NETWORK CENTRALITY AND THE CROSS SECTION OF STOCK RETURNS

KENNETH R. AHERN[†]

ABSTRACT

Industries that are more central in the network of intersectoral trade earn higher stock returns than industries that are less central. This finding is economically substantial and robust to controls for firm size, leverage, industrial concentration, standard asset pricing factors, and other return determinants. To explain this finding, I draw on recent research that shows that macroeconomic fluctuations are the aggregation of sector-specific shocks. For stock returns, this implies that systematic risk originates from idiosyncratic shocks. I argue that stocks in more central industries have greater systematic risk and earn higher returns because they have greater exposure to idiosyncratic shocks that transmit from one industry to another through intersectoral trade.

This version: 15 November 2012

Keywords: Systematic risk, intersectoral network, stock returns, diversification, input-output links

[†] University of Southern California, Marshall School of Business, 3670 Trousdale Parkway, BRI 204J, Los Angeles, CA 90089. E-mail: kenneth.ahern@marshall.usc.edu.

Network Centrality and the Cross Section of Stock Returns

ABSTRACT

Industries that are more central in the network of intersectoral trade earn higher stock returns than industries that are less central. This finding is economically substantial and robust to controls for firm size, leverage, industrial concentration, standard asset pricing factors, and other return determinants. To explain this finding, I draw on recent research that shows that macroeconomic fluctuations are the aggregation of sector-specific shocks. For stock returns, this implies that systematic risk originates from idiosyncratic shocks. I argue that stocks in more central industries have greater systematic risk and earn higher returns because they have greater exposure to idiosyncratic shocks that transmit from one industry to another through intersectoral trade.

In standard asset pricing models, including the capital asset pricing model (Sharpe, 1964; Lintner, 1965), the arbitrage pricing theory (Ross, 1976), and the three-factor model of Fama and French (1993), exposure to systematic risk determines a stock's expected returns. However, none of these models offers any guidance for understanding the determinants of a firm's exposure to systematic risk. Instead, risk is identified statistically using ex-post correlations between a firm's stock returns and market-wide factors. While this approach identifies which stocks are riskier than others, it doesn't explain *why* some stocks are riskier than others. Considering that systematic risk is the foundation for virtually all of financial economics, it is crucial that we understand the ex-ante causes of systematic risk, not just its symptoms.

In this paper, I propose that a large fraction of a firm's exposure to systematic risk can be identified from stable ex-ante fundamentals. In particular, I argue that systematic risk is formed from the aggregation of idiosyncratic shocks. A shock to one sector of the economy is likely to spillover to other sectors through economic transactions. Thus, an idiosyncratic shock could transmit through product market relations to create an aggregate shock. Take wars, for example. A war has differential effects on each sector of the economy. The most directly affected is the defense industry. But since the defense industry requires steel, computer hardware, and labor inputs, these sectors are also affected. The demand shock to these sectors could then affect their expenditures as well. As another example, consider oil shocks. An oil embargo is likely to have economy-wide effects, but the embargo does not affect all stocks simultaneously and uniformly. Instead, refineries produce less gasoline, shipping companies have higher costs, and the profitability of online retailers is affected. Thus, the impacts of oil shocks and wars are systematic risks, but they originate as idiosyncratic shocks that aggregate to the macroeconomy through economic trade.

Recent theoretical and empirical evidence supports this view. Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012) build upon the seminal work of Long and Plosser (1983) to advance a theory that sectoral shocks are transmitted to other sectors through a network of input and output linkages. If links in the input-output network are not uniformly distributed, 'volatility cascades' can arise from idiosyncratic shocks. Ahern and Harford (2012) empirically model the input-output trade network in the U.S. and show that the network is characterized by a relatively small number of central industries and many peripheral industries. They show that industry-level merger waves transmit across the network through product market links to form aggregate merger waves. Similarly, Carvalho and Gabaix (2012) presents empirical evidence that volatility in aggregate national output is driven by sectoral shocks.

If systematic risk is the aggregation of sector-level idiosyncratic shocks, then sectors that are more central in the network of intersectoral trade will have greater exposure to systematic risk. By definition, central sectors have stronger and greater numbers of connections to other sectors in the economy. Therefore, the likelihood that a central industry is affected by shocks that transmit through the network is higher than is the likelihood that a peripheral industry is affected. Thus, a sector's centrality in the intersector trade network is expected to be positively related to its average stock returns. This argument provides an exogenous and ex-ante identifiable source of variation in systematic risk.

To test this hypothesis, I construct an empirical model of trade flows between all sectors of the economy, including industry, households, government, capital, and a foreign sector. Following Ahern and Harford (2012), I first construct the network of inter-industry trade flows using the 1997 Input-Output tables from the U.S. Bureau of Economic Analysis (BEA). I then append households, government, capital, and the foreign sector to the network using data from the National Income and Product Account (NIPA) tables provided by the BEA. To my knowledge, this is the first network representation of an applied closed-economy general equilibrium model of the U.S. economy. This provides a complete picture of all economic transactions between a large number of dissagregated sectors (483 in total). Using techniques from graph theory, I then calculate the centrality of each sector in the economy.

The main finding in this paper is that more central industries earn higher stock returns, even after controlling for a wide range of alternative variables, including industry concentration, firm size, and known pricing factors. First, I find that industries in the highest quintile of centrality have average monthly stock returns that are 1.1 percentage points higher than industries in the lowest quintile of centrality. This result is highly statistically significant and economically meaningful. In addition, this effect is present in unlevered returns too, which suggests that the higher returns for more central industries are compensation for operating risk, not financial risk.

Second, I double sort industries into portfolios based on centrality and an industry's average firm size, its concentration of customers, and its concentration of suppliers. First, since central industries are larger with larger average firms, centrality may simply pick up a size effect. Second, because central industries have more connections to other sectors, their exposure to shocks through product market links may be more diversified. All else equal, industries with greater concentration of supplier and customer sectors are likely to face greater risks from exposure to systematic shocks. This presents a new view of diversification based on the concentration of customers and suppliers, rather than the concentration of a firm's operations, as is typically studied.

After accounting for average firm size and the concentration of suppliers and customer sectors, I still find a strong positive and significant relation between stock returns and centrality. In contrast, I find that an industry's average firm size does not explain the cross-section of stock returns. I also find that the concentration of customer segments is unrelated to stock returns. However, I do find evidence that industries with more concentrated suppliers earn higher stock market returns. In addition, there is a positive interaction between centrality and the concentration of supplier sectors. Those industries that both have the most concentrated suppliers and are the most central, earn the highest stock returns of all 25 double-sorted portfolios. This is consistent with higher compensation for greater exposure to systematic risk that is magnified through a lack of diversification.

I next show that the positive relation between centrality and stock returns is not explained by known pricing factors. I run time-series regressions of the Fama-French and Momentum factors on monthly returns sorted by centrality quintiles. The alpha of the regression of the lowest centrality quintile is 0.47% per month, compared to an alpha of 1.5% for the portfolio of industries in the highest quintile of centrality. This difference is statistically significant and economically important. This suggests that the standard pricing models do not explain the cross-sectional variation in returns across centrality quintiles. In support of this, I find no consistent significant differences between the loadings on the Fama-French and Momentum factors in the five centrality quintiles. Thus, sector centrality explains significant variation in returns that standard asset pricing models do not.

Finally, I run industry-level cross-sectional regressions of average stock returns on a host of explanatory variables. In addition to an industry's average firm size and concentration of customers and supplier sectors, described above, I also control for within-industry sales concentration, the scope of industry activities, the fraction of input costs paid to labor, and an alternative centrality measure based on asset complementarity as in Hoberg and Phillips (2010a) and Hoberg and Phillips (2010b). After controlling for all of these alternative explanations, I still find a strong positive relation between stock returns and industry centrality. In sum, in a wide variety of tests, I find consistent evidence that firms that are more central to the economy face greater exposure to systematic risk and earn higher stock returns.

The central contribution of this paper is to provide a new explanation for the cause of systematic risk, based on centrality in the network of intersectoral trade flows. This sheds light on the nature of systematic risk using ex ante exogenous characteristics and also connects systematic risk to recent macroeconomic models of sectoral volatility (Conley and Dupor, 2003; Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi, 2012). This is an improvement over ex-post statistical measures of risk in standard asset pricing models. In this vein, this paper contributes to recent research that connects stock returns to economic fundamentals. This includes Carlson, Fisher, and Giammarino (2004) and Zhang (2005), which relate asset prices to product market demand and the market for capital investment, and Hou and Robinson (2006), which finds that more concentrated industries have lower returns. Whereas these papers study within-industry characteristics, this paper emphasizes the importance of connections between industries to explain systematic risk.

This paper is also related to recent research that investigates the predictability of stock returns due to limited investor attention. Hong, Torous, and Valkanov (2007) finds that a subset of industries' returns lead the overall market return. The leading industries are those that are the most informative about macroeconomic fundamentals, such as real estate, retail, and financial. My results are consistent with these, since these industries are also the most central in the intersector network. In addition, other research shows that while leading industries are informative, investors do not anticipate the secondary effect on other industries. Cohen and Frazzini (2008) and Menzly and Ozbas (2010) find that firms related through customer and supplier relationships have predictable returns. As in Hong et al. (2007), they argue that though a shock to a customer will affect a supplier, stock prices respond with a lag due to limited attention by investors. My results support the idea that systematic risk transmits along a supply network and lend plausibility to the idea that investors' attention is constrained, given the extreme complexity of the trade network.

The remainder of the paper is organized as follows. Section I describes the construction of the intersector network of trade flows. Section II discusses the relation between systematic risk and centrality in the network. Section III presents the empirical results. Section IV concludes.

I. The Network of Economic Transactions

To understand the sources and transmission of systematic risk, I construct a network of all economic transactions within the economy from households, government, a foreign sector, and a large number of dissagregated industries. This network is based on a modified social accounting matrix (SAM). A SAM presents an account of the circular flow of transactions between a complete set of economic agents, including production activities, factors of production (capital and labor), and institutions (households, firms, and government). Each row of the SAM provides the receipts of an agent and each column provides the expenditures, with receipts equal to expenditures for each agent in the economy. A SAM can be thought of as an expanded input-output (IO) table, that includes flows between government, a capital account, and a foreign sector.

A. Social Accounting Matrix

Though SAMs are commonly used in other countries, to my knowledge, there is no published SAM matrix for the US. Therefore, I construct the SAM using data from the 1997 Input-Output (IO) tables and the National Income and Product Account (NIPA) tables provided by the Bureau of Economic Analysis (BEA). As discussed in Ahern and Harford (2012), the BEA updates the IO tables every five years to maintain consistency in industry homogeneity, accounting for technological advance. This means that the industry definitions change in each report, making it impossible to analyze a consistent set of industries over time, while accounting for changing trade patterns. Therefore, I focus on 1997 because it is the first year that the IO tables were based on more accurate NAICS codes, rather than SIC codes, and because it is centered in a time of stock market growth and then bust, which allows for more variation in stock returns.[1]

[1]Future versions of this paper will investigate additional years.

To construct the SAM for the US, I start by creating an IO table using the 1997 'Make' and 'Use' tables reported by the Bureau of Economic Analysis, as in Ahern and Harford (2012) (AH). These tables report the flow of 486 commodity outputs by 494 producing industries or final users. A commodity, as defined by the BEA, is any good or service that is produced, including all sectors of the economy, not just manufacturing. The 'Make' table records the dollar value of each commodity produced by the producing industry. The 'Use' table defines the dollar value of each commodity that is purchased by each industry or final user. I create a matrix that records industry trade flows (rather than commodity to industry flows), accounting for each industry's fraction of the total production of each commodity. See AH for a detailed description of the inter-industry matrix.

I make a few changes to the procedure in AH, because I wish to capture a complete representation of all trade flows, not just inter-industry flows. Therefore, I include imports and exports, used and secondhand goods, government enterprise, and other taxes and adjustment costs. This produces a matrix of 478 producing industries, plus an industry that accounts for scrap, inventory adjustments, and used goods.

To create the SAM, I append four additional large sectors to complete the economy: government, households, capital, and the foreign sector, for a total of 483 economic sectors. From the IO tables, I aggregate government expenditures (consumption and investment) for both Federal and local governments. I use data from the 1997 NIPA tables to record government receipts from industry, households, capital, and foreign sectors. Household expenditures across all industries is recorded from the 'Personal consumption expenditures' item of the IO tables. Household taxes are from the NIPA tables, as is household income from government transfers and the capital account. Flows to the capital and foreign sectors are computed using IO and NIPA data. For a complete description of the sources of data for the SAM and how each entry is calculated, see the Internet Appendix.

Table I presents the aggregated SAM, where for the sake of brevity, the economic activity of all 479 producing industries is compressed into one entry for firms. In later analysis, I use the disaggregated industries as the unit of observation.

Panel A of Table I shows that firms spent \$6.5 trillion on intermediate inputs, \$4.7 trillion on labor, invested \$2.6 billion in capital, and paid \$1.3 trillion to the government. Of the \$16 trillion in total firm output, \$6.5 trillion was purchased by other firms, \$5.6 trillion was consumed by households, and \$1.6 trillion was transfers from government. The majority of household income was from labor (\$4.7 trillion), followed by capital (\$1.4 trillion), and government transfers (\$0.9 trillion). Finally, the foreign sector purchased \$822 billion from firms, \$44 billion from households, and paid tax of \$5 billion to government. Foreign receipts include \$944 billion from firms in the form of imports, \$998 million in capital gains, and \$24.9 billion in government spending. Thus, the trade deficit is \$98.7 billion, recorded as foreign spending towards the capital account.

Panel B of Table I normalizes the dollar flows by total expenditures. Thus, 40.7% of firm expenditures went towards intermediate goods sold by other firms and 29% went to labor. Of total household expenditures, 80.1% went to firms and 18% went to the government. In Panel C, I normalize by total receipts. Firms received 40.7% of total receipts from other firms and 35% from households. Households received 66.4% of their income from firms, in the form of labor compensation, and 13.3% from government transfers. These representations of the SAM serve to normalize the large differences in sector-size.

Because the SAM is based on complete data from all economic sectors, not just industries, and not just publicly-traded firms recorded in Compustat and CRSP, this network is an empirical representation of a closed-economy applied general equilibrium model. In contrast to focusing on only inter-industry connections, the SAM provides a complete picture of the economy where all sectors' outputs and inputs balance. This provides an important benefit over the partial equilibrium approach in Ahern and Harford and Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012), because it allows for the identification of the full effect of a shock to any particular sector.

B. Network Representation of Intersector Transactions

The trade flows of a SAM can be interpreted as connections in a network of economic sectors. In general, a network is defined by a set of nodes that are connected through edges. A network is represented by a square 'adjacency' matrix, denoted A, where each entry a_{ij} for row i and column j indicates the connection between sector i and j. The matrix can be asymmetric, which allows directional relations, and the edges can be weighted to measure the strengths of the connections. Thus, to translate a SAM to a network setting is straightforward. The nodes in the network are the 479 industries, plus the household, government, capital, and foreign sectors. The trade flows recorded in the SAM represent the strength and direction of the connection between agents. Considering the SAM as a network, rather than simply a matrix of expenditures and receipts allows the use of a wide range of techniques developed in the graph theory and social network literature.[2]

To illustrate how the SAM network can be used to trace sector-specific shocks, consider a \$1 dollar exogenous increase in household income. Using the network of suppliers in Panel B of Table I, of this \$1.00, \$0.80 will be used to purchase goods from firms, \$0.02 will be spent on the capital sectors, and \$0.18 will be transfered to the government. If we continue to trace the effect of the shock, of the \$0.80 received by firms, 40.7% $\times$ \$0.80 = \$0.33 will be retained by firms through spending on intermediate goods, 29.0% $\times$ \$0.80 = \$0.23 will be spent on labor compensation, 16.3% $\times$ \$0.80 = \$0.13 will be invested in capital, 8.2% $\times$ \$0.80 = \$0.07 will be paid to the government, and 5.9% $\times$ \$0.80 = \$0.05 will be sent to the foreign sector. The same calculations will show where the \$0.02 received by the capital sector and the \$0.18 received by the government will be spent. Alternatively, the SAM network can be thought of as a Markov matrix, where each entry represents the likelihood that the entire dollar shock moves from one sector to another.

As the exogenous shock transmits across the SAM network, a stable outcome emerges. This outcome is calculated as A^{∞}, which shows that for a \$1 dollar exogenous shock anywhere in the economy, firms receive \$0.54, households receive \$0.24, the capital sector receives \$0.10, the government receives \$0.09, and the foreign sector receives \$0.03. Because this is a closed economy, tracing shocks through customer relations, as in Panel C of Table I, provides an identical result. Using the Markov matrix interpretation, the transition probabilities are directly proportional to the total output of each sector. Thus, larger sectors are proportionally more likely to receive a shock, or will receive a proportionally larger share of the exogenous shock.

This stable equilibrium concept is directly related to network centrality. Centrality can be measured in various ways, but as discussed in Ahern and Harford, eigenvector centrality is the most appropriate centrality measure for economic links. Eigenvector centrality is the principal eigenvector of the network's adjacency matrix (Bonacich, 1972). Nodes are more central if they are connected to other nodes that are themselves more central. As presented in Ahern and Harford, if we define the eigenvector centrality of node i as c_i, then c_i is proportional to the sum of the c_j's

[2]This insight was first made in Ahern and Harford (2012) using inter-industry IO relations. For a detailed introduction to networks in finance, see the Internet Appendix of their paper.

for all other nodes $j \neq i$: $c_i = \frac{1}{\lambda} \sum_{j \in M(i)} c_j = \frac{1}{\lambda} \sum_{j=1}^{N} A_{ij} c_j$, where $M(i)$ is the set of nodes that are connected to node i and λ is a constant. In matrix notation, this is $\mathbf{A}\mathbf{c} = \lambda\mathbf{c}$. Thus, $\mathbf{c}$ is the principal eigenvector of the adjacency matrix.

Network centrality provides a measure of how important a node is in an industry. In the case of the SAM network, it directly measures the strength and number of connections of a sector, considering the importance of the sectors to which it is connected. Equivalently, by tracing out all paths of a random shock in a network, as shown above, we get an equivalent measure of centrality. Thus, centrality in the SAM network is a measure of how important a sector is for the overall economy. Importantly, it also measures the likelihood that a sector will receive a random shock.

II. Systematic Risk and Network Centrality

Given that network centrality measures an industry's exposure to random shocks, I hypothesize that more central industries face greater risks, and hence, earn higher stock returns. There are two important assumptions of this hypothesis. First is the premise that aggregate shocks originate as idiosyncratic events. In contrast to the notion that aggregate shocks simultaneously affect all sectors, it is plausible that technological shocks, demand shocks, and natural resource shocks start in a single sector. For example, the invention of interchangeable parts had a profound impact on the productivity in manufacturing, but little direct impact on service sectors, such as hotels, doctors, and bookkeeping. Even macroeconomic events, such as currency and interest rate shocks, have a more direct effect on certain industries, such as banking, compared to others, such as legal services and car repair shops, for example.

Oil shocks illustrate how sector-specific shocks can lead to aggregate effects. Though oil shocks are often considered to be systematic risks, they originate locally. They first impact oil extraction firms. However, because oil is the key input in gasoline, oil shocks spillover to refineries. Gas prices affect delivery services and transportation, as well as general consumers. Thus, oil shocks are not 'blanket' shocks, affecting all sectors directly, but rather, they are sequential shocks that affect all sectors because of the prevalence of oil-related products as intermediary goods in the economy. Though oil-related goods are important intermediary inputs, they are just one of many. For instance, the services of the finance and real estate sectors are widely used by most other sectors, as are steel, electricity, and paper products. Since all sectors are connected to some degree, each sector has potential to influence the overall economy.

The second underlying assumption is that random local shocks do not simply cancel out. If a positive shock in one industry is countered by a negative shock in another industry, on average, the economy would be unaffected. However, this argument assumes that the network of connections between sectors is uniform and random. As Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012) show in a theoretical model and Ahern and Harford (2012) show empirically, because the true sector network is characterized by asymmetric sectors with non-uniform connections, randomly occurring sectoral shocks do not cancel each other out through diversification, but instead may aggregate to form economy-wide events.

Therefore, because local shocks do not cancel out as they transmit through customer-supplier links, I hypothesize that more central industries in the SAM network have greater exposure to systematic risk and thus earn higher stock returns. This approach is appealing because it is based on the premise that economic links are the conduit for economic shocks to transmit across sectors and that stock returns are based on economic fundamentals.

III. Empirical Results

A. Univariate Results

To test the hypothesis that more central industries earn higher returns, I compare industry-level centrality using the 1997 SAM network and industry-level stock returns over 1998 to 2002. First, I compute eigenvector centrality using the fully dissagregated SAM network of 483 sectors normalized by sector size, as described above. Because the SAM is balanced, with the a row's sum across columns equal to its column sum across rows, eigenvector centrality computed using either the supplier network or customer network is identical. Second, I collect stock price information from the Center for Research in Securities Prices (CRSP) Monthly Stocks dataset for stocks with share codes 10 or 11 from January 1998 to December 2002. I choose these dates to ensure that the network I construct is predetermined and exogenous to stock returns, but also not stale. I omit observations where the stock price is less than five dollars to avoid liquidity effects. For each firm in the dataset, I match the firm's 6-digit 1997 historical NAICS code to the IO industry codes based on the concordance tables provided by the BEA, following Ahern and Harford (2012). I then form value-weighted industry-level portfolios, rebalanced monthly. The availability of stock price information limits the sample to 385 industries.

Figure 1 presents industry-level average monthly returns and centrality. Each circle represents one of the 385 industries. The industries are grouped according to large sectoral divisions, such as Fabricated Metal Manufacturing, Electrical Equipment, and Services. Each large division's fraction of the total circle is proportional to the division's fraction of the total number of industries. The distance of each industry from the center of the circle is based on its log(centrality), normalized by subtracting the maximum log(centrality) to avoid emptiness in the center of the figure. Industry average monthly returns from 1998 through 2002 are indicated by the darkness of each circle, where lighter colors indicate higher returns.

This figure reveals a number of relationships. First, the industries with the lowest returns are the least central industries, represented by the black circles near the outside edge of the figure. Industries that are more central, and closer to the center of the figure, have higher average returns. Second, the centrality of industries varies by the large sectoral divisions. In particular, the Service industries are all relatively central in the economy, compared to industries in the Machinery Manufacturing division. Similarly, industries in the Wholesale, Retail, and Transportation division are all relatively central, with retail and wholesale being the most central industries. The Miscellaneous Manufacturing division is characterized by periphery industries, which is intuitive, given that the industries are not important enough to be classified in any large division. Other divisions are characterized by greater diversity in centrality and returns, such as the Electrical Equipment division and the Food, Textiles, and Leather Goods division.

In Table II, I statistically verify that more central industries have higher stock returns than less central industries. I sort industries into five quintiles based on centrality in 1997. There are 59 industries in the lowest quintile, compared to 84 in the highest quintile. The lower number of industry observations for the least central industries is caused by fewer industries with publicly available stock price data.

In Panel A of Table II, I present the average time-series monthly returns over 1998 through 2002. First, I report the raw levered returns where portfolios are formed based on value-weighted firm-level observations. Firms in the lowest quintile of centrality have average stock returns of 1.21% per month, compared to 2.44% per month for firms in the highest quintile, a statistically significant difference. I also present results based on equally-weighted firm-level portfolio returns and equal-weighted industry returns (where industry returns are value-weighted across firms in the industry). In both cases I find a strong positive relationship between centrality and average returns.

The economic magnitude of the relation between centrality and stock returns is substantial. Across the three methods of forming portfolios, I find a difference of roughly 1.1 percentage points per month between the highest and lowest quintiles of centrality. This translates into approximately 12 percentage points per year difference between the extreme quintiles. Even between the second and fourth quintiles, I find a substantial yearly difference of roughly 3 percentage points.

For further robustness, I calculate the same tests using unlevered returns. Because the shocks that are likely to pass through the economic network are shocks to operations, rather than financing, unlevered returns are perhaps a better estimate. I calculate unlevered returns following Bernardo, Chowdhry, and Goyal (2007), as $r_U = \frac{r_E}{1+(1-\tau)D/E}$, where r_U is the unlevered return, r_E is the unadjusted equity return, τ is the marginal tax rate, and D/E is the firm's debt-to-equity ratio. The D/E ratio is computed from Compustat data. I use firm-level marginal tax rates generously provided by John Graham.

The results for the unlevered returns are highly similar to the main results, though the returns are lower because they are unlevered. Across value weighted and equal weighted portfolios, I find that the most central industries have returns that are roughly 1 percentage point higher per month than the least central industries. This provides further evidence that returns are influenced by centrality and supplier concentration, while controlling for the effects of leverage on operating returns.

In Panel B of Table II, I examine other possible variables that may explain the strong relation between centrality and stock returns. First, diversification may be important. From a corporate decision-making point of view, diversification provides a co-insurance effect for firms that operate in different industries (Lewellen, 1971). In contrast, I consider diversification from an industrial organization point of view, where an industry's production function determines diversification exogenously. In particular, some industries are characterized by using a few, but relatively important inputs in production. Others use a wide variety of inputs, none of which is relatively large. Likewise, some industries have a wide customer base, whereas others sell primarily to a few large sectors. A shock that reaches an industry that has few, but important, customer or suppliers may have a greater impact than a shock that reaches an industry with many small customers or suppliers, where the industry is diversified.

To measure industrial concentration, I compute the four-sector concentration ratio for customer and supplier industries. These record the fraction of an industry's total receipts that are received by the industry's largest four customer sectors (Concentration of Customers), and the fraction of an industry's total expenditures that are sent to the industry's four largest suppliers. All sectors, including government, households, industry, capital and foreign sectors are included in the calculations.

Panel B shows that more central industries have a smaller concentration of customers, but a larger concentration of suppliers. In the first centrality quintile, 74% of the average industries sales are accounted for by the four largest customer sectors. In the highest centrality quintile, 67% of sales go to the four largest customer sectors. Thus, more central industries have a wider base of customers. In contrast, 61% of inputs used by the average industry in the lowest centrality quintile are purchased from the four largest suppliers, compared to 68% for the most central industries. This means that more central industries have more concentrated inputs in their production function.

Second, centrality is related to industry size. There is a monotonic relation between centrality and the total output of an industry. In fact, the two concepts are interchangeable, since the largest industries are those that have the most and strongest connections to other industries. This may imply that more central industries also have larger firms which could affect stock returns. Panel B shows that more central industries do have larger firms. In summary, more central industries are larger and have firms that have larger market equity on average. Therefore, given that firm size is related to stock returns, it is important to control for this effect in later tests.

B. Double-Sorted Portfolios

To better understand the interaction between industrial concentration, firm size, and centrality, I double sort industries into 25 portfolios. I first sort industries into quintiles by either an industry's average market equity, its concentration of customers, or its concentration of suppliers. Within each of these quintiles, I sort industries into five quintiles of network centrality. This two-stage sorting ensures that there are enough observations in each quintile-pair.

For illustration, Table III lists a sample of five large industries, by output, in the extreme quintiles of these double sorts, omitting non-industrial sectors such as households and government. Panel A presents the extreme quintiles of average market equity, measured in December 1997, and network centrality. Industries with the lowest centrality and also the lowest average firm size include household vacuum cleaners, musical instruments, and explosives. These industries do not have strong connections to other industries and also are dominated by small firms. Industries such as mattress manufacturing, elevators, and industrial gases are dominated by large firms, but are not central in the intersector network. More central industries include frozen food, ready-mix concrete, hospitals, and petroleum refineries. However, frozen food and ready-mix concrete industries are characterized by small firms on average, compared to hospitals and refineries, which are characterized by large firms.

Panel B sorts on concentration of customers and centrality. The industries with the least concentrated customers, and lowest centrality include iron forging, nonwoven fabric mills, and stationary manufacturing. The industries with the greatest centrality, but least concentrated customers include truck transportation, semiconductors manufacturing, and nondepository credit intermediaries, which includes pension funds and credit card companies. Because these industries do not have concentrated customers, they sell their goods and services to a wide range of customers, as is intuitive based on the types of goods and services provided. However, fabric mills and stationary are not centrally located in the economy, whereas truck transportation and semiconductors are. In contrast, Panel B shows that sawmill machinery manufacturers, sugarcane farms, hospitals, and aircraft manufacturers all have concentrated customers. However, hospitals and aircraft manufacturers are highly central industries in the economy, whereas sawmill machinery and sugarcane farms are peripheral industries.

Sorting on the concentration of suppliers in Panel C reveals some similarities and differences. For instance, sugarcane and aircraft manufacturing both have high concentrations of customers, but low concentrations of suppliers. However, in both cases, aircraft manufacturing is a highly central industry, but sugarcane farming is a peripheral industry. On the other hand, hospitals and automobile manufacturers are both in the highest quintile of customers concentration as well as supplier concentration. These samples of industries provide a sense of which industries are more central and how centrality interacts with firm size and industrial concentration.

Using these same sorts, Table IV presents average monthly returns for the 25 industry portfolios. Panel A shows that the strong positive univariate relation between centrality and average stock returns persists after controlling for the average firm size in an industry. For all but the industries with the smallest firms, there is a large and statistically significant difference between the most central and the least central quintiles. The magnitude is roughly 1.3 percentage points per month. In contrast, there is no significant relationship between the average firm size in an industry in 1997 and its stock returns over 1998 to 2002.

An even larger positive effect of centrality on stock returns is revealed in Panel B, where industries are first sorted by the concentration of customer sectors. For all quintiles of customer concentration, except the highest quintile, there is a statistically and economically significant increase in monthly returns from the lowest quintile to the highest quintile of centrality. Across these four quintiles, the annual return on the 'Low−High' strategy is 22.6 percent, on average, and 18.84 percent if the highest customer concentration quintile is included. In contrast, there is no statistical difference across the five customer concentration quintiles.

Finally, when portfolios are formed based on the concentration of suppliers, a similar pattern emerges. In the top three quintiles of supplier concentration, the portfolio returns of the most central industries have significantly higher monthly returns than the portfolio of the least central industries. The average yearly return on the 'Low−High' portfolio across the five quintiles is 15.70 percent. In addition, Table IV also reveals that when centrality is high (in the top two quintiles), greater concentration of suppliers is associated with significantly higher returns. For the highest quintile of centrality, the annual return for the 'Low−High' portfolio of supplier concentration industries is 16.32 percent. This is consistent with a positive interaction between centrality and industrial concentration. Industries that are more central and prone to receive a shock through product market relations and that also have less diversified suppliers experience the highest average stock returns (2.44% per month) of all of the double-sorted portfolios. This is consistent with a risk-based explanation of the importance of position within the intersector network.

While central industries with the highest returns include both traditional manufacturing industries, such as refineries, and service industries, such as hospitals, they also include financial services industries like depository credit institutions and insurance companies. To make sure these results are not driven by financial risk through leverage, in Internet Appendix Table I, I verify the results are robust to using unlevered stock returns. Though the results are weaker, the overall effect persists. This provides assurance that the results are not driven by financial risk through leverage, but instead are driven by exposure to systematic operating risk through intersector trade relations.

C. Factor Regression Tests

While the above results show clear patterns in average returns based on network centrality, the pattern may be captured by existing factor models, such as the market model or a four-factor model that combines the factors of Fama and French (1993) and the momentum factor of Carhart (1997). The market model includes one factor, $R_M - R_F$, the excess return of the market over the risk-free rate. The four-factor Fama-French and Carhart model includes $R_M - R_F$, as well as SMB, HML, and UMD. SMB is the difference in returns between small and large firms at the aggregate level. HML captures aggregated differences in returns between high and low book-to-market firms. UMD measures the difference between stocks with high returns in the past year with stocks with low returns.

It is reasonable to imagine that network characteristics may be related to these market-wide factors. For example, the high returns associated with centrality may be explained by exposure to market-wide stock returns. Or, centrality could be related to SMB, where central firms behave more like large firms than small firms. This is of special interest given that centrality is highly related to total industry output. Correlations with HML and UMD may reflect that concentration of customers and suppliers reflects valuation or momentum. In contrast, it is possible that position in the sector network is unrelated to these factors given that the sector network represents exogenous real trade flows, rather than aggregated stock return information.

Table V presents estimates of the time-series factor regressions on five sorted value-weighted centrality portfolios. The estimates reveal a clear pattern of increasing alphas as centrality gets higher. The alpha in the lowest centrality quintile is 0.47% in the market model and 0.08% in the four-factor model. In the highest centrality portfolio, the alpha is roughly 1.5% in both the market model regression and the four-factor regression. The differences in alphas between the highest and lowest centrality portfolios is statistically significant and economically meaningful. Thus, the explantory power of centrality for stock returns is not diminished even after controlling for known factors related to firm size, leverage, and momentum.

In contrast, the average coefficient estimates for $R_M - R_F$, SMB, HML, and UMD do not display clear patterns across centrality quintiles. The loadings on SMB, HML, and UMD are not statistically different across the centrality quintiles, reinforcing the point that known risk factors do not explain the role of centrality for stock returns. The loadings on $R_M - R_F$ increase with centrality, though not monotonically. In the market model, the market beta is significantly higher for the most central industries compared to the least central industries, but this difference vanishes after including the Fama-French and momentum factors. It is also important to note that while the magnitude and statistical significance of alpha increase with centrality, so does the adjusted R^2 of the regressions. The adjusted R^2 for the least central industries is less than 50%, compared to 96% for the most central industries. This implies that standard factor models have poor explanatory power for peripheral industries, though the estimates are unbiased, whereas they have high explanatory power for the most central industries, but produce biased estimates.

D. Cross-Sectional Regression Tests

In the last set of tests, I run cross-sectional regressions on value-weighted industry-level average monthly returns from 1998 to 2002 controlling for a host of exogenous explanatory variables recorded in December 1997. Consistent with all of the prior results, in column 1 of Table VI, I find that log(centrality) positively and significantly predicts average stock returns. In the next set of tests, I include various alternative variables.

In column 2, I regress average returns on the market model β coefficient estimated from monthly returns from 1992 through 1997. I find that it is positive and significantly related to future average returns.

In column 3, I include industry-level concentration ratios calculated as the eight-firm concentration ratio of sales, using data from the Economic Census of the U.S. Census Bureau in 1997.[3] I find a negative, but insignificant relation. The negative sign is consistent with Hou and Robinson (2006), who find that firms in more concentrated industries have statistically lower returns.

In column 4, I find no direct effect of an industry's concentration of customer sectors on average stock returns. In column 5, I find a positive and significant effect, consistent with the prior results. This indicates that industries with a greater concentration of suppliers have higher average returns, even at a disaggregated industry-level of analysis. As before, in column 6, I find no direct effect of the average size of firms in an industry on its average returns. This provides more evidence that centrality is not simply a proxy related to firm size.

In column 7, I include the variable 'Industry Scope' that captures the heterogeneity of the IO industry definitions, calculated as in Ahern and Harford (2012). This variable records the percent of all NAICS codes that map to a particular IO industry. This variable provides a measure of the variation of business activities of each IO industry (Economic Classification Policy Committee, 1993; Gollop, 1994). I find a positive and weakly significant relation between industry scope and stock returns. This indicates that industries that are more heterogeneous have higher stock returns.

In column 8, I include the fraction of each industry's total input that is accounted for by compensation of employees. This variable is designed to test whether the higher returns of service sector industries is based on their greater centrality, or on the importance of human capital. I find that industries with higher relative labor expenditures have higher returns, consistent with this hypothesis.

Finally, in column 9, I create an alternative measure of centrality for robustness based on the text-based similarity measure developed in Hoberg and Phillips (2010a) and Hoberg and Phillips (2010b). Hoberg and Phillips create a firm-level network of Compustat firms based on the similarity

[3]With the exception of agriculture and public administration, concentration measures are reported for all industries. Since these data cover firms of all sizes and the vast majority of industries, they provide the most comprehensive concentration ratios available.

of product descriptions in their 10-K filings. I aggregate the 1997 data, provided on Jerry Hoberg's, to IO industry-levels by recording the total number of firms in the Hoberg-Phillips database that are in each IO-industry-pair. This provides a measure of the connection between IO industries based on asset complementarity, rather than economic transactions. Using these inter-industry connections, I calculate the eigenvector centrality for each IO industry.

This measure has important differences to the SAM network centrality. In particular, sectors without Compustat firms in 1997 are omitted from the Hoberg-Phillips network, in contrast to the complete picture of the entire economy that the SAM network provides. Second, the Hoberg-Phillips network is not directional, as industry connections are based on product similarity, rather than the directional flow of real money between sectors. Nevertheless, the two centrality measures are positively correlated and present alternative industry network relations. In Table VI, I find that the Hoberg-Phillips centrality is also strongly related to stock returns. Thus, centrality in the network of product similarity and centrality in the network of trade flows are both positively related to average stock returns.

In column 10, I include all the variables, excluding Hoberg-Phillips centrality. The level and statistical significance of the SAM network centrality is unchanged after including all of the additional variables. This reinforces the evidence that centrality is an exogenous and powerful predictor of stock returns. The market model β remains significant and the concentration of customer sectors is marginally significant. Average firm size, industry scope are both negatively related to stock returns in the full model, and labor's fraction of inputs remains positive and significant. In column 11, I run the same tests, but include the Hoberg-Phillips measure of centrality and find the SAM network centrality measure remains positive and significant. Thus, even after accounting for a wide range of possible alternative explanations, I find a strong positive relation between centrality in the intersector network and average stock returns.

IV. Conclusion

This paper provides new evidence that stock returns are partially explained by an industry's position in the network of intersector trade flows. Industries that are more central in the network have greater exposure to idiosyncratic shocks that transmit from one sector to another through trade relations. For instance, oil shocks transmit from petroleum extraction companies, to refineries, to gasoline-dependent industries such as shipping, to online retailers. Central industries, such as automakers, paper mills, and hospitals, are at greater risk of exposure to idiosyncratic shocks, because they are embedded in the center of the economy. In comparison, peripheral industries, such as sugarcane farms, musical instruments, and mattress manufacturers have a lower likelihood of exposure to a random sector-specific shock. Thus, this paper shows that idiosyncratic sector-specific shocks can spill over into customer and supplier sectors which leads to higher stock returns for more central industries.

More broadly, this paper provides new evidence to support the view that sector-specific shocks may aggregate to form economy-wide volatility. This line of research has important implications in many areas of economics. For example, technological shocks do not necessarily have to have immediate macroeconomic consequences. Instead, the spread of small innovations through an economy can produce aggregate technology shocks. Second, this line of research identifies which sectors are more important for the overall economy. This allows policy makers to consider the aggregate implications of government intervention in a particular industry (e.g., the bailouts of the financial sector and automakers). Finally, as more detailed data become available, it will be possible to construct economy-wide networks at the firm-level or even the person-level to identify how micro-level events can have macro-level implications.

REFERENCES

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, 2012, The network origins of aggregate fluctuations, *Econometrica* 80, 1977–2016.

Ahern, Kenneth R., and Jarrad Harford, 2012, The importance of industry links in merger waves, *Journal of Finance, forthcoming.*

Bernardo, Antonio E., Bhagwan Chowdhry, and Amit Goyal, 2007, Growth options, beta, and the cost of capital, *Financial Management* 36, 5–17.

Bonacich, Philip, 1972, Factoring and weighting approaches to status scores and clique identification, *Journal of Mathematical Sociology* 2, 113–120.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, *Journal of Finance* 59, 2577–2603.

Carvalho, Vasco M., and Xavier Gabaix, 2012, The great diversification and its undoing, *American Economic Review, forthcoming.*

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *Journal of Finance* 63, 1977–2011.

Conley, Timothy G., and Bill Dupor, 2003, A spatial analysis of sectoral complementarity, *Journal of Political Economy* 111, 311–352.

Economic Classification Policy Committee, 1993, Criteria for determining industries, Discussion Paper 4 Economic Classification Policy Committee.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Gollop, Frank M., 1994, The heterogeneity index: A quantitative tool to support industrial classification, Discussion Paper 2 Economic Classification Policy Committee.

Hoberg, Gerard, and Gordon Phillips, 2010a, Dynamic text-based industry classifications and endogenous product differentiation, *Unpublished working paper. University of Maryland, College Park, MD.*

——— , 2010b, Product market synergies and competition in mergers and acquisitions: A text-based analysis, *Review of Financial Studies* 23, 3773–3811.

Hong, Harrison, Walter Torous, and Rossen Valkanov, 2007, Do industries lead stock markets?, *Journal of Financial Economics* 83, 367–396.

Hou, Kewei, and David T. Robinson, 2006, Industry concentration and average stock returns, *The Journal of Finance* 61, 1927–1956.

Lewellen, Wilbur G., 1971, A pure financial rationale for the conglomerate merger, *Journal of Finance* 26, 521–537.

Lintner, John, 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *The Review of Economics and Statistics* 47, 13–37.

Long, John B., and Charles I. Plosser, 1983, Real business cycles, *Journal of Political Economy* 91, 39–69.

Menzly, Lior, and Oguzhan Ozbas, 2010, Market segmentation and cross-predictability of returns, *Journal of Finance* 65, 1555–1580.

Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425–442.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

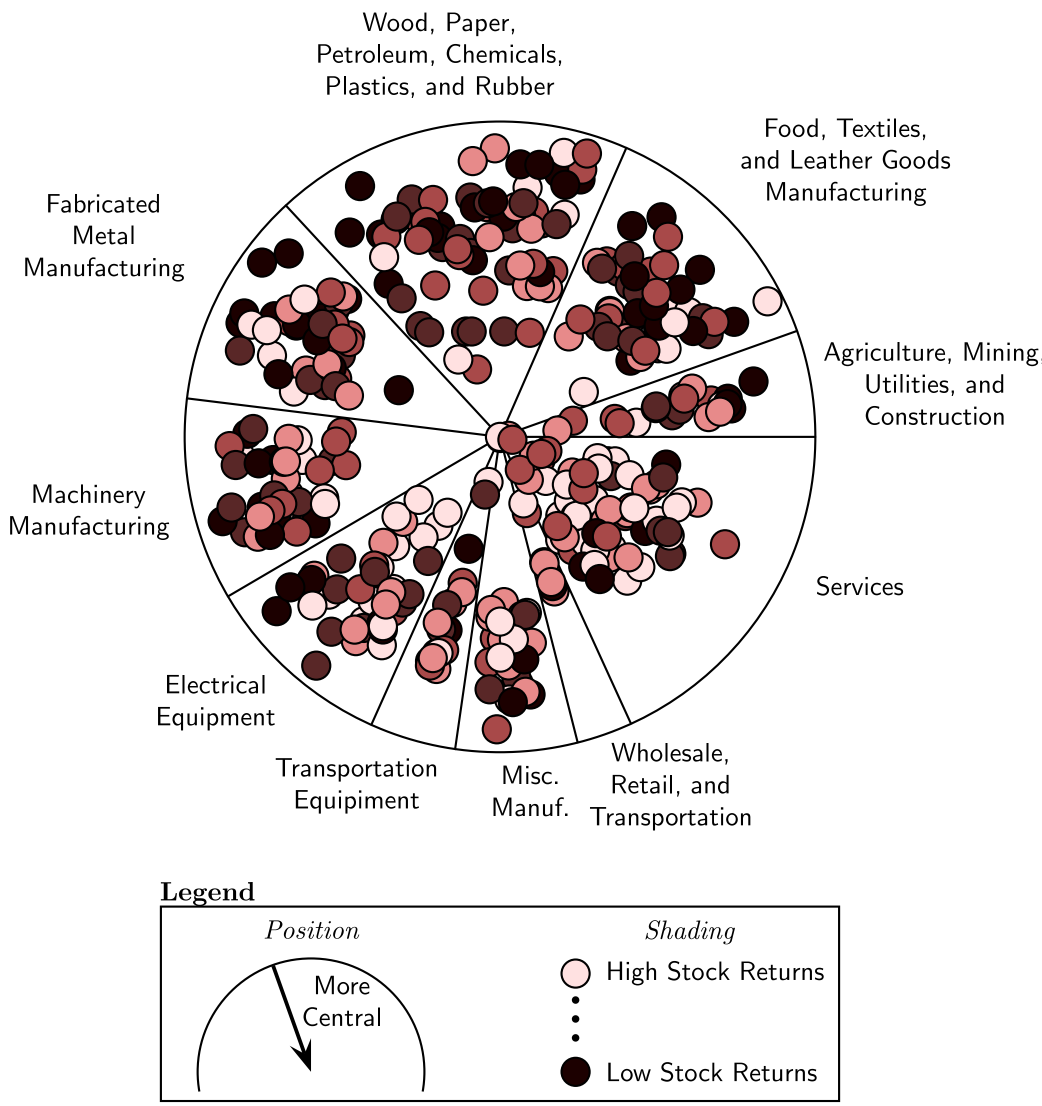

Figure 1
Network Centrality and Average Stock Market Returns by Industry
Each circle represents an industry using industry definitions from the 1997 detail-level Input-Output tables of the Bureau of Economic Analysis. Industries are categorized according to aggregate production groupings. The size of each grouping's slice of the total circle is the ratio of the number of industries contained in the grouping to the total number of industries. Within each grouping, the distance from each industry's placement in the circle to the center of the circle is proportional to log(Centrality). Centrality is the eigenvector centrality in the inter-sector network of trade flows, based on 483 sectors. Sectors are defined using the 1997 detail-level Input-Output and National Income and Product Account tables of the Bureau of Economic Analysis. Only sectors with stock return data are presented in the figure and the distance to the center of the circle is normalized to be zero for the most central industry with stock price data. The lighter is the shading of an industry circle, the higher is its time-series average value-weighted stock returns, where firm-level returns are aggregated based on firm market equity, over January 1998 to December 2002.

Table I
Social Accounting Matrix of the U.S. Economy in 1997
This table presents the economic transactions between five aggregate sectors of the economy. Panel A reports dollar transactions in millions of 1997 dollars. Panel B reports the fraction of total expenditures by column sectors received by row sectors. Panel C reports the fraction of the total receipts of row sectors that are contributed by column sectors. Data are from the 1997 detail-level Input-Output and National Income and Product Account tables of the Bureau of Economic Analysis. The 'Firms' sector is an aggregate of 479 separate industries. 'Government' does not include government enterprises, such as local transit and electric utilities, which are included under 'Firms.' The government sector in this table includes tax collections and consumption and investment expenditures for Federal and local governments. See the Internet Appendix for a complete description of each sector.

Panel A: Dollar Flows

	Expenditures				
Receipts	Firms	Households	Capital	Government	Foreign
Firms	6,530,867	5,616,840	1,472,776	1,609,666	822,163
Households	4,656,650	0	1,382,900	929,800	44,753
Capital	2,610,103	137,712	0	113,724	98,735
Government	1,309,840	1,259,550	103,600	0	5,100
Foreign	944,853	0	998	24,900	0
Total	16,052,313	7,014,102	2,960,274	2,678,090	970,750

Panel B: Supplier Network
(% of Column Expenditures Received by Row)

	Expenditures				
Receipts	Firms	Households	Capital	Government	Foreign
Firms	40.7	80.1	49.8	60.1	84.7
Households	29.0	0.0	46.7	34.7	4.6
Capital	16.3	2.0	0.0	4.2	10.2
Government	8.2	18.0	3.5	0.0	0.5
Foreign	5.9	0.0	0.0	0.9	0.0

Panel C: Customer Network
(% of Row Receipts Spent by Column)

	Expenditures				
Receipts	Firms	Households	Capital	Government	Foreign
Firms	40.7	35.0	9.2	10.0	5.1
Households	66.4	0.0	19.7	13.3	0.6
Capital	88.2	4.7	0.0	3.8	3.3
Government	48.9	47.0	3.9	0.0	0.2
Foreign	97.3	0.0	0.1	2.6	0.0

Table II
Industry Characteristics Sorted by Centrality

This table presents industry-level averages across five quintiles of centrality. Centrality is the eigenvector centrality in the inter-sector network of trade flows, based on 483 sectors. Sectors are defined using the 1997 detail-level Input-Output and National Income and Product Account tables of the Bureau of Economic Analysis. Centrality quintiles are formed using all 483 sectors, but only sectors with stock return data are presented in the table. 'Concentration of Customers' is the four-firm concentration ratio of sales outputs per industry. 'Concentration of Suppliers' is for purchases per industry. 'Log(Industry Output)' is the total dollar volume of output produced by an industry in 1997 according the BEA IO reports. 'Log(Industry Average Market Equity)' is the cross-sectional average market equity of all firms listed on CRSP in each industry in December 1997. In Panel B, time series average returns over January 1998 to December 2002 are presented for portfolios formed on centrality quintiles. Levered returns are the raw returns reported on CRSP. Unlevered returns are unlevered using debt-to-equity ratios and marginal tax rates in December 1997. Value weighted returns are value-weighted monthly using firm-level market-equity. 'Firm-level equal weighted' are portfolios formed monthly based on firm-level returns. 'Industry-level equal weighted' are portfolios formed monthly using industry-level value-weighted returns (value-weighted across firms in the industry). Statistical significance of the difference between the highest and lowest centrality quintiles is reported by t-tests assuming unequal variances. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***.

	Centrality						
	Low 1	2	3	4	High 5	1−5	t-statistic
Panel A: Monthly Returns (%)							
Levered							
Value weighted	1.21	1.93	2.25	2.29	2.44	−1.23**	(−2.24)
Firm-level equal weighted	0.28	0.97	1.30	1.23	1.42	−1.14***	(−4.69)
Industry-level equal weighted	0.10	0.68	0.95	0.85	1.22	−1.12***	(−4.14)
Unlevered							
Value weighted	1.09	1.72	2.10	2.11	2.07	−0.99*	(−1.84)
Firm-level equal weighted	0.29	0.91	1.19	1.15	1.29	−1.00***	(−4.33)
Industry-level equal weighted	0.16	0.66	0.84	0.74	1.04	−0.88***	(−3.56)
Number of Industries	59	77	82	83	84		
Panel B: Industry Characteristics							
Centrality	0.03	0.06	0.11	0.23	1.44	−1.41***	(−6.97)
Concentration of Customers	0.74	0.72	0.72	0.67	0.67	0.08*	(1.95)
Concentration of Suppliers	0.61	0.58	0.60	0.62	0.68	−0.07***	(−4.68)
Log(Industry Output)	7.74	8.53	9.11	9.84	11.27	−3.53***	(−34.67)
Log(Industry Average Market Equity)	12.33	12.65	13.05	13.25	13.82	−1.48***	(−6.41)
Log(Industry Median Market Equity)	12.15	12.34	12.51	12.46	12.57	−0.42*	(−1.89)

Table III

Industry Lists by Centrality, Firm Size, and Concentration of Customers and Suppliers

This table presents a sample of the five largest industries grouped into either the first or fifth quintile of centrality, average market equity, concentration of customers, or concentration of suppliers. In Panel A, industries are first sorted into five quintiles based on the average market equity of firms in each industry. Within each quintile, firms are then sorted into five quintiles of centrality. Centrality is the eigenvector centrality in the inter-sector network of trade flows, based on 483 sectors. Sectors are defined using the 1997 detail-level Input-Output and National Income and Product Account tables of the Bureau of Economic Analysis. Centrality quintiles are formed using all 483 sectors, but only industries with stock return data are presented in the table. In Panel B, industries are first sorted into five quintiles of the concentration of customer industries, defined as the four-sector concentration ratio of sales outputs per industry. Within each of these sectors, five quintiles are formed based on centrality. In Panel C, the first sort is based on the concentration of supplier sectors, defined as the four-sector concentration ratio of purchase inputs per industry.

	Panel A: Average Market Equity	
	Smallest Average Market Equity	Largest Average Market Equity
Lowest Centrality	Household vacuum cleaner manufacturing	Industrial gas manufacturing
	Fabric coating mills	Gold, silver, and other metal ore mining
	Musical instrument manufacturing	Mattress manufacturing
	Laboratory apparatus and furniture manufacturing	Copper, nickel, lead, and zinc mining
	Explosives manufacturing	Elevator and moving stairway manufacturing
Highest Centrality	Automotive repair, except car washes	Hospitals
	Travel arrangement and reservation services	Automobile and light truck manufacturing
	Frozen food manufacturing	Insurance carriers
	Ready-mix concrete manufacturing	Petroleum refineries
	Textile and fabric finishing mills	Pharmaceutical and medicine manufacturing
	Panel B: Concentration of Customers	
	Lowest Concentration	Highest Concentration
Lowest Centrality	Iron and steel forging	Sugarcane and sugar beet farming
	Nonwoven fabric mills	Sawmill and woodworking machinery
	Stationery and related product manufacturing	Small arms manufacturing
	Flexible packaging foil manufacturing	Burial casket manufacturing
	Software reproducing	Tortilla manufacturing

Highest Centrality	Truck transportation	Hospitals
	Semiconductors and related device manufacturing	Automobile and light truck manufacturing
	Nondepository credit intermediation	Insurance agencies, brokerages, and related
	Accounting and bookkeeping services	Aircraft manufacturing
	Commercial printing	Cigarette manufacturing
	Panel C: Concentration of Suppliers	
	Lowest Concentration	Highest Concentration
Lowest Centrality	Household laundry equipment manufacturing	Water, sewage and other systems
	Prefabricated wood building manufacturing	Ophthalmic goods manufacturing
	Ammunition manufacturing	Tobacco stemming and redrying
	Sugarcane and sugar beet farming	Musical instrument manufacturing
	Elevator and moving stairway manufacturing	Software reproducing
Highest Centrality	Motor vehicle parts manufacturing	Hospitals
	Paper and paperboard mills	Monetary authorities and depository credit
	Electronic computer manufacturing	Insurance carriers
	Aircraft manufacturing	Automobile and light truck manufacturing
	Frozen food manufacturing	Petroleum refineries

Table IV
Mean Portfolio Returns by Centrality, Size, and Product Market Concentration
This table reports average monthly portfolio returns, where portfolios are formed based on sorts into five quintiles of the average market equity of firms in the industry (Panel A), the concentration of customers (Panel B), or the concentration of suppliers (Panel C), all recorded in December 1997, using data from CRSP and the BEA (see the text for a complete description of each variable). Within each of these quintiles, industries are sorted into five quintiles of centrality. Industry-level returns are value-weighted monthly from January 1998 to December 2002, where stocks with a price less than five dollars are excluded. The 25 sorted portfolio returns are equal-weighted portfolios of industry-level returns. The table reports average monthly portfolio returns over the 60 months from January 1998 to December 2002; t-statistics in parentheses are adjusted for autocorrelation. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***.

	Panel A: Average Market Equity Quintile						
Centrality Quintile	Low 1	2	3	4	High 5	Low−High	t-statistic
1 Low	0.33	1.21	0.71	0.26	0.64	−0.31	(−0.27)
2	0.18	1.80	1.46	1.04	1.09	−0.91	(−0.88)
3	1.06	1.18	1.81	1.68	0.66	0.40	(0.63)
4	2.31	1.35	1.38	1.90	1.33	0.97	(1.53)
5 High	1.49	2.36	2.31	1.90	1.99	−0.50	(−0.75)
Low−High	−1.16	−1.15*	−1.60***	−1.64**	−1.35*		
t-statistic	(−1.59)	(−1.97)	(−3.04)	(−2.43)	(−1.86)		

	Panel B: Concentration of Customers Quintile						
Centrality Quintile	Low 1	2	3	4	High 5	Low−High	t-statistic
1 Low	0.59	0.06	−0.02	0.03	1.48	−0.89	(−1.00)
2	0.98	1.35	1.21	1.62	0.57	0.40	(0.71)
3	0.83	1.49	1.05	1.57	1.56	−0.73	(−1.25)
4	1.57	1.02	1.77	1.48	1.74	−0.17	(−0.41)
5 High	2.05	2.30	1.66	2.19	1.79	0.26	(0.48)
Low−High	−1.46***	−2.24***	−1.68**	−2.16***	−0.31		
t-statistic	(−2.89)	(−2.99)	(−2.41)	(−3.32)	(−0.34)		

	Panel C: Concentration of Suppliers Quintile						
Centrality Quintile	Low 1	2	3	4	High 5	Low−High	t-statistic
1 Low	1.66	0.72	−0.24	−0.23	0.92	0.74	(1.15)
2	0.62	0.42	1.24	1.41	0.69	−0.08	(−0.12)
3	1.44	1.27	1.76	1.50	1.71	−0.27	(−0.55)
4	1.11	1.76	1.51	0.38	2.25	−1.14**	(−2.41)
5 High	1.08	1.66	1.88	2.31	2.44	−1.36**	(−2.39)
Low−High	0.58	−0.94	−2.12***	−2.54***	−1.52**		
t-statistic	(0.77)	(−1.28)	(−3.05)	(−3.50)	(−2.56)		

Table V
Factor Sensitivities by Centrality
This table reports estimates of the Market Model, $E(R_{it} - R_{ft}) = \alpha_i + \beta_i(R_{mt} - R_{ft}) + \varepsilon_{it}$, and the Fama-French three-factor model including a fourth momentum factor, $E(R_{it} - R_{ft}) = \alpha_i + b_i(R_{mt} - R_{ft}) + s_i SML_t + h_i HML_t + u_i UMD_t + \varepsilon_{it}$. Portfolio returns are value-weighted inudstry returns formed for five quintiles of network centrality based on the 1997 BEA IO reports (see prior tables). Data are from Kenneth French's website. Observations are monthly returns from January 1998 to December 2002. Stocks with a price less than five dollars are excluded. Newey-West adjusted t-statistics are reported in parentheses. 'High−Low' reports the difference between coefficient estimates from the fifth and first centrality quintiles. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***.

		Factor Loadings				
Portfolio	Alpha (%)	$R_M - R_F$	SMB	HML	UMD	Adj. R^2 (%)
Low Centrality	0.470	0.667				41.75
	(0.845)	(5.757)				
	0.080	0.926	0.061	0.441	0.067	46.33
	(0.134)	(6.241)	(0.585)	(2.279)	(1.077)	
2	1.110	0.826				58.17
	(2.055)	(8.122)				
	0.880	1.110	0.366	0.627	−0.160	73.79
	(2.101)	(10.387)	(4.107)	(3.983)	(−3.139)	
3	1.270	1.144				80.52
	(3.035)	(16.022)				
	1.450	1.094	−0.015	−0.072	−0.098	80.45
	(3.030)	(13.016)	(−0.158)	(−0.554)	(−1.458)	
4	1.430	0.907				83.64
	(4.974)	(17.187)				
	1.420	0.932	0.142	0.100	−0.050	84.36
	(4.925)	(16.127)	(2.512)	(1.331)	(−1.076)	
High Centrality	1.550	0.969				95.99
	(9.575)	(33.620)				
	1.530	0.978	−0.039	−0.001	0.013	95.93
	(9.231)	(26.920)	(−0.947)	(−0.020)	(0.504)	
High−Low	1.080*	0.302**				
	(1.951)	(2.558)				
	1.450**	0.052	−0.100	−0.442**	−0.054	
	(2.424)	(0.373)	(−0.898)	(−2.516)	(−0.711)	

Table VI
Cross Sectional Regressions of Industry-Level Average Monthly Returns

This table presents regression coefficients and robust standard errors in parentheses. The dependent variable is average monthly value-weighted industry returns from January 1998 to December 2002. Industries are based on BEA Input-Output (IO) definitions from 1997. 'Centrality' is the eigenvector centrality of the industry in the economy-wide sector network. 'Market model β' is the coefficient on $R_M - R_F$ in a industry-level market-model regression using monthly returns from January 1993 to December 1997. 'Industry Concentration' is the four-sector concentration ratio of sales in an industry. 'Concentration of Customers' is the four-sector concentration ratio of sales outputs per industry. 'Concentration of Suppliers' is for purchases per industry. 'Log(Industry Average Market Equity)' is the cross-sectional average market equity of all firms listed on CRSP in each industry, averaged over 1/1993 to 12/1997. 'Log(Industry Scope)' is the fraction of all NAICS codes that map to the IO industry code. 'Labor's Fraction of Inputs' is the total dollars spent on labor compensation divided by total input costs per industry. 'Log(Hoberg Phillips Centrality)' is the eigenvector centrality of industries in an inter-industry network based on the text-based similarity measures of Hoberg and Phillips (2010a) and Hoberg and Phillips (2010b). Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log(Centrality)	0.339***									0.364***	0.328***
	(0.073)									(0.112)	(0.121)
Market model β		0.715***								0.597***	0.581***
		(0.210)								(0.210)	(0.212)
Industry Concentration			−0.007							0.001	0.000
			(0.005)							(0.005)	(0.005)
Concentration of Customers				0.511						0.607*	0.551*
				(0.401)						(0.316)	(0.317)
Concentration of Suppliers					1.647*					0.966	0.842
					(0.847)					(0.941)	(0.957)
Log(Industry Average Market Equity)						−0.016				−0.147*	−0.145*
						(0.075)				(0.088)	(0.088)
Log(Industry Scope)							0.218*			−0.302*	−0.313*
							(0.117)			(0.165)	(0.166)
Labor's Fraction of Inputs								1.549**		1.349**	1.289*
								(0.657)		(0.658)	(0.659)
Log(Hoberg Phillips Centrality)									0.645***		0.150
									(0.147)		(0.132)
Constant	3.498***	0.800***	1.613***	0.922***	0.259	1.523	2.712***	0.806***	4.089***	1.690	2.174
	(0.450)	(0.209)	(0.264)	(0.293)	(0.524)	(1.006)	(0.779)	(0.204)	(0.631)	(1.365)	(1.413)
Observations	385	334	357	385	385	377	385	385	385	309	309
Adjusted R^2	0.048	0.035	0.003	0.001	0.005	−0.003	0.005	0.009	0.035	0.108	0.107

Internet Appendix for
"Network Centrality and the Cross Section of Stock Returns"

This Internet Appendix provides more details on the data used in the paper. To construct the social account matrix (SAM) I combine data from the 1997 Input-Output (IO) tables and the National Income and Product Account (NIPA) tables, both provided by the Bureau of Economic Analysis. The IO tables record real economic trade flows between commodity producers and industry and final users, including government and foreign sectors. Commodities are defined as goods and services, not just manufacturing goods. Thus the IO tables provide a picture of inter-industry trade flows, final uses, and all production inputs including labor. A SAM extends the IO tables to include complete expenditures and receipts for government, a foreign sector, and a capital sector. A typical SAM separates purchasing agents into producing agents (industries), factors of production (labor and capital), institutions (households and government), capital account, and a foreign sector, each with various degrees of aggregation (e.g., high income and low income households, local and federal governments). I modify the standard SAM to combine labor and households into one agent, since I will use the SAM to form a network. Having two nodes for the same agent would distort the network analysis.

I. Input-Output Tables

To construct the IO matrix, I follow the procedure in Ahern and Harford (2012) (AH), with a few minor exceptions. See AH and the Internet Appendix of AH for a complete and detailed description of the IO tables.

Like in AH, I account for the fact that some commodities are produced by multiple industries, and some industries do not have a corresponding commodity that can be identified with an industry, by recording each industry's output as a fraction of the total production of a commodity. I also combine the construction industries in the IO tables because they all match to the same NAICS codes.

Also, like in AH, I modify the 'Make' table to include employee compensation as an industry. Though the IO tables record the compensation of employees as a commodity input in production, there is no corresponding industry that produces compensation. Because of this, employee compensation gets dropped from the industry matrices. Therefore, I create an artificial labor industry to make sure that we account for labor as an input in the industry matrices. The same problem also occurs for 1) 'Noncomparable imports,' 2) 'Used and secondhand goods,' 3) 'Rest of world adjustment to final uses,' 4) 'Indirect business tax and nontax liability,' and 5) 'Other value added.' Therefore, I alter the Make table to make sure these elements of the IO matrix are not lost when converting from industry-by-commodity accounts to a directed industry-by-industry matrix.

A second difference with AH, is that I include all industries and final users from the IO tables, including government enterprise, exports and imports, scrap, and various adjustments. In total, I create a IO matrix with 478 industries, plus the following sectors: households, capital, government, foreign, and used/inventory adjustments, for a total of 483 unique economic agents in the economy.

II. Social Accounting Matrix

I account for government receipts and expenditures by creating a government agent that collects taxes and makes consumption and investment expenditures. From the IO matrix, I aggregate government expenditures by summing the eight IO column entries for 1) 'Federal Government consumption expenditures, national defense,' 2) 'Federal Government gross investment, national defense,' 3) 'Federal Government consumption expenditures, nondefense,' 4) 'Federal Government gross investment, nondefense,' 5) 'State and local government consumption expenditures, education,' 6) 'State and local government gross investment, education,' 7) 'State and local government consumption expenditures, other,' and 8) 'State and local government gross investment, other.' The IO tables record expenditures by these agencies, but not tax receipts, which will be described below.

It is important to note that the industries classifed in the IO tables as 1) 'Federal electric utilities,' 2) 'Other Federal Government enterprises,' 3) 'State and local government passenger transit,' 4) 'State and local government electric utilities,' 5) 'Other State and local government enterprises,' and 6) 'General government industry' remain disaggregated in my analysis. These industries are represented in the IO tables as completely separate and balanced industries, with total receipts equal to total expenditures, just as are private industries. Therefore, the government agent described in the prior paragraph can be thought of as the redistributive aspect of government, whereas government enterprise is grouped with other indutstrial production.

To record household expenditures and consumption from the IO tables, I record expenditures as 'Personal consumption expenditures' and income receipts as 'Compensation of employees.' The calculation of taxes and capital investment are recorded from the NIPA tables, described below.

The expenditures of the capital account are composed of the IO entry for 'Private fixed investment', and the income receipts for capital are composed of the IO entry for 'Other value added,' which is the residual claim or profit after accounting for all other input costs.

I adjust the IO values of imports and exports to create a balanced foreign sector. The IO tables record a column 'Imports of goods and services' as negative values of final use. I transpose these entries to be positive flows to the row entries for the foreign sector. I then calculate the receipts of the foreign sector by summing this row with the IO input rows 1) 'Noncomparable imports,' and 2) 'Rest of the world adjustment to final uses.' The expenditures column for the foreign sector is recorded as the IO column entry for 'Exports of goods and services.'

Next, I aggregate a few industry categories that record small adjustments in usage and consumption. These include 1) 'Change in private inventories,' 2) 'Scrap,' 3) 'Used and secondhand goods,' and 4) 'Inventory valuation adjustment. '

Finally, I transpose all negative dollar flows from industry i to j recorded in the IO tables to positive flows from industry j to i. This allows the dollar flows in the SAM to be interpreted as the strength of the connection between two agents in a network setting.

A. Government

To incorporate taxes paid to the government, I use the 1997 NIPA tables provided by the BEA. First, from NIPA Table 3.1 on Government Current Receipts and Expenditures, I calculate the total flows from industry to the government as the sum of taxes on production and imports, taxes on corporate income, 50 percent of the contribution to government social insurance (assuming households and firms split this evenly), and current (net) transfer receipts from business. This equals \$1.174 trillion. The IO tables only record total indirect taxes paid by business to government. I take the difference between total government receipts from businesses the NIPA table and the total indirect taxes from the IO table to calculate the additional taxes not recorded in the IO table. I then allocate this extra tax across all production industries by subtracting it from each industry's value added, in proportion to each industry's total value added.

I next record taxes paid by households as the sum of Personal current taxes, 50% of contribution to government social insurance, and current transfer receipts from persons, as recorded on NIPA Table 3.1 Government Current Receipts. This is $1.259 trillion. Taxes paid by the capital sector are $103.6 billion, as recorded as Income receipts on assets from NIPA Table 3.1. Finally, taxes paid by the foreign sector are $5.1 billion, from NIPA 3.1 entry for Taxes from the rest of the world. Since I have kept government enterprises as separate entities, I do not include their surplus as part of the redistributive government agent.

Government expenditures for production output are as recorded in the IO tables. Government expenditures (transfers) to households are taken from NIPA Table 3.1, Government social benefits paid to persons for a value of $929.8 billion. Transfers to the foreign sector are recorded as the sum of Government social benefits to the rest of the world and Other current transfer payments to the rest of the world, for a total of $24.9 billion. The difference of $113 billion between government receipts and expenditures, is recorded as government expenditure paid to the capital account.

B. Households

Household income from firms is provided as Compensation of employees in the IO tables, which is equal to $4.656 trillion. Household receipts from government are as recorded above, $929.8 billion. Household receipts from capital is from NIPA Table 2.1, and is recorded as total personal income minus compensation of employees minus personal current transfer receipts. This total of $1.382 trillion represents all non-wage and non-government transfer income, including the NIPA items, 1) Proprietors' income with inventory valuation, 2) Rental income of persons with capital consumption adjustment, and 3) Personal income receipts on assets. Household income from the foreign sector is recorded from employee compensation in the IO tables.

C. Capital

Receipts to the capital account from firms is taken from 'Private fixed investment' in the IO tables. Capital receipts from households is recorded as the difference in household receipts and household expenditures. Thus this represents the saving of households of $137 billion. The contribution of the foreign sector to the capital account, or foreign savings, is the difference between foreign receipts and foreign expenditures, or, in other words, the trade deficit. Using the data from the IO tables and the other NIPA accounts, I calculate this as $98.7 billion in 1997. According to NIPA Table 4.1 Foreign Transactions in the NIPA, it was $101.4 billion. Thus, these estimates are very similar.

REFERENCES

Ahern, Kenneth R., and Jarrad Harford, 2012, The importance of industry links in merger waves, *Journal of Finance, forthcoming.*

Internet Appendix Table I
Mean Portfolio Returns by Centrality, Size, and Product Market Concentration: Unlevered Returns

This table reports average monthly portfolio returns, where portfolios are formed based on sorts into five quintiles of the average market equity of firms in the industry (Panel A), the concentration of customers (Panel B), or the concentration of suppliers (Panel C), all recorded in December 1997, using data from CRSP and the BEA (see the text for a complete description of each variable). Within each of these quintiles, industries are sorted into five quintiles of centrality. Industry-level returns are unlevered and value-weighted monthly from January 1998 to December 2002, where stocks with a price less than five dollars are excluded. The 25 sorted portfolio returns are equal-weighted portfolios of industry-level returns. The table reports average monthly portfolio returns over the 60 months from January 1998 to December 2002; t-statistics in parentheses are adjusted for autocorrelation. Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***.

	Panel A: Average Market Equity Quintile						
Centrality Quintile	Low 1	2	3	4	High 5	Low−High	t-statistic
1 Low	0.40	1.13	0.73	0.37	0.62	−0.22	(−0.24)
2	−0.01	1.55	1.33	0.95	0.96	−0.97	(−1.13)
3	0.93	1.08	1.64	1.47	0.57	0.36	(0.66)
4	1.87	1.22	1.14	1.53	1.21	0.66	(1.30)
5 High	1.14	2.06	1.99	1.61	1.57	−0.42	(−0.80)
Low−High	−0.75	-0.93^{*}	-1.26^{***}	-1.24^{**}	-0.95^{*}		
t-statistic	(−1.23)	(−1.91)	(−2.92)	(−2.42)	(−1.68)		
	Panel B: Concentration of Customers Quintile						
Centrality Quintile	Low 1	2	3	4	High 5	Low−High	t-statistic
1 Low	0.49	0.03	−0.01	−0.01	1.31	−0.82	(−1.07)
2	0.89	1.16	1.13	1.49	0.60	0.29	(0.64)
3	0.73	1.27	0.95	1.35	1.36	−0.64	(−1.30)
4	1.34	0.80	1.59	1.28	1.49	−0.16	(−0.44)
5 High	1.76	2.04	1.39	1.81	1.45	0.31	(0.66)
Low−High	-1.27^{***}	-2.01^{***}	-1.39^{**}	-1.82^{***}	−0.14		
t-statistic	(−3.28)	(−3.43)	(−2.52)	(−3.57)	(−0.17)		
	Panel C: Concentration of Suppliers Quintile						
Centrality Quintile	Low 1	2	3	4	High 5	Low−High	t-statistic
1 Low	1.48	0.63	−0.16	−0.12	0.71	0.77	(1.47)
2	0.65	0.46	1.02	1.24	0.64	0.02	(0.03)
3	1.20	1.09	1.60	1.36	1.46	−0.26	(−0.64)
4	0.98	1.52	1.35	0.37	1.93	-0.96^{**}	(−2.25)
5 High	0.88	1.43	1.55	1.91	2.00	-1.12^{**}	(−2.36)
Low−High	0.61	−0.80	-1.70^{***}	-2.03^{***}	-1.29^{**}		
t-statistic	(1.00)	(−1.44)	(−3.02)	(−3.35)	(−2.52)		

Internet Appendix Table II
Cross Sectional Regressions of Industry-Level Average Monthly Returns: Unlevered Returns

This table presents regression coefficients and robust standard errors in parentheses. The dependent variable is the average monthly value-weighted unlevered industry returns from January 1998 to December 2002. Industries are based on BEA Input-Output (IO) definitions from 1997. 'Centrality' is the eigenvector centrality of the industry in the economy-wide sector network. 'Market model β' is the coefficient on $R_M - R_F$ in a industry-level market-model regression using unlevered monthly returns from January 1993 to December 1997. 'Industry Concentration' is the four-sector concentration ratio of sales in an industry. 'Concentration of Customers' is the four-sector concentration ratio of sales outputs per industry. 'Concentration of Suppliers' is for purchases per industry. 'Log(Industry Average Market Equity)' is the cross-sectional average market equity of all firms listed on CRSP in each industry, averaged over 1/1993 to 12/1997. 'Log(Industry Scope)' is the fraction of all NAICS codes that map to the IO industry code. 'Labor's Fraction of Inputs' is the total dollars spent on labor compensation divided by total input costs per industry. 'Log(Hoberg Phillips Centrality)' is the eigenvector centrality of industries in an inter-industry network based on the text-based similarity measures of Hoberg and Phillips (2010a) and Hoberg and Phillips (2010b). Statistical significance at the 10%, 5%, and 1% level is indicated by *, **, and ***.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log(Centrality)	0.254***									0.260***	0.246**
	(0.060)									(0.091)	(0.098)
Market model β		1.044***								0.940***	0.933***
		(0.197)								(0.206)	(0.207)
Industry Concentration			−0.005							0.001	0.001
			(0.004)							(0.004)	(0.004)
Concentration of Customers				0.455						0.441*	0.421
				(0.335)						(0.264)	(0.267)
Concentration of Suppliers					1.426*					1.033	0.987
					(0.755)					(0.812)	(0.818)
Log(Industry Average Market Equity)						−0.007				−0.114	−0.113
						(0.063)				(0.072)	(0.072)
Log(Industry Scope)							0.160			−0.232*	−0.236*
							(0.099)			(0.139)	(0.141)
Labor's Fraction of Inputs								1.280**		0.972*	0.950*
								(0.550)		(0.560)	(0.563)
Log(Hoberg Phillips Centrality)									0.482***		0.056
									(0.146)		(0.116)
Constant	2.788***	0.477***	1.370***	0.807***	0.241	1.250	2.178***	0.734***	3.228***	0.879	1.056
	(0.380)	(0.153)	(0.223)	(0.241)	(0.461)	(0.848)	(0.663)	(0.162)	(0.636)	(1.132)	(1.179)
Observations	385	334	357	385	385	377	385	385	385	309	309
Adjusted R^2	0.038	0.091	0.002	0.002	0.006	−0.003**	0.003	0.009	0.028	0.150	0.147