0.053
27	TGT	TARGET CORP	0.053
28	IT	GARTNER INC	0.052
29	CBRE	CBRE GROUP INC	0.052
30	GM	GENERAL MOTORS CO	0.052
31	BA	BOEING CO	0.052
32	PFE	PFIZER INC	0.051
33	LMT	LOCKHEED MARTIN CORP	0.051
34	ABT	ABBOTT LABORATORIES	0.050
35	AXP	AMERICAN EXPRESS CO	0.050
36	VZ	VERIZON COMMUNICATIONS INC	0.050
37	TSLA	TESLA INC	0.050
38	NKE	NIKE INC -CL B	0.049
39	DXC	DXC TECHNOLOGY CO	0.049
40	NOC	NORTHROP GRUMMAN CORP	0.049
41	VMW	VMWARE INC -CL A	0.049
42	GE	GENERAL ELECTRIC CO	0.049
43	USB	U S BANCORP	0.049
44	CVS	CVS HEALTH CORP	0.049
45	TMO	THERMO FISHER SCIENTIFIC INC	0.049

46	SYK	STRYKER CORP	0.049
47	PNC	PNC FINANCIAL SVCS GROUP INC	0.049
48	HD	HOME DEPOT INC	0.049
49	ALL	ALLSTATE CORP	0.048
50	ADBE	ADOBE INC	0.048
51	CMCSA	COMCAST CORP	0.048
52	F	FORD MOTOR CO	0.048
53	SBUX	STARBUCKS CORP	0.048
54	HPQ	HP INC	0.048
55	UPS	UNITED PARCEL SERVICE INC	0.048
56	PG	PROCTER & GAMBLE CO	0.048
57	RTN	RAYTHEON CO	0.047
58	MMM	3M CO	0.047
59	BAH	BOOZ ALLEN HAMILTON HLDG CP	0.047
60	KO	COCA-COLA CO	0.046
61	LOW	LOWE'S COS INC	0.046
62	MRK	MERCK & CO	0.046
63	MAR	MARRIOTT INTL INC	0.046
64	KELYA	KELLY SERVICES INC -CL A	0.046
65	DIS	DISNEY (WALT) CO	0.045
66	RHI	ROBERT HALF INTL INC	0.045
67	AIG	AMERICAN INTERNATIONAL GROUP	0.045
68	BSX	BOSTON SCIENTIFIC CORP	0.045
69	XOM	EXXON MOBIL CORP	0.045

70	KFRC	KFORCE INC	0.045
71	SCHW	SCHWAB (CHARLES) CORP	0.045
72	AMP	AMERIPRISE FINANCIAL INC	0.045
73	CAH	CARDINAL HEALTH INC	0.045
74	IQV	IQVIA HOLDINGS INC	0.045
75	V	VISA INC	0.045
76	WDAY	WORKDAY INC	0.044
77	BBY	BEST BUY CO INC	0.044
78	AMGN	AMGEN INC	0.044
79	AAL	AMERICAN AIRLINES GROUP INC	0.044
80	MET	METLIFE INC	0.044
81	PRU	PRUDENTIAL FINANCIAL INC	0.044
82	S	SPRINT CORP	0.044
83	LLY	LILLY (ELI) & CO	0.044
84	BDX	BECTON DICKINSON & CO	0.044
85	M	MACY'S INC	0.044
86	PYPL	PAYPAL HOLDINGS INC	0.044
87	CHTR	CHARTER COMMUNICATIONS INC	0.044
88	CVX	CHEVRON CORP	0.044
89	MCK	MCKESSON CORP	0.044
90	CMCSA2	NBCUNIVERSAL MEDIA LLC	0.044
91	CTL	CENTURYLINK INC	0.044
92	BK	BANK OF NEW YORK MELLON CORP	0.043
93	CERN	CERNER CORP	0.043

94	AFL	AFLAC INC	0.043
95	DAL	DELTA AIR LINES INC	0.043
96	FIS	FIDELITY NATIONAL INFO SVCS	0.043
97	INTU	INTUIT INC	0.043
98	ABBV	ABBVIE INC	0.043
99	JWN	NORDSTROM INC	0.043
100	BMY	BRISTOL-MYERS SQUIBB CO	0.043

Table 3 Correlation between Connectivity (Social Capital) Measures and Firm Characteristics

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. Eigenvector centrality	**1.00**	**0.51**	**0.36**	**0.07**	**0.35**	**0.35**	**0.35**	**0.35**	**0.36**	**0.35**	**0.35**	**0.44**	**0.44**	**0.34**	**0.06**	**0.48**	**0.41**	**-0.07**	**0.40**	**-0.02**	**0.45**
2. Number of employees on LinkedIn	**0.51**	**1.00**	**0.88**	**0.07**	**0.90**	**0.85**	**0.88**	**0.86**	**0.84**	**0.88**	**0.87**	**0.52**	**0.53**	**0.42**	0.02	**0.62**	**0.54**	**-0.03**	**0.49**	-0.01	**0.50**
3. Total number of employee connections	**0.36**	**0.88**	**1.00**	**0.13**	**0.97**	**0.99**	**1.00**	**1.00**	**0.97**	**0.99**	**1.00**	**0.37**	**0.42**	**0.26**	0.02	**0.62**	**0.43**	-0.02	**0.38**	0.00	**0.35**
4. Average number of connections	**0.07**	**0.07**	**0.13**	**1.00**	**0.11**	**0.14**	**0.13**	**0.13**	**0.13**	**0.12**	**0.13**	-0.02	0.01	**-0.04**	**-0.05**	**0.05**	0.01	**0.04**	**0.03**	**0.03**	**-0.03**
5. Total number of employee internal connections	**0.35**	**0.90**	**0.97**	**0.11**	**1.00**	**0.95**	**0.97**	**0.97**	**0.91**	**0.99**	**0.97**	**0.35**	**0.41**	**0.25**	0.01	**0.59**	**0.44**	-0.02	**0.38**	0.00	**0.33**
6. Total number of employee connections with people who have the same education level	**0.35**	**0.85**	**0.99**	**0.14**	**0.95**	**1.00**	**0.99**	**0.99**	**0.98**	**0.98**	**1.00**	**0.36**	**0.41**	**0.25**	0.02	**0.64**	**0.43**	-0.02	**0.37**	0.00	**0.34**
7. Total number of employee connections with people who have the Same experience level	**0.35**	**0.88**	**1.00**	**0.13**	**0.97**	**0.99**	**1.00**	**1.00**	**0.95**	**0.99**	**1.00**	**0.36**	**0.41**	**0.25**	0.02	**0.61**	**0.42**	-0.02	**0.38**	0.00	**0.34**
8. Total number of connections with people who have the same job functions	**0.35**	**0.86**	**1.00**	**0.13**	**0.97**	**0.99**	**1.00**	**1.00**	**0.97**	**0.99**	**1.00**	**0.36**	**0.41**	**0.25**	0.02	**0.62**	**0.43**	-0.02	**0.37**	0.00	**0.34**
9. Total number of connections with people who have gone to the same school	**0.36**	**0.84**	**0.97**	**0.13**	**0.91**	**0.98**	**0.95**	**0.97**	**1.00**	**0.95**	**0.97**	**0.40**	**0.43**	**0.26**	0.02	**0.69**	**0.46**	-0.02	**0.37**	0.00	**0.37**
10. Total number of connections with people who have worked at the same institution.	**0.35**	**0.88**	**0.99**	**0.12**	**0.99**	**0.98**	**0.99**	**0.99**	**0.95**	**1.00**	**0.99**	**0.35**	**0.41**	**0.25**	0.01	**0.62**	**0.44**	-0.02	**0.38**	0.00	**0.33**
11. Exp(Average Similarity)	**0.35**	**0.87**	**1.00**	**0.13**	**0.97**	**1.00**	**1.00**	**1.00**	**0.97**	**0.99**	**1.00**	**0.36**	**0.41**	**0.25**	0.02	**0.63**	**0.43**	-0.02	**0.37**	0.00	**0.34**
12. Revenue (in millions)	**0.44**	**0.52**	**0.37**	**-0.02**	**0.35**	**0.36**	**0.36**	**0.36**	**0.40**	**0.35**	**0.36**	**1.00**	**0.72**	**0.80**	0.02	**0.43**	**0.58**	**-0.04**	**0.48**	-0.01	**1.00**

13. Assets (in millions)	**0.44**	**0.53**	**0.42**	0.01	**0.41**	**0.41**	**0.41**	**0.41**	**0.43**	**0.41**	**0.41**	**0.72**	**1.00**	**0.46**	0.02	**0.55**	**0.61**	**-0.03**	**0.73**	-0.01	**0.72**
14. Number of employees (in thousands)	**0.34**	**0.42**	**0.26**	**-0.04**	**0.25**	**0.25**	**0.25**	**0.25**	**0.26**	**0.25**	**0.25**	**0.80**	**0.46**	**1.00**	0.02	**0.22**	**0.37**	**-0.03**	**0.31**	-0.01	**0.80**
15. ROA	**0.06**	**0.02**	**0.02**	**-0.05**	0.01	0.02	0.02	0.02	0.02	0.01	0.02	**0.02**	**0.02**	**0.02**	**1.00**	**0.02**	**0.02**	**-0.73**	**0.02**	**-0.63**	**0.02**
16. R&D Expense	**0.48**	**0.62**	**0.62**	**0.05**	**0.59**	**0.64**	**0.61**	**0.62**	**0.69**	**0.62**	**0.63**	**0.43**	**0.55**	**0.22**	**0.02**	**1.00**	**0.54**	-0.01	**0.42**	-0.01	**0.43**
17. Net Income (in millions)	**0.41**	**0.54**	**0.43**	0.01	**0.44**	**0.43**	**0.42**	**0.43**	**0.46**	**0.44**	**0.43**	**0.58**	**0.61**	**0.37**	**0.02**	**0.54**	**1.00**	**-0.02**	**0.49**	-0.01	**0.57**
18. R&D/Assets	**-0.07**	**-0.03**	**-0.02**	**0.04**	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	**-0.04**	**-0.03**	**-0.03**	**-0.73**	-0.01	**-0.02**	**1.00**	**-0.03**	**0.62**	**-0.04**
19. Intangible assets (in millions)	**0.40**	**0.49**	**0.38**	**0.03**	**0.38**	**0.37**	**0.38**	**0.37**	**0.37**	**0.38**	**0.37**	**0.48**	**0.73**	**0.31**	**0.02**	**0.42**	**0.49**	**-0.03**	**1.00**	-0.01	**0.47**
20. Tobin's Q	**-0.02**	-0.01	0.00	**0.03**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.01	**-0.63**	-0.01	-0.01	**0.62**	-0.01	**1.00**	-0.01
21. Real sales	**0.45**	**0.50**	**0.35**	**-0.03**	**0.33**	**0.34**	**0.34**	**0.34**	**0.37**	**0.33**	**0.34**	**1.00**	**0.72**	**0.80**	**0.02**	**0.43**	**0.57**	**-0.04**	**0.47**	-0.01	**1.00**

Figure 1 Connection Heat Map for S&P 100 Firms

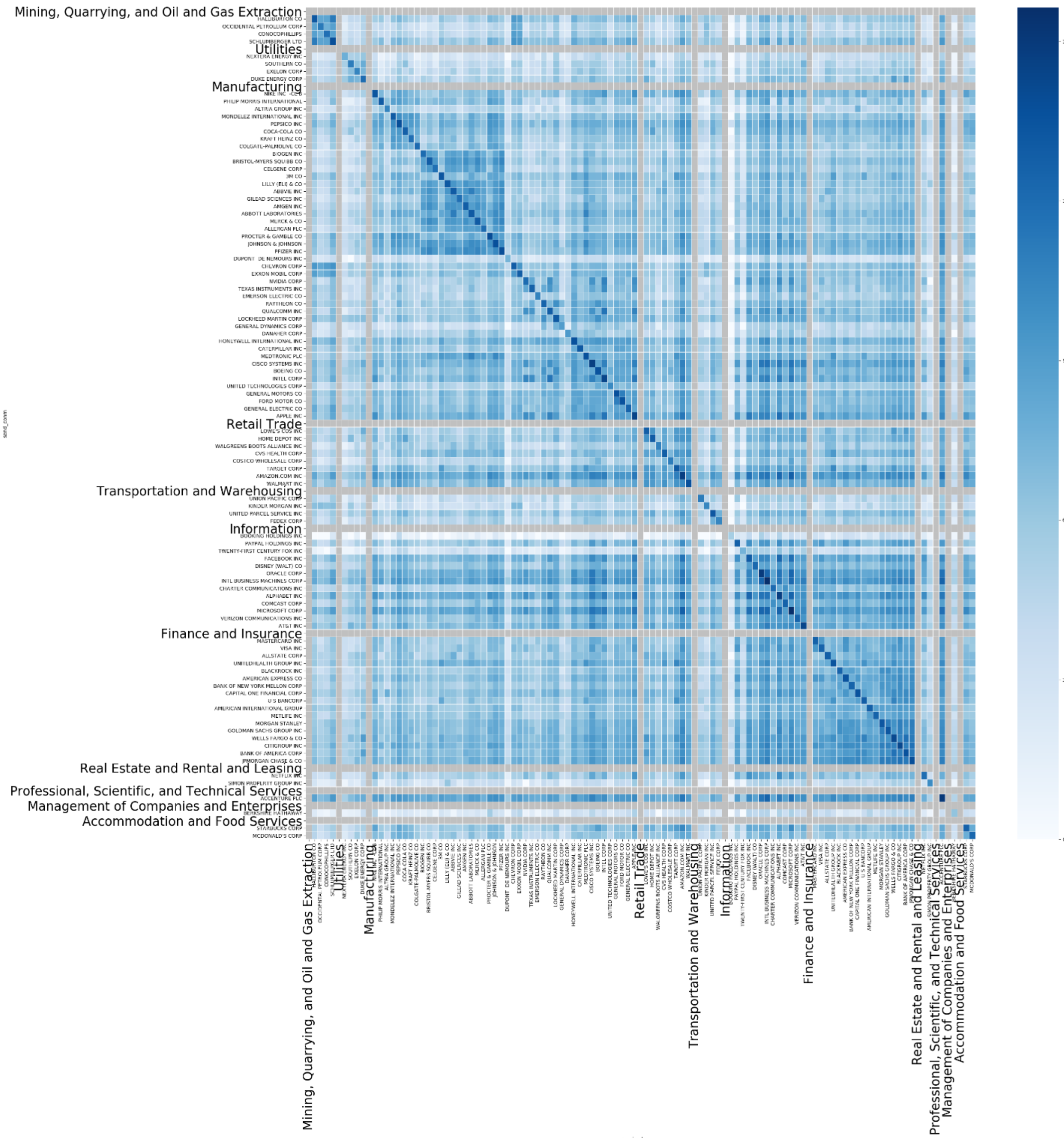
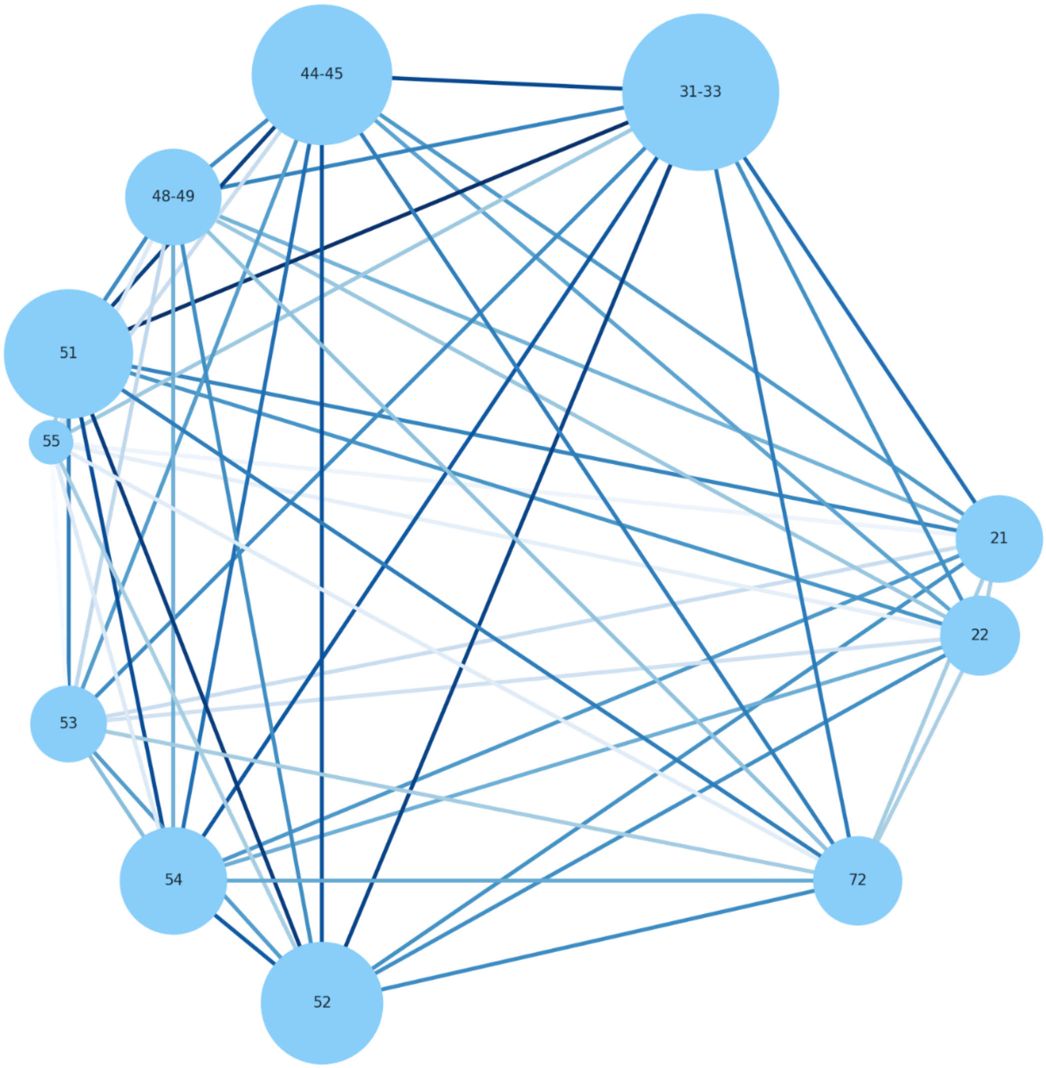


Figure 2 Industry Network Diagram for S&P 100 Firms (2018)



2-digit NAICS code	Industry Name
11	Agriculture, Forestry, Fishing and Hunting
21	Mining, Quarrying, and Oil and Gas Extraction
22	Utilities
23	Construction
31-33	Manufacturing
42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing

51	Information
52	Finance and Insurance
53	Real Estate and Rental and Leasing
54	Professional, Scientific, and Technical Services
55	Management of Companies and Enterprises
56	Administrative and Support and Waste Management and Remediation
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services (except Public Administration)
92	Public Administration

Figure 3 Industry Network Diagram for S&P 500 Firms (2018)

An explanation for the 2-digit industry codes is provided on the previous page (after **Figure 2**).

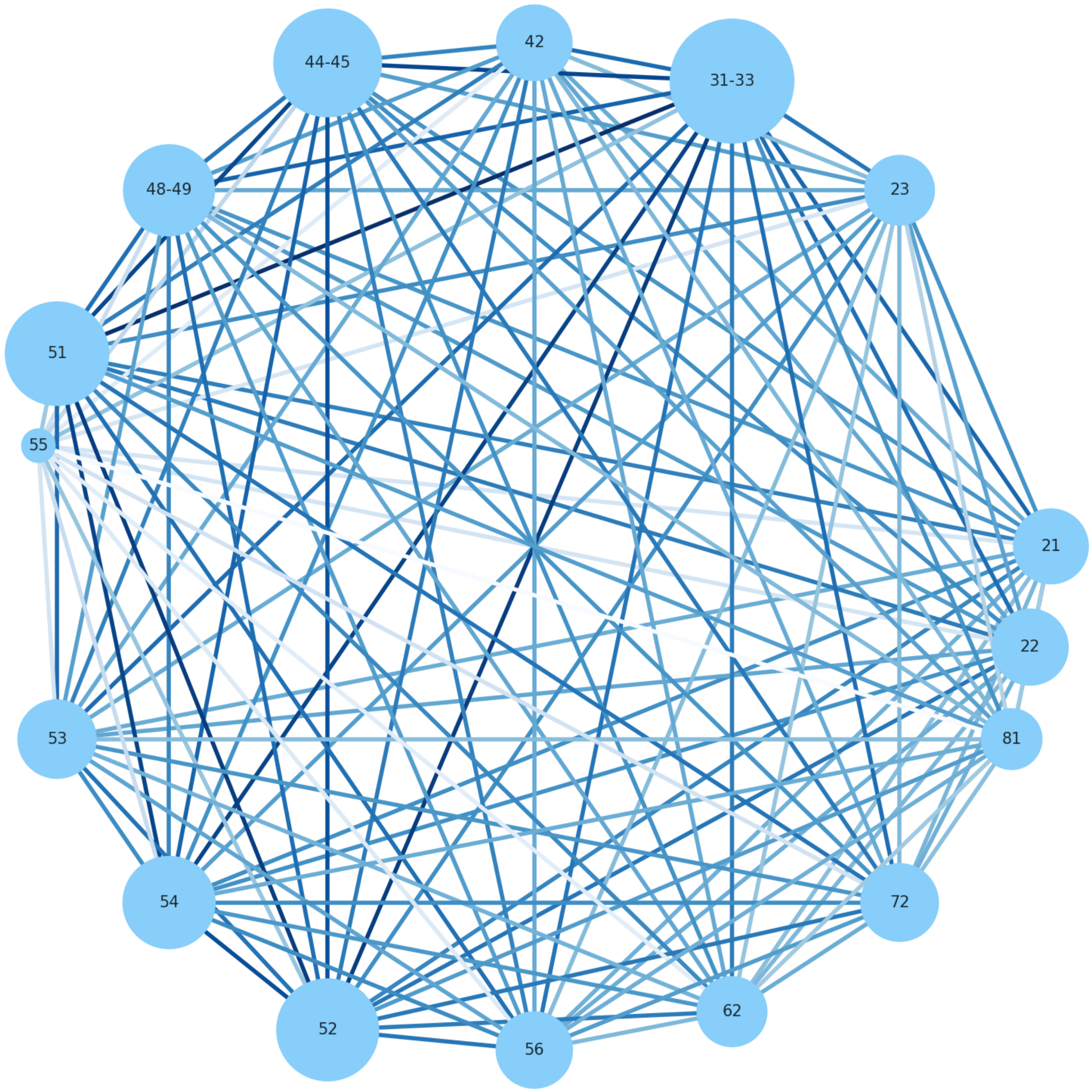


Figure 4 Industry Network Diagram for All US Public Firms (2018)

An explanation for the 2-digit industry codes is provided on the previous page (after **Figure 2**).

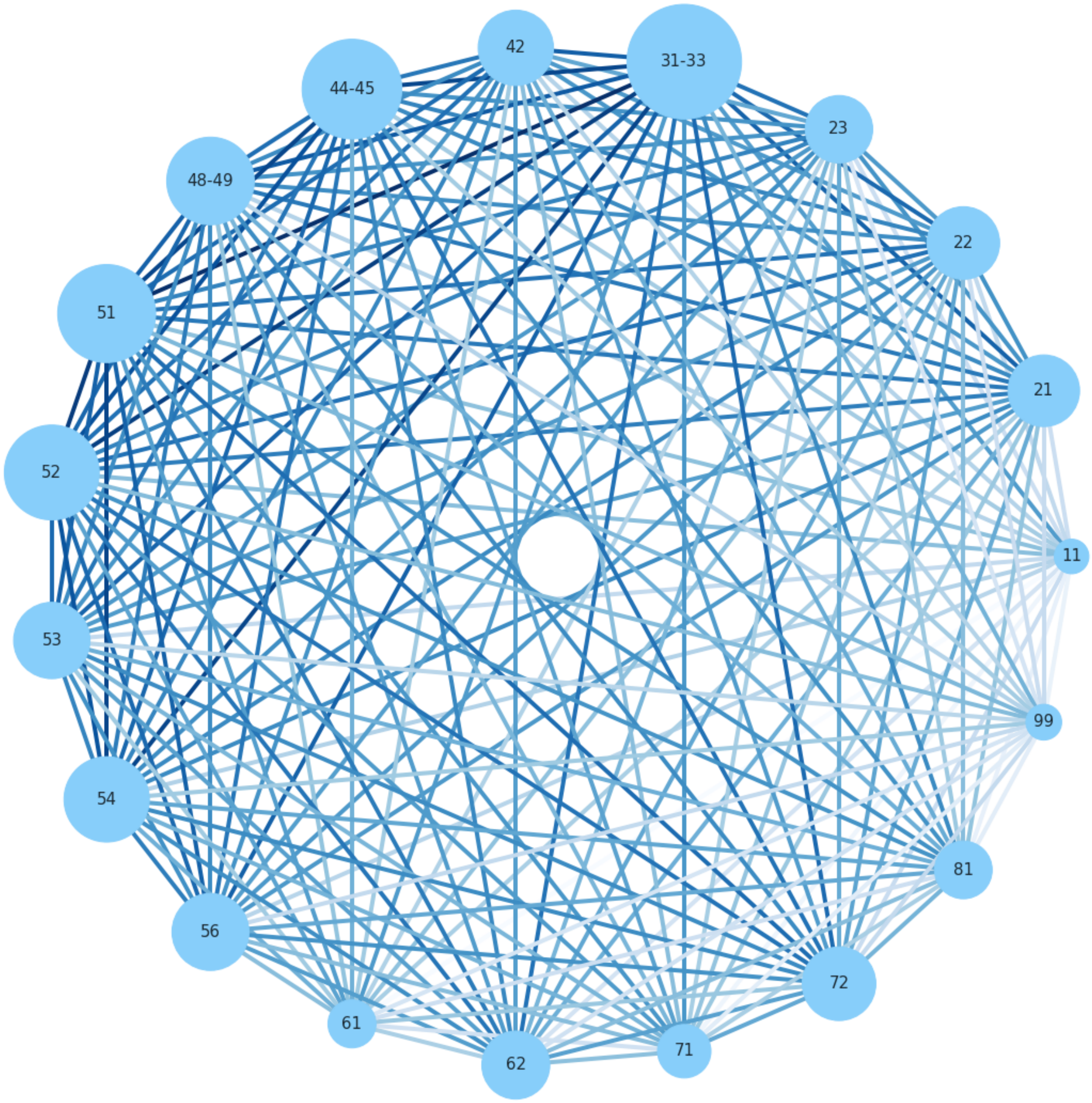


Table 4 Role of Social Capital in Productivity

	1	2	3	4
DV	**Real Sales**	**Real Sales**	**Tobin's Q**	**Tobin's Q**
Model	**OLS**	**Firm FE**	**OLS**	**Firm FE**
Total # of Employee Connections	0.050***	0.004	0.053***	0.004
Average # of Employee Connections	0.087***	0.001	0.093***	-0.011
Total # of Internal Connections	0.037***	0.006	0.035***	0.004
Eigenvector Centrality	8.834***	2.038***	13.734***	2.501***
Average Similarity	0.068**	0.030*	0.116***	0.053***
Control Variables	K, L	K, L	Assets, R&D/Assets, ROA, Leverage	Assets, R&D/Assets, ROA, Leverage
Industry FE	Y	-	Y	-
Year FE	Y	Y	Y	Y

Note: ***p<.01, **p<.05, *p<.1. Each SC variable is included in the regression one at a time. Analysis of Average Similarity includes a control for total # of employee connections.

Table 5 Eigenvector Centrality By Job Level

	1	2	3	4
DV	**Real Sales**	**Real Sales**	**Tobin's Q**	**Tobin's Q**
Model	**OLS**	**Firm FE**	**OLS**	**Firm FE**
High to High	4.799***	0.432	8.879***	0.560
Medium to Medium	10.716***	18.732***	19.209***	11.227***
Low to Low	5.866***	11.549***	13.829***	11.003***
Control Variables	K, L	K, L	Assets, R&D/Assets, ROA, Leverage	Assets, R&D/Assets, ROA, Leverage
Industry FE	Y	-	Y	-
Year FE	Y	Y	Y	Y

Note: ***p<.01, **p<.05, *p<.1. Each eigenvector variable is included In the regression one at a time.