Supply Chain Network Structure and Firm Returns

Jing Wu
Booth School of Business, University of Chicago, Chicago, IL 60637, wujing@chicagobooth.edu

John R. Birge
Booth School of Business, University of Chicago, Chicago, IL 60637, john.birge@chicagobooth.edu

Firm choices on internal or external sources, supplier selection, and coordination mechanisms play a fundamental role in the management of firm operations. The complexity and opacity of the network of connections, however, inhibit understanding of these decisions concerning the boundaries of the firm and its relationships with suppliers and customers. This paper investigates the effect of these connections on firm performance using two levels of interactions and reactions: first-order effects from direct connections in which changes in supplier and customer performance expectations may have a delayed impact on their partners' expectations, and second-order effects from the systemic exposure to the overall market network that may impact the perceived risk of an individual firm. We measure performance using firm stock returns as representing both expectations of future performance and exposure to systematic risk. Using data on the relationships of public US firms, for the first-order effect, we show that a firm's return can be predicted by its supplier lagged returns, whereas customer lagged returns have little impact. For the second-order effect, by grouping firms according to their centrality measures in the supply chain, we find a market anomaly that may represent different incentives, depending on firms' supply chain levels, to have redundant sources to increase the reliability of supply. Specifically, upstream manufacturing firms earn lower returns when more central in the network, while downstream firms in the transportation, wholesale, and retail sectors that are more central in the network earn higher returns. Our results are robust after controlling for common asset pricing factors, industry effects, and industry concentration. The results indicate that expectations of future performance may proceed from an upstream supply chain level to a downstream partner and that risk propagation in supply chain networks depends on more than just centrality, suggesting that upstream firms may have greater motivation to form connections to reduce risk than do downstream firms.

Key words: supply chain, lead-lag effect, network centrality, systematic risk

---

1. Introduction

Firms do not exist in isolation but are linked to each other through supply chain relationships. The firms and their supply chain relationships compose the supply chain network, in which the links transmit idiosyncratic shocks[1], such as changes in a firm's individual performance expectations. Assessing the relative costs and benefits of adding, deleting, and absorbing supply chain connections

[1] "Idiosyncratic shocks" in this paper means firm-level shocks, which may be correlated across firms depending on the business characteristics such as industry sector and geographic location.

naturally gives rise to many questions such as the following that we pose in categories as first-order and second-order effects respectively. First, from the shock transmission perspective, since shocks may be transmitted at different speeds and at different intensities, what are the effects of these shock transmissions and how do upstream and downstream transmissions differ? Second, from the risk management perspective, since the idiosyncratic shocks transmitted along the supply chain network may depend on each other, do firms strategically choose a supply chain network structure to mitigate risk and how does this effect depend on the firms' industry and market positions?

Previous literature has studied the first question both at the industry level and at the firm level. At the industry level, for example, Menzly and Ozbas 2010 find strong own lagged effect and both upstream and downstream cross-prediction effects across industries using BEA (the U.S. Bureau of Economic Analysis) input-output data; Shahrur et al. 2010 extend that methodology to international trade. Using recent observations, Fruin et al. 2012 study different time horizons for trailing cross-industry lagged effects and find that longer-term (more than three-month) frequency signals are not statistically significant. While industry relationships may affect an individual firm, they also reflect within-industry lag effects in which large firm returns generally lead those of smaller firms (see, e.g., Menzly and Ozbas 2010), possibly masking the impact of a firm's direct relationships. For literature at the firm level, Hendricks and Singhal 2003 find evidence that firm returns decrease at the announcements of supply chain glitches, particularly production or shipment delays. In addition, Cohen and Frazzini 2008 find evidence of return predictability in the supply chain, providing a test of investors' attention constraints, while Kelly el at. 2013 build a model of upstream shock transmission for firm level volatility and find that size dispersion and volatility dispersion move together. At a more refined level of analysis, Atalay et al. 2013 examine firms' ownership of production chains and find no clear evidence for intra-firm trade (suggesting different reasons for vertical integration). To the best of our knowledge, our results are significantly different from the previous studies as we are the first to examine the differences between supplier firm shock and customer firm shock transmission, for both the intensity and the speed. We also show a structural diffusion mechanism at the firm level compared to the industry level result by Menzly and Ozbas 2010.

To address the question of relative upstream and downstream impact, we develop a theoretical framework in which shocks propagate through the supply chain in both directions, with possible contemporaneous and lead-lag effects. Using cross-sectional supply chain data, we construct a relationship-weighted map quantifying firm-level supply chain structure within the U.S. economy. We first test for the customer lagged effect documented by Cohen and Frazzini 2008 using recent data and find that the customer lagged effect is no longer significant. Interestingly, we still observe significant own lagged effect and supplier lagged effect. We also find that a supplier lagged effect trading strategy yields significant abnormal excess returns in back-testing. We further investigate the return information diffusion for firms operating in different industries according to the first two digits of the North American Industry Classification System (NAICS) standard, which define the large industry sectors[2], and find that the supplier lagged effect exists in most industries.

We study the shock transmission as reflected in firm returns information for two principle reasons. First, firm return data has higher frequency than operational measures such as revenues and profit that are generally only reported quarterly. The frequency of trades of a firm's shares provides us with a sufficient number of samples in the chosen horizon to conduct tests of relationship impact. Second, firm return data endogenizes operations information and thus gives cleaner information on the expectation and the riskiness of firm earnings than real economic measures. Since stock returns reflect information updating, the lagged effect between supplier and customer firms is a joint test of both investor inattention to supplier chain information and the real effect of supply chain shock transmission delay. To consider alternative mechanisms for the lag effect we observe, in robustness tests, we control for common asset pricing factors and rule out alternative explanations as reported in previous literature, including institutional holding, trading volume, analyst coverage, and market capitalization.

The second question we address is related to systematic risk as a second-order factor in risk transmission reflecting global properties of the network. The standard asset pricing models suggest that exposure to systematic risk determines stocks' expected returns. Those models, including the capital asset pricing model (CAPM) of Sharpe 1964 and Lintner 1965, the Fama-French three-factor model, and the extension to a fourth factor by Carhart 1997, all propose common factors that measure firms' exposure to systematic risk. CAPM treats the market risk as the factor of non-diversifiable risk, generally proxied by the market premium, the difference between the market return and the risk-free rate. Those components of returns that cannot be explained by CAPM have been traditionally referred to as "anomalies," among which the most well known are the size effect, the value effect, and momentum. Recognition of the size effect dates back at least to Banz 1981, who finds that average returns on small stocks are too high in the cross-section of returns given their market betas. The value effect is first recorded by Rosernberg et al. 1985, who find that average returns of stocks in the cross-section are positively related to the ratio of a firm's book value to its market value. Building on these observations, Fama and French 1993 proposed the three-factor model including a portfolio's exposure to the small-cap class and the high book-to-market ratio class. The additional momentum effect refers to the positive relation between an

[2] On the one hand, we wish to use fine-grained industry classifications so that firms in unrelated lines of business are not grouped together. On the other hand, using too fine an industry classification results in portfolios that are statistically unreliable. Choosing first two-digit classifications strikes a balance between these two concerns.

asset's current returns and its recent historical performance, which is based on the observation that stocks that performed relatively well in the past tend to have higher returns in the short run. Momentum was first studied by Jegadeesh and Titman 1993 and was incorporated as a fourth factor by Carhart 1997.

Even though the standard asset pricing models explain a portfolio's return quite well, other factors (in particular, liquidity as shown in Pástor and Stambaugh 2003) may also influence systematic risk. More importantly, since the standard asset pricing models generally identify risk using ex-post correlation between a portfolio's returns and market factors, they do not reason the ex-ante determinants of a firm's exposure to systematic risk. To address this question, we argue that the correlated supply chain relationships in aggregate determine systematic risk. Specifically, holding the supply chain network structure sufficiently stable for a short period of time, this structure is an exogenous and ex-ante identifiable source of cross-sectional variation. In line with this logic, the fundamental assumption we make is that a firm's systematic risk is formed from the aggregation of idiosyncratic shocks, which are likely to be transmitted to supply chain partners. Recent theoretical and empirical evidence supports this view. Based on a theory of network transmission of sectoral shocks, Acemoglu et al. 2012, for example, show that microeconomic idiosyncratic shocks may lead to aggregate fluctuations. In addition, Carvalho and Gabaix 2013 present empirical evidence that volatility in aggregate national output is driven by sectoral shocks. Kelly el at. 2013 also show evidence that the supplier chain network is an important determinant of firm-level volatility. While these observations at aggregate levels give an indication of systematic risk transmission at aggregate levels, they do not address how shocks propagate across individual firms and how firms' operational decisions about suppliers are related to risk mitigation motives. This paper aims to help fill this information gap.

Using a network constructed by the supply chain connections to understand systematic risk is appealing because it mirrors the intuition of most asset pricing models, where systematic risk is not driven by an asset's own idiosyncratic risk. Instead, an asset's exposure to systematic risk is based on its relationship with the entire economy. Following this logic, the underlying source of systematic risk should also reflect the relationship between an asset's economic fundamentals and overall economic fundamentals. This relationship is precisely what the supply chain network captures. The position of a firm in the supply chain network can be constructed as a proxy for its exposure to the overall economy.

To address the hypothesis that supply chain network structure is associated with systematic risk, we group firms in quintiles according to their network centrality. The most similar research to ours is that of Ahern 2013, which argues that industries that are more central in the economic network of intersectoral trade earn higher stock returns than industries that are less central. This is because, at the industry level, links are hardly substitutable; thus, operational hedging (substitution of different inputs or outputs in response to shocks) is difficult. Taking input links as an example, if an industry requires inputs from multiple other industries, it is exposed to higher risk because any shock to its supplier industries affects its production. However, we argue this finding may not be identical at the firm level since now links may be substitutable; thus, the correlation among idiosyncratic shocks matters at this level.

It has been well known that operational hedging can be used to mitigate idiosyncratic noise in the supply chain, as shown, for example, in Anupindi and Akella 1993. On one hand, if the idiosyncratic shocks of supply chain partners are positively correlated, a firm with more links is exposed to higher systematic risk due to aggregation of shocks; thus, it should have higher returns on average. On the other hand, if the idiosyncratic shocks of supply chain partners are hedged away due to their independence, a firm with more links is actually exposed to lower systematic risk and should have lower returns. Interestingly, both possible phenomena are observed in our results after controlling for common pricing factors and other alternative explanations. While more numerous suppliers and centrality are associated with lower returns for manufacturing firms, increased input links correlate with higher returns for logistics and transportation firms. We interpret these differences as manufacturers' relative ability to hedge and to take advantage of competencies not directly related to specific products (as shown in Atalay et al. 2013).

The two above questions examining the supply chain network structure's implications on firm returns can be unified in the basic net present value formula as follows, which determines a firm's valuation, as well as its return performance:

$$p_t = \sum_{s=0}^{\infty} e^{-(r_s+\delta_s)s} d_s, \tag{1}$$

where d_s is the expected dividend paid, r_s is the expected discount factor, and δ is the risk premium. The first-order effect changes in the expectations of a firm's performance in each future period, i.e. d_s. The second-order effect captures the exposure of that performance to market risk premium, i.e. δ_s, the firm faces. Those two effects together jointly affect a firm's returns. Our objective is to see how supply chain position and structure affect these two aspects of firm valuations.

The rest of the paper is structured as follows. Section 2 introduces the theoretical model and hypotheses for the first-order effect from direct connections and the second-order impact from systemic exposures through the network. Section 3 describes the supply chain data set we use in this study. Particularly, we introduce a data set from a major financial data company, which captures much richer cross-sectional information than the commonly known Compustat segment data. Section 4 examines the empirical test results. We show that a firm's return can be explained by its one-month supplier lagged returns and that more central manufacturing firms earn lower returns on average, while the opposite is true for logistics firms[3]. Section 5 concludes the paper.

2. Models of Supply Chain Network Effects on Firm Returns

In this section, we present our first-order (direct) and second-order (indirect) effect models of firm performance as reflected in stock returns. For the first-order effect, we propose a model in which the supply chain network transmits firm return shocks through direct firm connections both contemporaneously and with a one-month lag. With this model, we can then investigate the speed and the intensity of shock transmission for both upstream and downstream directions and formulate hypotheses on the relative importance of supplier influence versus customer influence for the current period and the one-month forward period. For the second-order effect[4], we propose that network centrality in the supply chain network can explain firms' exposure to systematic risk. We hypothesize that some network positions may be aggregators of correlated idiosyncratic shocks, leading to higher systematic risk, while others may be connected to relatively independent sources, reducing systematic risk effects.

2.1. First-order effects

For this network model, we suppose that firms compose the nodes of the network and that their sales relationships form directed links. We let sales determine the link strength, which is similar to what is proposed by Menzly and Ozbas 2010, in which the relationship weight is computed using the flow from one industry sector to another, and in Kelly el at. 2013 for relative firm influence on growth. This relationship is intuitive since firms are likely to be affected more if a major supplier or customer experiences a shock than if the shock comes from a minor supplier or customer. For the annual sales from firm i to firm j, we use $sales_{ij}$, which is then an output from firm i and input to firm j and will be weighted by the total sales of firm i as an output and by the total sales of firm j as an input. In this model, we assume that the supply chain relationships are sufficiently stable for a short period of time. Particularly, for our empirical tests, we assume that the supply chain structure is predetermined and exogenous to stock returns for the monthly window from July 2011 to June 2013, a total of 24 time series observations, and that this information should also be accessible to investors ex ante.

We let w_{ij}^{in} denote the input supplier weight for j as a fraction of i's procurement and let w_{ij}^{out} denote the output customer weight for j as a fraction of i's sales:

[3] From now on, we use a broad definition of logistics firms to include all firms that add value in the logistics process such as the storage, transfer, and distribution to consumers, which includes firms in the transportation, wholesale, and retail sectors.

[4] We call this effect *second-order* because it reflects not just the effect of individual connections but of the multiplicity of connections and their potential interactions..

$$w_{ij}^{in} = \frac{sales_{ji}}{Total\ Procurement_i} = \frac{sales_{ji}}{\sum_{k=1}^{N} sales_{ki}}, w_{ij}^{out} = \frac{sales_{ij}}{Total\ Sales_i} = \frac{sales_{ij}}{\sum_{k=1}^{N} sales_{ik}}.$$

We propose that these weights relate to the propagation of return shocks through the network with common damping parameters $\beta_k, k = 1, \ldots, 5$, which correspond to the rate of propagation from own lagged effect (one-period lagged own returns), supplier lagged effect (one-period lagged weighted output returns), customer lagged effect (one-period lagged weighted input returns), concurrent supplier weighted returns, and concurrent customer weighted returns. We then define $r_{i,t}$ as the return of firm i in month t, which is a linear combination of its own one-month lagged effect, weighted sum of supplier and customer one-month lagged effect, weighted sum of supplier and customer returns, as well as its own idiosyncratic shocks:

$$r_{i,t} = \alpha + \beta_1 r_{i,t-1} + \beta_2 \sum_j w_{ij}^{in} r_{j,t-1} + \beta_3 \sum_j w_{ij}^{out} r_{j,t-1} + \beta_4 \sum_j w_{ij}^{in} r_{j,t} + \beta_5 \sum_j w_{ij}^{out} r_{j,t} + \epsilon_{i,t}. \quad (2)$$

The coefficients α and β_k, $k = 1, \ldots, 5$ are then to be estimated; $\sum_j w_{ij}^{in} r_{j,t-1}$ is the one-month supplier lagged effect, $\sum_j w_{ij}^{out} r_{j,t-1}$ is the one-month customer lagged effect, $\sum_j w_{ij}^{in} r_{j,t}$ is the concurrent supplier return, and $\sum_j w_{ij}^{out} r_{j,t}$ is the concurrent customer returns. This model is in accordance with the valuation model in (1), since it explains the relative changes in expected dividends as a result of expected cash flow shocks to customers and suppliers. The lag effects represent delays in the diffusion of these expectations . In our empirical tests by both pooled OLS and Fama-MacBeth, we also introduce common risk factors into (2) to examine the independent effects of the relationships.

From the above definition, both the in-degree weights and the out-degree weights are normalized such that $\sum_j w_{ij}^{in} = \sum_j w_{ij}^{out} = 1$ and $w_{ii}^{in} = w_{ii}^{out} = 0$. For firms which do not have a supplier or customer recorded in our data, we use industry supplier returns or customer returns to avoid possible singularity in the ordinary least square estimation. The industry returns are value-weighted by other firms in the same industry according to the full NAICS code classification.

From basic operations theory, firm cash flows depend on reliable inputs from suppliers and orders from customers. We, therefore, expect to find strong positive relationships between firm performance expectations, hence, contemporaneous stock returns and those of suppliers and customers. Independent changes in stochastic discount factors should also affect firms and their supply chain partners in the same directions.

Theory for the presence of a lagged effect is less consistent. The form of lagged effect we consider here is serial autocorrelation (as opposed to consistent relative performance of winners versus losers as in the common definition of a momentum factor). For individual firms, some rational theories predict positive serial correlation (e.g., Johnson 2002), while others predict negative serial correlation (e.g., Berk et al. 1999). Sagi and Seasholes 2007 also provide a firm model that allows for either positive or negative serial correlation depending on the firm's growth prospects and costs of operation. Behavioral theories generally support positive autocorrelation (underreaction, e.g., Barberis et al. 1998) or negative autocorrelation (overreaction, e.g., Bondt and Thaler 1985). Empirical findings generally indicate short-term (and longer term over one year) negative autocorrelation (e.g., Fama and French 1988) and intermediate term positive autocorrelatioourn (e.g., Jegadeesh 1990, Jegadeesh and Titman 1993). Portfolios of firms sorted by size (Brennan el at. 1993) and industry groups (Moskowitz and Grinblatt 1999) also exhibit short-term positive autocorrelations. In addition, other issues, such as trading inactivity, can create autocorrelation.

Our focus here is on the relationships among firms, which makes direct predictions about lagged effects even more ambiguous. For example, positive one-month autocorrelation across industry groups may imply observed positive serial correlation with suppliers (who may serve the entire industry) but negative one-month serial autocorrelation for individual firm returns may imply negative one-month serial correlation for suppliers without a diversified customer base.

If investors pay limited attention to supply chain relationships, both supplier and customer lagged effects can be supported by operations management theory. When a supplier receives an idiosyncratic shock, its customer firms may be affected by the supplier's disruption due to the delivery lead time and the friction in switching suppliers. Such supply disruptions have significant and lasting impacts on the customer's share price as Hendricks and Singhal 2003 show in event studies. Effects may also appear first with suppliers if buyers observe private signals of future prospects and pass on these expectations to suppliers in the form of new contract terms or order quantities which change cash flows of the supplier before the buyer. If investors pay limited attention to such events or information about the relationship is slowly diffused, then we may observe a lag in the shock effect from the supplier to the customer.

A possible example of this form of supplier lagged effect appears in the aftermath of the Philips semiconductor fabrication plant fire in March 2000 (Wall Street Journal 2001). While the severity of the disruption was not immediately known, potential market reaction appeared in the price of Philips's shares (PHG), which dropped 13% in value in March 2000. The stock of Ericsson (ERIC), a major customer of Philips which relied on this plant for cell phone chips, was only down slightly (2%) in March 2000 but then declined by 6% in April and 7% in May 2000 (and steadily for the next several years), indicating a possible lagged effect from the supplier disruption. Another Philips customer supplied by this plant, Nokia (NOK), had a much different experience, rising 12% in March 2000 and another 1% in each of April and May 2000. In contrast to Ericsson, Nokia also had alternative suppliers (who may in fact have benefited from the Philips disruption) and did not experience a significant production disruption. In this case, Nokia appears to have benefited from the second-order interaction effect of having multiple supplier relationships. We explore this effect in more detail in the next section.

In addition to the supplier disruption effect, when a customer firm receives an idiosyncratic shock, its supplier firms may be affected by the customer lagged effect due to a change in future production orders. Cohen and Frazzini 2008 give an example of Callaway Golf Corporation and Coastcast, a manufacturer of golf club heads, in which Callaway's price dropped significantly in June 2001 while Coastcast's dropped proportionally in July 2001. While Cohen and Frazzini 2008 found a significant customer lagged effect, their exposure of this return relationship and associated trading strategy may have motivated investors to pay greater attention to these supply chain relationships and to eliminate this effect. Potential evidence of increased awareness include news media (e.g., Yahoo Finance 2013), which have started to cover the customer lagged effect strategy, and the creation of investment products using such strategies for retail investors (e.g., Motif Investing 2014). Moreover, investment banks have published white papers on the past performance of the customer lagged effect (e.g., Salvini et al. 2012, Cahan et al. 2013) and have developed relevant research products on their trading platforms (e.g., Balch 2013). Even if investors corrected the inefficiency regarding customer relationships revealed in 10Q filings as used in Cohen and Frazzini 2008, we still would not know whether investors have recognized all customer information. We may or may not, therefore, observe a significant customer lagged effect in our tests using more complete supply chain information. Therefore, we consider the possibility that underreaction or investor inattention may persist or that the publication of these results may have alerted investors sufficiently to devote greater attention to customer connections[5].

Overall, we hypothesize significant supplier and customer effects for current period returns but have alternative hypotheses of positive or insignificant customer effects. We do, however, hypothesize that it is still possible to observe returns predictability using supplier lagged returns since that effect may be less salient to investors.

HYPOTHESIS 1. *Concurrent supplier and customer returns explain a firm's returns.*

HYPOTHESIS 2. *Supplier lagged returns predict a firm's returns significantly.*

HYPOTHESIS 3. *(A) Customer lagged returns predict a firm's returns significantly. (B) Customer lagged returns do not significantly predict a firm's returns.*

[5] According to proprietary information from some anonymous hedge fund managers, customer lagged effect has been fully exploited after the appearance of Cohen and Frazzini 2008.

2.2. Second-order (systematic risk) effects

In this section, we investigate the supply chain network and firms' exposure to systematic risk, the second-order impact of aggregate shocks across multiple relationship levels. We particularly model network centrality and its risk implications. Following the variety of patterns of shock transmission that appear in models such as in Acemoglu et al. 2012, we assume that some network positions may be aggregators of correlated idiosyncratic shocks while others have connections that tend to dissipate idiosyncratic shocks and reduce systematic risk.

Our fundamental underlying assumption is that a firm's systematic risk is formed from the aggregation of idiosyncratic shocks, which are then likely to be transmitted to their supply chain partners. Those idiosyncratic shocks may not be independent of each other and may be correlated exogenously. The exogenous correlation is irrelevant for the supply chain relationships, meaning that idiosyncratic shocks are correlated with each other even if there is no sales link between the firms. Geographical proximity and sector proximity are examples of such exogenous factors that may produce correlation, e.g., geographically close firms may tend to have correlated idiosyncratic shocks. An earthquake or regional political unrest is likely to affect all firms that operate in the area, regardless of their industrial sectors. Sector proximity, on the other hand, may produce correlation as firms in the same industry face similar changes in resources or technologies. For example, a discovery of a large gold mine would possibly affect all mining firms in precious metal, or the new release of a popular tablet or a smart phone may be a simultaneous negative shock to other competing firms. Therefore, even assuming the network structure is uniformly distributed, where the no-connection network and the fully connected network are two extreme cases, idiosyncratic shocks may not be independent of each other.

Firms can mitigate supply risk or demand risk by choosing partners with which the idiosyncratic shocks are less correlated. As we observed regarding Nokia, their having multiple supplier relationships apparently helped them absorb the shock of the Philips fire, which, while idiosyncratic, could have had a ripple effect, as in Acemoglu et al. 2012, across the economy. We suppose this may be the case for other manufacturing firms, which often seek multiple less correlated suppliers to provide input materials, i.e., multiple sourcing, and which tend to take advantage of efficient organizational processes to enter different levels of the supply chain even when those entries have no physical (direct input or output) connections to parts of the firm operating at upstream or downstream supply chain levels. This observation in Atalay et al. 2013 of the prevalence of firms with disconnected production units at distinct supply chain levels suggests a natural risk mitigation mechanism in manufacturing that reduces the systematic risk of a firm that creates such connections.

However, not all firms are able to diversify their suppliers or customers (e.g., diversifying geographically linked shocks) or to enter different levels of the supply chain that may mitigate sectoral risks, resulting in a systematic risk exposure. Firms in the logistics industry may be such examples. Logistics firms such as transportation and warehousing usually serve other businesses which are close in geographical or sector distance. Their input resources (direct equipment and supplies) may also be limited in geographical diversity as may be their abilities to employ their organizational capabilities from this industry at different levels of the supply chain. They also do not face the hold-up problem of a manufacturer, such as Ericsson, where a disruption to a single supplier can shut down all production. This multiplier effect creates an incentive for creating uncorrelated relationships that is not present for wholesalers, retailers, and logistics firms whose individual suppliers rarely can hold up all of their operations.

As noted, manufacturers also may have more opportunities than logistics firms to exploit management expertise in different sectors. For example, while an automotive components manufacturer may be able to exploit its manufacturing expertise to move up the supply chain to fabricate plastic molded parts, a trucking firm that consumes automotive components for service parts may not have a particular advantage in entering that or other supplier markets. For firms in the trucking company's position, idiosyncratic shocks at partners may be more likely to be correlated, thus causing a ripple effect. As a result, they may be exposed to higher risks if they are in more central positions of the logistic firms' supply chain network.

To illustrate better, we use the following model to show a demonstrative example. Suppose an economy with 2 regions (A and B) and 3 potential future states with equal probability ($Prob\,(S = S_i) = \frac{1}{3}$, $\forall i \in \{1, 2, 3\}$): S_1: both A and B function; S_2: A cannot produce while B can; S_3: B cannot produce while A can.

Next, suppose we have 4 firms in the economy, 3 manufacturers and 1 distributer. The manufacturers have limited production capacity and produce a payoff of 1 (due to the fixed production capacity) as long as one of their input regions functions. Firm 1, 2 and 3 are manufacturers. Firm 1 only sources input from region A, Firm 2 only sources input from region B, and Firm 3 sources from both regions. Firm 4 is the distributor which connects to both region A and region B with a fixed cost of 1 in all states. Therefore, in each of the states mentioned above, the payoff for these 4 firms are as given below:

$$\Pi_1 = \{1, 0, 1\}, \Pi_2 = \{1, 1, 0\}, \Pi_3 = \{1, 1, 1\}, \Pi_4 = \{1, 0, 0\}. \tag{3}$$

Suppose we have a representative mean-variance investor. Let $\mu = [\mu_1, \mu_2, \mu_3, \mu_4]$ denote the firms' expected return. Then we will have $\mu_3 < \mu_1 = \mu_2 < \mu_4$[6], i.e. the manufacturers have lower risk

[6] See Appendix for proof

than the distributor, and the dual sourcing manufacturer is less risky than the single sourcing manufacturers.

In sum, our arguments support the presence of lower systematic risk for better connected manufacturing firms and higher systematic risk for more central logistics firms. We then state the following hypotheses.

HYPOTHESIS 4. *For the manufacturing industry, more central firms earn lower stock returns on average due to their exposure to lower systematic risks.*

HYPOTHESIS 5. *For the logistics industry, more central firms earn higher stock returns on average due to their exposure to higher systematic risks.*

We use equity returns over other metrics to focus on systematic risks alone since other factors such as product variety are endogenous in the returns information. In Section 4, we present measures of centrality that we then use to test these hypotheses.

3. Supply Chain Data

A major difficulty in studying supply chain networks is the observability of the network. For tractability, we limit our attention to the supply chain network formed by publicly listed firms in the U.S. Therefore, we omit private firms, the foreign sector, government, and household consumption from our consideration. Public firms disclose supply chain data in a variety of ways, including but not limited to public filings, conference call transcripts, capital markets presentations, sell-side conferences, firm press releases, product catalogs, and firm websites. Some information is disclosed mandatorily, while other is disclosed voluntarily due to value-maximizing managers' incentive to accommodate the capital markets, as shown, for example, in Ellis et al. 2012.

Mandatory supply chain disclosure requirements among public firms vary globally. In the United States, under the Securities and Exchange Commission's (SEC's) Statement of Financial Accounting Standards No. 14 (SFAS 14), "if 10% or more of the revenue of an enterprise is derived from sales to any single customer, that fact and the amount of revenue from each such customer shall be disclosed" in interim financial reports issued to shareholders (including annual and other quarterly reports). The segment part of the Compustat database, which has about 30 years of time-series records, captures this information. In addition, some non-major customers, which compose less than the 10% threshold of a firm's sales, are also voluntarily disclosed in public filings and thus captured by Compustat.

In recent years, financial data firms such as Bloomberg and Standard & Poor's have endeavored to fill in the missing relationships beyond the public filings. The Bloomberg Supply Chain Data (SPLC) function, available on the Bloomberg terminal application, provides the business relationships between many firms in terms of the flow of sales. More than half of the relationships in

Bloomberg SPLC are not, however, quantified (with only the existence of a directed link, i.e., the names of the supplier firm and the customer firm, indicated), but other firm pairs include an estimate of sales based on one (or more) of the possible public sources. We do not use the unquantified relationships in this paper (leaving that for future research). For the quantified relationships with actual sales amounts, Bloomberg computes the relationship percentage between firms on both a customer (revenue) and supplier (cost) basis. Bloomberg SPLC uses a variety of sources, including the public filings, for the quantified relationships. The reliability of the data set is documented in that every quantity captured is backed up by a source, which is accessible on the Bloomberg terminal.

Bloomberg keeps track of about 26,000 public firms worldwide in their universe, among which about 4,500 are US firms. Of this number, a total of 2,152 U.S. firms in SPLC have quantified supply chain data. This reduction in coverage from all public firms to those with quantified relationships underscores the difficulty in collecting supply chain information, even after investigating other sources beyond the public filings.

Since Bloomberg SPLC also uses public filings, the Compustat segment data is a subset of SPLC, which we validate by data merging. The public filings represented in the Compustat segment only contribute to fewer than 10% of the relationships in the Bloomberg SPLC data, as most quantified relationships are created by Bloomberg's estimates. According to Bloomberg documentation available on its terminal (SPLC<GO>), to create supply chain estimates, Bloomberg first constructs an exhaustive list of customers and suppliers to a firm based on disclosures found in all sources. Analysts then review the company's business model to understand how the individual segments are tied into its customers and/or suppliers, then break the revenue stream (as disclosed in company filings) down to its most granular level and match customers/suppliers to specific revenue or product streams where the relationship most likely resides. For example, the analyst would typically connect a semiconductor manufacturer with the personal computer segment of an electronics manufacturing firm.

The advantage of Bloomberg SPLC is that it captures richer cross-sectional information than public filing data alone. Unfortunately, Bloomberg SPLC is, however, only a cross-sectional data set with the latest annual relationships; so it does not offer archival data as in the Compustat segment. This is mainly due to the fact that estimates of historical sales are both arduous and difficult. Due to the time series data limitation, we use a two year sample period by assuming the supply chain network remains unchanged. Since our data have richer cross sectional information, we have a more detailed model specification than previous literature[7]. Since SPLC is a newly created product

[7] In unreported tables we replicate our findings at 12-month (July 2012 to June 2013), 18-month (Jan 2012 to June 2013) and 30-month (April 2011 to September 2013) windows. The results are qualitatively identical, showing the robustness of our assumption on the stable supply chain relationships.

and Bloomberg updates the information on firms in its universe frequently, including supply chain news, we may, however, anticipate time series data in the future.

We merge the 2012 cross-sectional data from Bloomberg SPLC and the Compustat segment, both as of June 2, 2013. Since the Bloomberg terminal is designed mainly for practitioners, the natural identifier for firms is the ticker symbol. The ticker symbol, however, tends to change frequently over time and to have duplicates; hence, we first automatically merged the dataset using both ticker and CUSIP and then hand-matched those if at least one of the identifiers did not match. As expected, Bloomberg SPLC captures the relationships in Compustat but with some newer updates using the estimates. For such situations, we average the values from both data sets and delete the duplicate relationship. We note that Bloomberg SPLC includes a few customer relationships above the 10% threshold that do not appear in the Compustat data, suggesting that it is possible that firms may conceal major customers in public filings to mitigate the costs of aiding competitors as discussed in Ellis et al. 2012.

After data cleaning, 11,819 U.S. domestic relationships are left, of which 865 are from public filings and 10,954 from Bloomberg estimates. This set then provides richer cross-sectional information than the Compustat segment data, which only captures an average of 1,124 relationships per year in the past 30 years according to Cohen and Frazzini 2008. Since the majority of the data is based on the Bloomberg database, we use SPLC to refer to our merged supply chain network data.

Even though our data is downloaded contemporaneously, actual report dates for both public filings and proprietary estimates vary due to different reporting and estimation dates. Figure 1 shows the distribution of SPLC's report dates. The earliest report date for our data set is April 3, 2012, while the latest report date is June 2, 2013. The median report date is Feb. 19, 2013 while 52.9% of the report dates concentrate in the first four months of 2013. Since supply chain relationships are sufficiently stable over short horizons, we assume the cross-sectional data set reflects supply chain network structure for the monthly window from July 2011 to June 2013, a total of 24 time series observations. We downloaded the monthly firm returns from the Center for Research in Security Prices (CRSP) within that window, which covers three exchange platforms in the U.S. market, NYSE, AMEX and NASDAQ, and 99.72% of the firms in the SPLC. The 6 tickers missing in CRSP for the selected period do not affect our results since they are missing, either due to recent listings (DXM and ENVS) or delistings due to bankruptcy or otherwise very low stock prices (CRCV, FOHL, PCXCQ and VLTC), and might have undesired liquidity effects if included.

Since our data does not capture the complete supply chain network, it is important to understand any systematic biases. Using the closing market value on the last day of 2012, we compare the coverage of our data to the CRSP universe in terms of firm size distribution. The log-size

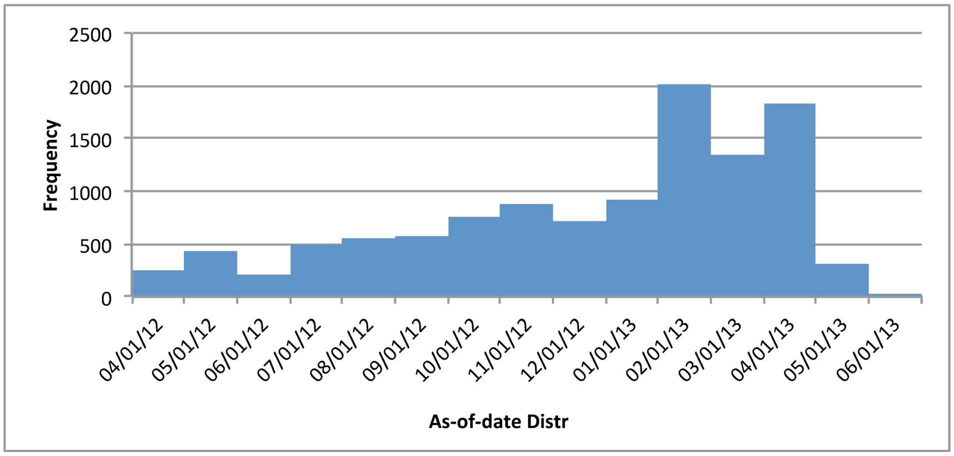

Figure 1 **Sales Report As-of-date Distribution**

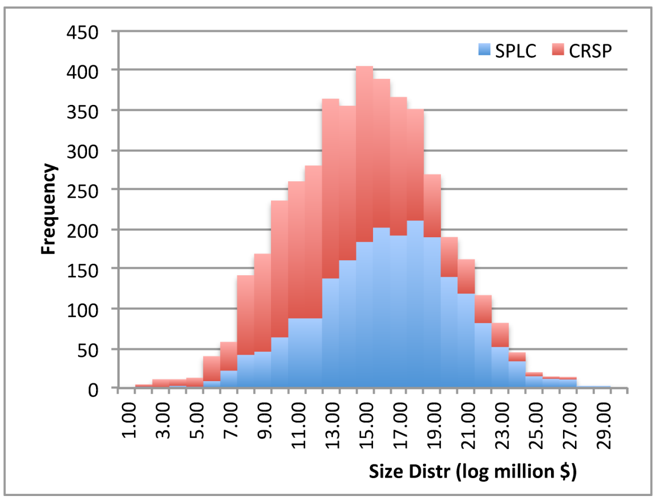

Figure 2 **Firm Log-size Distribution**

distribution is shown in Figure 2. We use red for firms in CRSP and blue on top of the red for the SPLC firms. Both the SPLC data and CRSP universe seem to have approximately lognormal size distributions. The firm size distribution of SPLC is, however, clearly biased towards larger firms, which intuitively makes sense. This suggests that the supply chain relationships involving large firms are easier to capture than those involving only small firms. Firms, especially small ones, also have incentive to not disclose, or even hide their supply chain relationships for competition concerns, as discussed in Ellis et al. 2012. Given this observation, we would anticipate that small firms would exhibit more bias from intentional concealment or voluntary disclosure than large firms and that SPLC's greater large firm representation reduces this bias.

Supplier relationships may also have different importance for firms in different industries. A car manufacturer relies on its supply chain partners heavily to produce cars just-in-time, while a bank may still be able to operate properly if the ordered office laptops are delayed. Therefore,

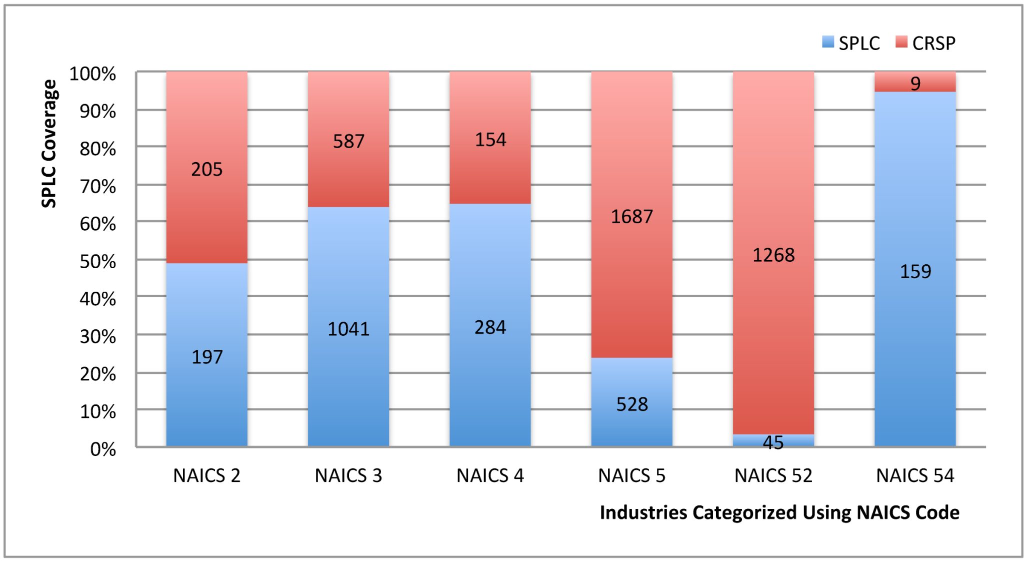

Figure 3 **Firm Coverage of Industry Breakdown**

it is important to see the coverage bias in terms of industry breakdown. In Figure 3, we plot the total firms captured in our data according to the first digit of the NAICS code and compare these numbers to the total firms in CRSP. We use blue to indicate the number of firms in our data and red to indicate the the number of firms not captured. The first bar represents industries starting with Code 2, including mining, utilities, and construction, of which we can see that 197 out of 402 firms in this large sector are captured by SPLC, a coverage ratio of approximately 50%. The second bar represents industries starting with Code 3, i.e., manufacturing, and the third bar represents industries starting with Code 4, i.e. the logistics sector which includes wholesale, retail, warehousing, and transportation. Our data have about 65% coverage for both manufacturing and logistics. This coverage ratio is consistent even if we further break down these categories using the first two digits of the NAICS code. The fourth bar represents industries starting with Code 5, i.e., various service industries. While overall coverage in this grand service sector is almost one quarter, the coverage ratios vary dramatically within groups selected. For example, the fifth bar shows that our data only covers 3.4% of firms in finance and insurance (NAICS 52), compared to coverage of 94.6% of firms in professional, science, and technology (NAICS 54) as shown in the last bar. Overall, the manufacturing and logistics sectors have the most consistent cross-sectional firm coverage in our data.

We further investigate the distributions of the captured relationships. Figure 4 shows the histograms of in-degree and out-degree per firm, which seem to follow a power law distribution. Characterizing the exact degree distribution is beyond the scope of this paper, but we note that other research, such as that of Atalay et al. 2011, argues that the power law distribution may over-predict the number of minimally connected firms. It is also worth mentioning that not all firms

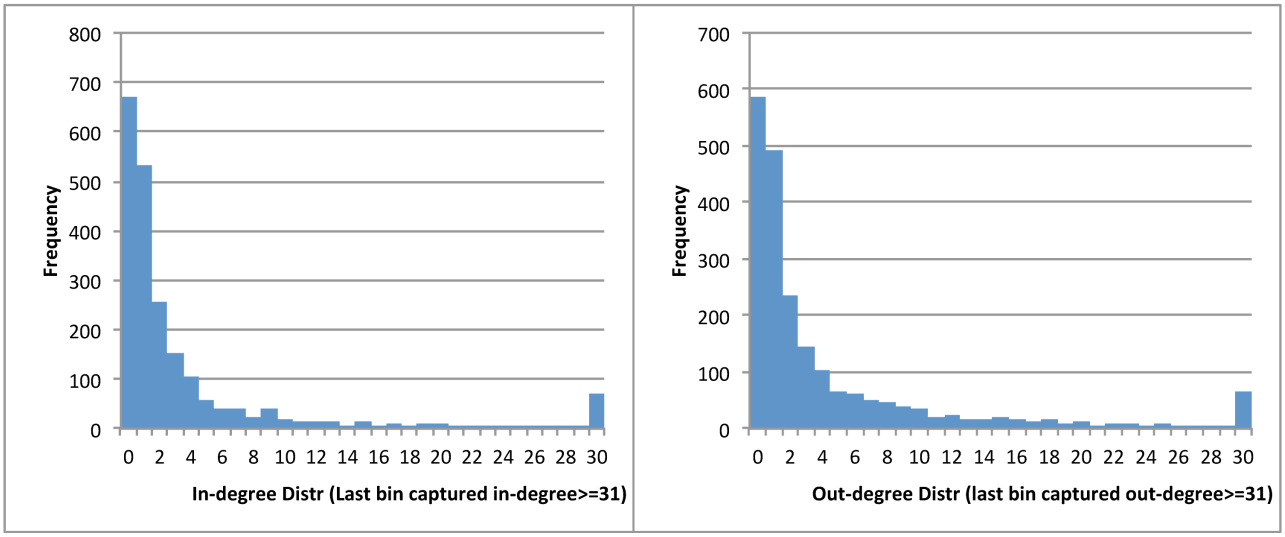

Figure 4 **In-degree and out-degree distribution**

Notes. The last bars in both distributions represent the number of firms that have no less than 30 in-degree (or out-degree) relationships. Descriptors of the data in this figure, mean, median, and power law coefficients, are given in Table 1. For reference, firms with large degree are listed in Table 2 as "Top 10 most connected firms."

have both supplier and customer relationships captured in our data; 670 firms do not have supplier information, while 587 firms do not have customer information. We need special treatment for these firms, as discussed in the next section.

Since the Compustat dataset captures sales that are more than 10% of suppliers' revenue, we consider the extent of the sales below the 10% threshold in our data. Figure 5 shows the distribution of sales contribution percentages, which are the ratios of captured sales quantity to the total revenue made by the supplier firm. The left figure shows the distribution of the 865 relationships above the 10% threshold; the right figure shows that of the 10,954 relationships below the 10% mark. We note that the sales contribution here also seems to follow a power law distribution.

Table 1 shows summary statistics of our data. In Panel A, we report firm coverage. Among the 2,152 firms in our dataset, 1,576 firms function as suppliers to other firms, while 1,496 firms function as customers to other firms. The total market capitalization of the firms in our dataset is about 14.2 trillion dollars. For comparison to the CRSP universe, CRSP has 5,090 firms in our chosen time window and a total market capitalization of about 19.3 trillion dollars according to 2012 annual fundamentals. Thus, our dataset covers 42.3% of the total number of publicly listed firms in the U.S. market and 75.0% of the total market capitalization. The fact that SPLC has a larger coverage over the market cap than the number of firms indicates again that SPLC is tilted toward large cap firms, which can also be seen from the mean and median firm sizes. The average firm size in SPLC is 6,740 million dollars, compared to the average size in CRSP of 4,447 million

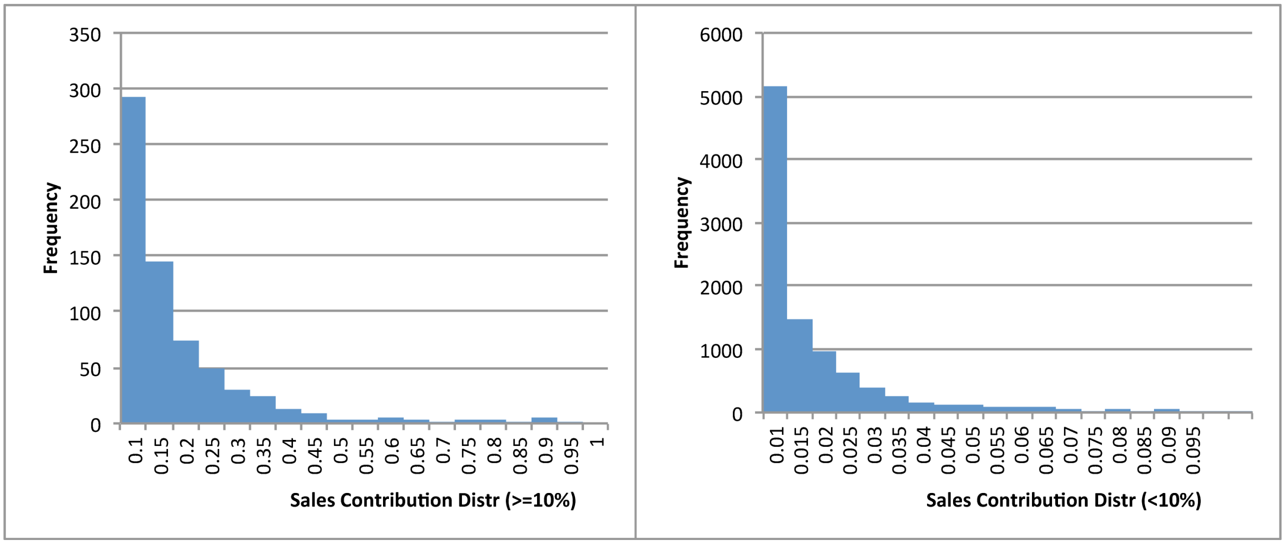

Figure 5 Sales Contribution Distribution

Notes. This figure shows the sales contribution of all relationships captured in our data. The sales contribution is the ratio of captured sales quantity to the total revenue of the supplier firm.

dollars. The median in SPLC is 1,112 million dollars, compared to the median in CRSP of 550 million dollars. Overall, we conclude that SPLC covers a significant portion of public firms in the U.S. economy.

In Panel B we report summary statistics on the link information. The mean of supplier / customer per firm is 5.16, while the median is only 1, indicating a sparse network in general, in which, many firms are actually on supply chain paths instead of networks. We estimate the degree distribution using the maximum likelihood method described in Clauset et al. 2009, and find coefficients of 1.88 for out-degree customer and 2.76 for in-degree supplier; therefore, the out-degree customer distribution has a heavier tail than the in-degree supplier distribution. Since smaller sales relationships are more likely to be missing compared to larger sales, the true degree distributions should have even heavier tails and our coefficient estimates should be overestimated relative to the actual power law coefficients.

For every firm, we also compute the ratio of the total captured sales to total revenue. We find on average, a firm only has 16.09% of its total sales identified in our data. If we use revenue-weighted averages and consider the whole economy, we find an even lower ratio of 11.01%. This means that in aggregate, a large portion of sales relationships are still missing, which has an implication for the centrality measure we use in the next section. Overall, we believe that our data may compute a relatively realistic order in terms of a first-order centrality measure for firms, such as eigenvector centrality and degree centrality, but may be biased for higher order centrality measures such as supplier concentration or customer concentration.

We argue that a significant part of missing sales are due to the omission of private firms, household consumers, and government, as well as foreign sectors, which may be significant suppliers or customers for many firms as in the examples below.

1. Lockheed Martin Corporation has 9.67 billion dollars in sales to the public sector, i.e., the U.S. government, which is 82.0% of its 2012 annual revenue.

2. Intel Corporation sold 1.41 billion dollars, 11% of Intel's 2012 annual revenue, to Lenovo Group Ltd., a Chinese firm and the 2nd largest personal computer manufacturer in the world.

3. Best Buy Company purchased 1.33 billion dollars, 10.41% of Best Buy's COGS in 2012, from Samsung Electronics, a Korean firm.

Table 1 Summary Statistics

	SPLC	CRSP	% Coverage of CRSP
Panel A: Firms			
Number of all firms	2,152	5,090	42.3
Number of supplier firms	1,576	-	31.0
Number of customer firms	1,496	-	29.4
Market value of all firms (million $)	14,229,214.35	18,983,256.21	75.0
Market value of suppliers (million $)	11,622,294.74	-	61.2
Market value of customers (million $)	13,085,195.03	-	68.9
Mean size of all firms (million $)	6,740.00	4,497.34	-
Mean size of suppliers (million $)	7,498.25	-	-
Mean size of customers (million $)	8,901.49	-	-
Median size of all firms (million $)	1,112.18	577.01	-
Median size of suppliers (million $)	1,048.68	-	-
Median size of customers (million $)	1,827.66	-	-
Panel B: links			
Number of links captured	11,819	-	-
Number of sales contribution $\geq 10\%$	865	-	-
Number of sales contribution $< 10\%$	10,954	-	-
Mean supplier / customer per firm	5.16	-	-
Median supplier / customer per firm	1	-	-
Out-degree power-law coefficient r	1.88†	-	-
In-degree power-law coefficient r	2.76†	-	-
% Equal weighted sales captured	16.09	-	-
% Revenue weighted sales captured	11.01	-	-

† Power law coefficients are fit to the function $N(k) = k^{-r}$ (meaning the probability for a node to have no smaller than k degrees) by maximum likelihood using the goodness-of-fit based method described in Clauset et al. 2009.

Notes. The SPLC column lists cross-sectional observations as of June 2, 2013. The CRSP column provides cross-sectional observations of 2012 annual fundamentals. The percent coverage is the number of stocks with a valid supplier-customer link in SPLC divided by the total number of CRSP stocks. The market cap percent coverage is the total market capitalization of stocks with a valid supplier-customer link in SPLC divided by the total market value of the CRSP stock universe. Size is the firm's market value of equity.

4. Cargill, a privately held firm, had 133.9 billion dollars in sales in 2012. Its customer base includes retail giants such as Wal-Mart and Target, although the exact quantities in these relationships are unknown.

Overall, we conclude that the firms covered in the SPLC account for a major part of the U.S. economy. The basic distribution patterns discussed suggest the measures of supply chain network captured by our data are meaningful. Since our main interest is to observe the effects of firm centrality and systematic risk, we believe that missing end-customer nodes, such as government and household consumers, and less-connected segments, such as foreign sectors, would have relatively little influence on risk propagation. We also believe that the omission of private firms, few of which would appear among the largest firms in the most heavily covered NAICS segments, also introduces little bias to the measured centrality-risk relationships.

To further show the economic network in the data, we plot the cross-sectional supply chain network of the 2,152 firms in Figure 6 using a force-directed layout algorithm proposed in Fruchterman and Reingold 1991. In this algorithm, spring-like attractive forces based on Hooke's law are used to attract pairs of endpoints of the graph's edges towards each other, while simultaneously repulsive forces like those of electrically charged particles based on Coulomb's law are used to separate all pairs of nodes. In equilibrium states for this system, the edges tend to have uniform length, and nodes that are not connected by an edge tend to be drawn further apart. As a result, well-connected nodes tend to be placed in more central positions while less-connected nodes are placed at the periphery. This is useful to show companies with different positions in the supply chain network.

We consider two firms, Apple and CVS, which are both highlighted in red in Figure 6. Apple has a total of 135 relationships (30 out-degrees, 105 in-degrees, which ranks 11 in terms of total degree in the dataset) while CVS has a total of 127 relationships (10 out-degrees, 117 in-degrees, which ranks 12 in terms of total degree) captured by SPLC. Since both firms have many links in our data, the nodes representing Apple and CVS both tend to be placed near the center of the network. However, CVS connects to more peripheral firms than Apple, which can be seen from the length of the links. As a result, Apple has a eigenvector centrality of 6.784×10^{-3}, much higher than CVS's eigenvector centrality of 2.028×10^{-3}.

Table 2 shows the 10 most connected firms in the SPLC data. Wal-Mart is the most connected public firm in the US economy, but it does not have a single customer firm captured in our data since it sells primarily to household consumers. IBM is the second most connected firm in the US economy and is the fourth most connected firm in terms of both in-degree and out-degree. This level of centrality for IBM stems from its position in supplying business information solutions, which require inputs from upstream semiconductor and device firms, and sales to various business

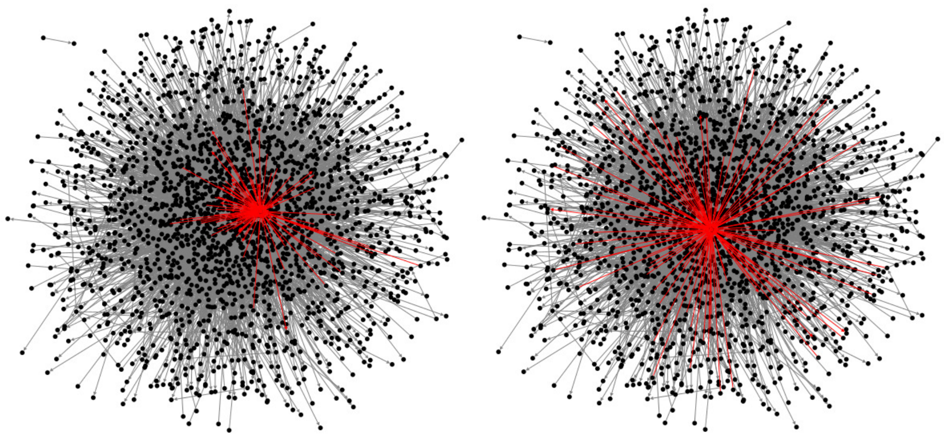

Figure 6 Supply Chain Network Captured by SPLC (Left: Apple, Right: CVS).

Notes. These diagrams depict the Bloomberg SPLC dataset as of June 2, 2013 using the Fruchterman-Reingold layout algorithm. In equilibrium states for this system, the edges tend to have uniform length; nodes that are not connected by an edge tend to be drawn further apart. Apple's links are colored red on the left, while CVS's links are colored red on the right.

customers. For example, IBM's top US supplier is Intel Corporation, which sold 242.9 million dollars of goods to IBM in 2012, including Intel's Xeon® CPU, a major input for IBM's business server products, which are then sold to many downstream business customers such as the US Postal Service, Verizon Communications, and AT&T Inc.

We can also observe from Table 2 that most of the top connected firms belong to manufacturing (NAICS code 31-33) and the logistics (NAICS code 42-49) industries. The degree data availability for those two large industry sectors agrees with the industry coverage result as shown in Figure 4.

4. Empirical Results

In this section, we use the SPLC dataset to test both our first-order (direct) effect hypotheses and our second-order (indirect) effect hypotheses.

Table 2 Top 10 most connected firms.

Rank	in-degree	k	out-degree	k	Total degree	k
1	Wal-Mart	249	Oracle	110	Wal-Mart Stores	249
2	Target	152	VMware	107	IBM	228
3	Hewlett-Packard	150	Microsoft	83	Hewlett-Packard	214
4	IBM	145	IBM	83	Cisco Systems	201
5	Lockheed Martin	140	Kansas City Southern	76	Microsoft	177
6	Boeing	138	Rackspace Hosting	74	Dell	171
7	Cisco Systems	132	Salesforce.com	74	Boeing	156
8	Dell	127	Manhattan Associates	74	Target	152
9	Costco Wholesale	126	Citrix Systems	72	Lockheed Martin	147
10	CVS Caremark	117	Cisco Systems	69	Oracle	139

4.1. First-order Effects: Panel Data Regression

We first run a pooled OLS regression of the network model of returns for the full panel data, using variants of the regression (1). Our primary interest is the explanatory or predictive power of $\sum_j w_{ij}^{in} r_{j,t-1}$, $\sum_j w_{ij}^{out} r_{j,t-1}$, $\sum_j w_{ij}^{in} r_{j,t}$ and $\sum_j w_{ij}^{out} r_{j,t}$. To be considered in the following tests, a firm must meet a minimum liquidity threshold of $5 share price in the chosen horizon. This ensures that portfolio returns are not driven by micro-capitalization effects for illiquid securities. This also helps to avoid delisting (which generally occurs when stock prices fall below one dollar) and infrequent trading issues that can lead to stale pricing effects such as inflated serial correlation.

Table 3 summarizes the results. Observing each column, we see that the effects of the concurrent supplier and customer returns are significant for both univariate and multivariate regressions, supporting Hypothesis 1. In the first row, the concurrent supplier returns have a coefficient of 0.370, close to the concurrent customer returns coefficient of 0.387. In the univariate cases, the coefficients are respectively 0.517 and 0.587. The magnitudes of these coefficients show that our data provide economically meaningful supplier chain relationships.

We next investigate lagged effects, i.e., one-month lagged responses to own, supplier, and customer shocks. For all cases, the one-month own lagged effect is significant with slightly negative coefficients, meaning high past own returns predict low future own returns. As we noted above, this effect also appears in Fama and French 1988, Jegadeesh 1990, and other studies without the presence of supplier and customer returns terms. For the cross-firm lagged effect, we find that in all cases the supplier lagged effect is statistically significant, but that the customer lagged effect is not significant. This supports Hypothesis 2 and 3(B). The supplier lagged effect has a statistically significant coefficient of 0.025 when current-period connections are also included. Comparing the first row with the second row, the supplier lagged effect has a higher coefficient of 0.044 when we omit the contemporaneous effects. The rows with at least one concurrent variable all have an adjusted R^2 greater than 13%, while the cases with no concurrent variable have an adjusted R^2 less than 0.2%. This shows that variations in the dependent variable are mostly explained by concurrent cross-firm returns.

Overall, the panel data regression results suggest that both customers and suppliers have significant concurrent effects, of which the first is slightly stronger than the second, but only suppliers have a significant one-month lagged effect. The cross-lagged effect results have two important implications for the time window we choose. First, from the financial market perspective, investors may be subject to limited attention to suppliers as opposed to customers. Another reason could be that firms are more reluctant to disclose supplier information than their customer information; thus, supplier information is more difficult to obtain for investors. Second, from an operations management perspective, the gradual diffusion of information in the downstream direction may indicate lack of downstream supply chain coordination, i.e., supplier firms withholding proprietary operational news from downstream firms. The asymmetric information may be attributed to different market power that upstream players and downstream players possess in the supply chains. Another possible reason for the gradual downstream information diffusion is that customer firms may order less, foreseeing a demand shock; thus, supplier firms would show a decrease in sales before the customer firms due to the delivery lead time. Overall, the cross-lagged effect results can be explained by a combination of supply chain operations and investors' insufficient perception of supplier information.

We construct similar tests with different horizons, finding that the significance drops as the horizon increases, with the 2-month trailing returns coefficients being significantly weaker than the one month signal, and the 3-month trailing returns coefficients being insignificant. In the following we focus on one month lagged returns to avoid biasing the t-tests with overlapping forward periods.

4.1.1. Fama-MacBeth Regression Pooled OLS results may be biased since the residuals may not be independently and identically distributed. Since the residuals of a given year may be

Table 3 Pooled OLS of Concurrent and Lagged Returns.

	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$	Adj.R^2(%)
Coef	0.000	-0.028***	0.025**	0.007	0.370***	0.387***	19.04
(T-Stat)	(0.49)	(-3.78)	(2.42)	(0.56)	(36.48)	(31.98)	
Coef	0.008***	-0.036***	0.044***	0.010			0.13
(T-Stat)	(9.10)	(-4.30)	(3.80)	(0.72)			
Coef	0.008***	-0.021***					0.03
(T-Stat)	(9.45)	(-2.81)					
Coef	0.008***		0.030***				0.04
(T-Stat)	(9.14)		(2.96)				
Coef	0.008***			0.013			0.00
(T-Stat)	(9.11)			(1.08)			
Coef	0.004***				0.517***		14.45
(T-Stat)	(5.16)				(55.65)		
Coef	0.001					0.587***	13.09
(T-Stat)	(0.91)					(52.56)	
Coef	0.004***		0.022**		0.517***		14.47
(T-Stat)	(5.02)		(2.32)		(55.61)		
Coef	0.001			0.002		0.587***	13.09
(T-Stat)	(0.88)			(0.21)		(52.54)	
Coef	0.000				0.370***	0.387***	18.98
(T-Stat)	(0.54)				(36.52)	(32.03)	

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the pooled OLS result of the regression (1) using concurrent supplier and customer returns, supplier and customer lagged effect, as well as firm's own lagged effect. The result shows that the concurrent supplier returns, the concurrent customer returns, the own lagged effect and the supplier lagged effect are significant in explaining firm returns, but not the customer lagged effect. The result is consistent for both univariate and multivariate regressions.

correlated across firms, we use Fama-MacBeth regression as in Fama and MacBeth 1973 to deal with the time effect. As discussed in Petersen 2009, the Fama-MacBeth method estimates the loadings on risk factors in two steps to avoid problems of correlation across contemporaneous residuals in panel data. The first step runs T cross sectional regressions to obtain estimated coefficients for assets while the second step uses the coefficient estimates to find the loading estimates. A detailed discussion of the Fama-MacBeth method is provided in the Appendix. We also assume the correlation of firm residuals in different years is weak and proceed with the Fama-MacBeth regression as follows[8].

Each month in our time window has its own set of monthly regression coefficients. We calculate the average coefficient for each signal across months and then calculate the t-statistic to test whether each coefficient is statistically different from 0. The results are shown in Table 4.

Similar to Table 3, in both the univariate and multivariate cases, the coefficients for concurrent supplier and customer returns and the supplier lagged effect are significant, but the coefficients for the customer lagged effect are not significant. Own momentum is significant with slightly negative coefficients, again consistent in all cases. Comparing Table 4 to Table 3, the concurrent customer has a much larger impact than the concurrent supplier. In the first row, the downstream coefficient of 0.755 is almost twice as large as the upstream coefficient of 0.399. Our results then suggest that investors should pay more attention to a firm's customers than to its suppliers for the contemporaneous effect but should mainly care about its suppliers for cross-firm lead-lag effects.

4.1.2. Robustness test To further explore the robustness of our results, we want to see whether we still observe the same results after controlling for the common factors of market premium, size, value, and momentum. We add these common factors to form the following regression (2):

$$
\begin{aligned}
r_{i,t} = {} & \alpha + \beta_1 r_{i,t-1} + \beta_2 \sum_j w_{ij}^{in} r_{j,t-1} + \beta_3 \sum_j w_{ij}^{out} r_{j,t-1} + \beta_4 \sum_j w_{ij}^{in} r_{j,t} + \beta_5 \sum_j w_{ij}^{out} r_{j,t} \\
& + b_i \left(R_{mt} - R_{ft}\right) + s_i SMB_t + h_i HML_t + u_i MOM_t + \epsilon_{it},
\end{aligned} \tag{4}
$$

where SMB stands for "Small (market capitalization) Minus Big", HML stands for "High (book-to-market ratio) Minus Low", and MOM stands for "Momentum" of average returns on the two high prior returns portfolios minus the average returns on the two low prior return portfolios. Those factors measure the stock's exposure to small caps over big caps, value stocks over growth

[8] The average firm return auto-correlation is -0.020, and the average firm return residual auto-correlation is -0.011.

stocks, and winner stocks over loser stocks. All the factors are defined by self-financing portfolios. The factor data is readily available from the Kenneth French data library[9].

Table 5 summarizes the results[10]. We see similar qualitative results to those in Tables 3 and 4, i.e., both current suppliers and customers explain a firm's return, while customers are more important than suppliers for the current period effect; for the lagged effects, both the supplier one-month lagged effect and the own lagged effect are significant. The customer lagged effect is only slightly significant for the univariate case. Comparing the columns of Table 5 to the corresponding columns of Table 4, the coefficients for concurrent supplier and customer returns are smaller. The weaker sensitivities for current cross-firm returns are due to the fact that some concurrent cross-firm effects are explained by the common factors.

[9] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[10] A complete table with loadings on the common factors is in Appendix.

Table 4 Fama-MacBeth Regression of Concurrent Returns and Momentum.

	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$
Ave. Coef	-0.001	-0.088***	0.036**	0.024	0.399***	0.755***
(T-Stat)	(-0.96)	(-11.06)	(2.17)	(0.95)	(20.90)	(3.12)
Ave. Coef	0.009***	-0.090***	0.057***	0.004		
(T-Stat)	(10.38)	(-9.08)	(2.96)	(0.09)		
Ave. Coef	0.009***	-0.047***				
(T-Stat)	(10.53)	(-6.96)				
Ave. Coef	0.008***		0.022**			
(T-Stat)	(11.09)		(1.83)			
Ave. Coef	0.008***			-0.040		
(T-Stat)	(10.92)			(-0.66)		
Ave. Coef	0.003***				0.619***	
(T-Stat)	(3.61)				(37.25)	
Ave. Coef	-0.002**					0.992***
(T-Stat)	(-2.26)					(4.54)
Ave. Coef	0.004***		0.018*		0.625***	
(T-Stat)	(4.51)		(1.57)		(36.44)	
Ave. Coef	-0.002*			0.001		1.001***
(T-Stat)	(-1.92)			(0.0274)		(4.51)
Ave. Coef	-0.001*				0.393***	0.744***
(T-Stat)	(-1.80)				(22.48)	(3.20)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the Fama-MacBeth results of the regression (1) using concurrent returns and one-month momentum as independent variables. We have the same result as the pooled OLS after controlling for the time effect, i.e., the concurrent supplier returns, the concurrent customer returns, the own momentum and the supplier lagged effect are significant in explaining firm returns. The result is consistent for both univariate and multivariate regressions.

Since Menzly and Ozbas 2010 find strong upstream and downstream industry effects, we further examine whether the supplier and customer firm effects are different from the upstream and downstream industry effects. After replacing the firm returns on the right hand side of the regression (1) by the industry returns that the firm resides in based on the first 3 digits of the NAICS codes, we observe smaller coefficients and the respective t-stats of the cross-sectional regression in an unreported table, while the signs are the same. The adjusted R-square also reduces significantly by 6.2% to 12.8% compared to that in Table 3. This robustness check indicates that the supplier and customer firm effects explain returns better than the supplier and customer industry effects. We also find positive own lagged effect for the 1-month returns as in Grinblatt and Moskowitz (1999).

Although results such as the supplier lagged effect are consistent with the investor's limited attention hypothesis, there are a number of other plausible explanations of the data. We next present results for a series of robustness tests for investor inattention.

A number of papers find that larger firms, or firms with higher levels of analyst coverage, institutional ownership, and trading volume, lead smaller firms

Table 5 Fama-MacBeth Regression after Controlling for Common Factors.

	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$
Ave. Coef	-0.000	-0.086***	0.063***	0.010	0.111***	0.503*
(T-Stat)	(-0.45)	(-9.16)	(3.42)	(0.23)	(4.28)	(1.78)
Ave. Coef	-0.001	-0.091***	0.050***	0.029		
(T-Stat)	(-1.09)	(-10.43)	(3.02)	(0.70)		
Ave. Coef	-0.002*	-0.054***				
(T-Stat)	(-1.80)	(-7.93)				
Ave. Coef	-0.001		0.029**			
(T-Stat)	(-1.60)		(2.29)			
Ave. Coef	-0.002**			0.034*		
(T-Stat)	(-2.50)			(2.05)		
Ave. Coef	-0.001**				0.126***	
(T-Stat)	(-1.75)				(6.24)	
Ave. Coef	-0.002***					0.501*
(T-Stat)	(-2.83)					(1.69)
Ave. Coef	-0.001		0.029**		0.130***	
(T-Stat)	(-0.886)		(2.14)		(5.93)	
Ave. Coef	-0.003***			0.041		0.492**
(T-Stat)	(-2.91)			(1.66)		(2.19)
Ave. Coef	-0.002***				0.114***	0.485*
(T-Stat)	(-2.16)				(5.45)	(1.79)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the Fama-MacBeth results after controlling for common asset pricing factors. We have similar results to those in Table 4. The results are consistent for both univariate and multivariate cases. All factors are defined by self-financing portfolio. Factor data is from the Kenneth French data library.

or firms with lower levels of analyst coverage, institutional ownership, and trading volume (e.g., Lo and MacKinlay 1990, Brennan el at. 1993 Badrinath el at. 1995, Chordia and Swaminathan 2000, Hou and Moskowitz 2005, Hou 2006). The supplier lag effect results could be caused by firms of different size, analyst coverage, institutional ownership, and trading volume. To ensure that our results are not driven by those alternative explanations, we conduct the following robustness tests by constructing filters. For checking the firm size effect for example, we only pick the firms that have their market capitalization larger than the input supplier weighted firms' market capitalization. In other words, the firms we pick are all larger firms compared to their average supplier firms weighted by their purchase orders. Since smaller supplier firms are less noticeable to investors, then, if we still see a significant supplier lagged effect, this should not be due to larger sizes of upstream firms. We apply similar filters using levels of analyst coverage, institutional ownership, and trading volume. A detailed description of the alternative explanation robustness tests is provided in the Appendix. A brief result is given in Table 6. The supplier lagged effect is still significant after different filters, which means those possible alternative explanations cannot alone explain the supplier lagged effect. Different from Menzly and Ozbas 2010, we find negligible changes after removing the top analyst coverage stocks, which implies that analyst coverage does not explain diffusion of information about a firms supply chain connections. A detailed description of the test for alternative explanations is provided in the Appendix.

The supplier lead-lag effect is also documented in the appendix of Cohen and Frazzini 2008. They use 30 years of Compustat data from 1980 to 2004, which is not overlapped with our data and thus serves as an out-of-sample test to our finding.

4.1.3. Backtest Since the above results suggest that the one-month supplier lagged effect has predictive power for firms' returns, we perform a backtest using a value-weighted portfolio based on the following supplier prediction strategy. Specifically, every month we rank the firms according to their one-month supplier lagged effect and assign them to one of five quintile portfolios. All stocks are value weighted within a given portfolio and rebalanced every month. The strategy is to build a zero-cost portfolio that longs the top quintile supplier lagged effect stocks and shorts the bottom quintile supplier lagged effect stocks. This investment rule should earn zero abnormal returns in an efficient market.

We present the result in Table 7. We compute the abnormal returns using several definitions, including the excess returns above market, and the alpha in factor models. The rightmost column shows that this strategy delivers an excess return of 0.62% per month. After controlling for the four factors, it delivers an abnormal return of 0.56% per month or approximately 6.7% per year. This result is highly statistically significant, suggesting the economic magnitude of the supplier lagged effect is large.

4.1.4. Industry Breakdown After examining the whole market, we wish to test whether similar results can be found for each industry sector since different industries may have different sensitivities to supplier and customer concurrent returns and lagged effects. Using the NAICS

Table 6 Fama-MacBeth Regression of Alternative Explanations.

Filter Criteria	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$
Market	0.065***	-0.091***	0.070**	0.025	0.391***	0.370***
Capitalization	(6.29)	(-5.99)	(2.77)	(0.88)	(15.18)	(12.92)
	0.014***	-0.103***	0.105***	0.045		
	(14.87)	(-5.21)	(3.14)	(1.27)		
Institution	0.002*	-0.090***	0.084***	0.027	0.414***	0.566***
Ownership	(1.77)	(-6.84)	(3.45)	(0.79)	(13.14)	(14.07)
	0.013***	-0.101***	0.119***	-0.003		
	(10.09)	(-5.71)	(3.89)	(-0.08)		
Analyst	-0.000	-0.081***	0.047**	-0.007	0.377***	1.024*
Coverage	(-0.11)	(-6.49)	(2.25)	(-0.17)	(16.15)	(1.89)
	0.008***	-0.077***	0.071***	-0.067		
	(7.94)	(-4.95)	(2.90)	(-0.75)		
Trading	0.000	-0.081***	0.054**	0.060	0.429**	0.887***
Volume	(0.30)	(-7.94)	(2.53)	(0.17)	(17.39)	(2.41)
	0.010***	-0.081***	0.087***	-0.045		
	(8.56)	(-6.32)	(3.45)	(-0.69)		

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the Fama-MacBeth robustness test results after filtering market capitalization (ME), i.e. $ME_i > \sum_j w_{ij}^{in} ME_j$, institution ownership, i.e. $\left(\frac{InstitutionOwnedShares}{TotalShareOutstanding}\right)_i > \sum_j w_{ij}^{in} \left(\frac{InstitutionOwnedShares}{TotalShareOutstanding}\right)_j$, analyst forecast, i.e. $AnalystForecastCount_i > \sum_j w_{ij}^{in} AnalystForecastCount_j$, and trading volume, i.e. $\left(\frac{TradingVolume}{TotalShareOutstanding}\right)_i > \sum_j w_{ij}^{in} \left(\frac{TradingVolume}{TotalShareOutstanding}\right)_j$. Institution ownership data is from Thomson-Reuters Institutional Holdings (13F) Database. Analyst coverage data is from the IBES dataset. Share trading volume data comes from the CRSP dataset.

Table 7 Supplier Prediction Strategy, Abnormal Returns. (%)

	1(Low)	2	3	4	5(High)	L/S
Excess returns	-0.09	-0.09	0.37	0.47	0.53	0.62*
(T-Stat)	(-0.05)	(-0.06)	(0.27)	(0.42)	(0.40)	(1.69)
CAPM	-0.45	-0.49	-0.28	0.06	0.16	0.61**
(T-Stat)	(-1.06)	(-1.48)	(-0.89)	(0.66)	(0.33)	(2.25)
Three-factor	-0.39	-0.40	-0.28	0.06	0.16	0.55**
(T-Stat)	(-0.86)	(-1.05)	(-1.52)	(0.14)	(0.44)	(2.10)
Four-factor	-0.41	-0.42	-0.30	0.12	0.15	0.56**
(T-Stat)	(-0.91)	(-1.06)	(-1.59)	(0.29)	(0.39)	(2.09)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the backtest result using the supplier prediction strategy. The zero-cost portfolio constructed by holding the top quintile and selling short the bottom quintile yields significant abnormal returns as is shown in the rightmost column. Every quintile portfolio has 352 firms.

codes, we conduct Fama-MacBeth regressions for each industry. We use the first two digits of the NAICS code to identify large sectors to strike a balance between fine-grained industry classifications and statistical reliability. Note that the NAICS codes for a few firms tend to change over our chosen monthly window. For example, Cameron International Corporation (ticker CAM), a firm that provides flow equipment products, systems, and services worldwide, has its NAICS code as 332912 (Fluid Power Valve and Hose Fitting Manufacturing) in June 2011, while the code changes to 423830 (Industrial Machinery and Equipment Merchant Wholesalers) in July 2011. For sake of simplicity, we use the NAICS code as of December 31, 2012.

The results appear in Table 8. The number of firms in each industry is recorded in the second column. Note that there are 2,139 firms in total in Table 8, fewer than the 2,152 firms in our data set. This is because some industry sectors, such as Education Services (NAICS code 61) and Public Administration (NAICS code 92), have fewer than 5 firm observations captured in our data and are thus excluded from consideration in this study. A few firms also have an industry code listed as "Non-Classified" and are thus omitted. From the results, we can assign most industries to one of two groups: Group 1 (those with concurrent relationship effects and supplier lagged effects) and Group 2 (those without supplier lagged effects).

Group 1 has the same results as in Tables 3-5, that concurrent upstream and downstream returns as well as the supplier lagged effect are significant. Group 1 includes Agriculture & Forestry (NAICS code 11), Manufacturing (NAICS code 31-33), Transportation & Warehousing (NAICS code 48-49), Information (NAICS code 51), and Health Care (NAICS code 62). This group has a total of 1,413 firms which is 66% of the sample size and more than 60% of the U.S. economy. Thus, Group 1 drives our result for the economy-wide observations in Tables 3-5. We note that Manufacturing (NAICS code 31-33) in this group also has a weakly significant customer lagged effect.

Group 2 only has concurrent effects and no lagged effects. Group 2 includes Mining (NAICS code 21), Construction (NAICS code 23), Wholesale (NAICS code 42) and Retail (NAICS code 44-45), Finance & Insurance (NAICS code 52), Real Estate & Leasing (NAICS code 53), Professional & Science (NAICS code 54), and Arts & Entertainment (NAICS code 71).

Other sectors exhibit unique behavior. Specifically, Utilities (NAICS code 22) are sensitive to suppliers' concurrent performance and one-month lagged effect, but not to customer effects. One possible reason may be that they are sensitive to the prices of their input materials such as oil and gas, but the downstream demand is relatively stable since their customer base is well diversified. Support, Waste & Remediation (56) seems to only have statistically significant relations with downstream customers and not with suppliers. This may be due to the fact that their market performance is mainly determined by the quantity of services purchased by downstream firms. Accommodation & Food (NAICS code 72) is only sensitive to its concurrent supplier performance,

Table 8 Fama-MacBeth Regression of Industry Breakdowns

NAICS	# Firms	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$
11	6	0.021***	-0.209**	0.297*	-1.048	0.412***	0.204**
Agriculture & Forestry		(3.38)	(-2.74)	(1.47)	(-1.52)	(72.15)	(2.49)
21	127	-0.011	-0.136***	0.035	0.030	0.597***	0.516***
Mining		(-2.18)	(-3.34)	(0.32)	(0.30)	(7.13)	(3.90)
22	52	0.013***	-0.163**	0.184**	-0.044	0.199	0.065
Utilities		(4.07)	(-2.63)	(1.85)	(-0.91)	(1.50)	(0.52)
23	18	-0.008	-0.093	0.101	0.100	0.747***	0.642**
Construction		(-0.61)	(-1.00)	(0.36)	(1.00)	(3.81)	(6.29)
31-33	1041	-0.003	-0.085***	0.024*	0.044*	0.376***	0.969**
Manufacturing		(-2.31)	(-8.04)	(1.73)	(1.68)	(17.05)	(2.26)
42	59	0.012***	-0.096**	-0.057	0.042	0.698***	0.230**
Wholesale Trade		(3.33)	(-2.03)	(-0.40)	(0.36)	(4.41)	(2.04)
44-45	146	0.006	-0.116***	-0.040	0.046	0.493***	0.392***
Retail Trade		(0.15)	(-4.19)	(0.38)	(0.58)	(4.48)	(4.87)
48-49	79	-0.002	-0.040	0.100*	-0.079	0.438***	0.513***
Transportation & Warehousing		(-0.28)	(-0.80)	(1.72)	(-0.69)	(4.77)	(3.02)
51	245	0.001	-0.076***	0.087**	-0.079	0.319***	0.544***
Information		(0.54)	(-3.22)	(2.09)	(-1.62)	(6.69)	(8.16)
52	45	0.013	-0.069	-0.142	0.063	0.361**	0.503**
Finance & Insurance		(1.27)	(-0.59)	(-0.65)	(0.76)	(2.49)	(1.99)
53	33	0.012	-0.014	0.060	-0.012	0.516***	0.694***
Real Estate, Rental & Leasing		(2.51)	(-0.19)	(0.47)	(-0.09)	(3.43)	(3.43)
54	159	-0.001	-0.099***	0.050	0.071	0.422***	0.478***
Professional, Science & Tech		(-0.24)	(-3.56)	(0.78)	(0.79)	(4.94)	(5.72)
56	46	0.006	-0.116**	-0.096	0.262**	0.031	0.607***
Support, Waste & Remediation		(0.88)	(-2.09)	(-0.84)	(1.90)	(0.12)	(3.22)
62	39	-0.012	-0.181***	0.271*	-0.005	0.884***	0.930***
Health Care & Assistance		(-1.13)	(-2.90)	(1.74)	(-0.02)	(2.59)	(2.68)
71	14	0.007	-0.057	0.032	-0.035	0.499***	0.427
Arts, Entertainment & Recreation		(1.59)	(-0.70)	(0.498)	(-0.413)	(4.82)	(1.60)
72	30	0.029***	-0.111	0.086	-0.177	0.558***	0.091
Accommodation & Food Service		(3.03)	(-1.59)	(0.31)	(-0.45)	(3.04)	(0.32)

*p-value<10%, **p-value<5%, ***p-value<1%

*Notes.*This table summarizes the Fama-MacBeth results at the industry level using the first 2 digits of the NAICS codes (1997 standard).

possibly because their major customers, household consumers, are omitted from our consideration. Note that the supply chain relationship is not a major factor for firms in some industries such as finance and insurance, so some results here are exploratory but included for completeness.

4.2. Second-order Effects

To test the hypotheses concerning second-order effects and systematic risk, we group firms in the same large industry sectors (based on the first digit of the NAICS code) into five quintiles according to centrality, which can be measured in various ways. Common measures to quantify centrality in networks include degree, closeness, betweenness, and eigenvector centrality. To use the correct measure for the sales in the supply chain network, we must consider the characteristics of importance that underlie each measure. Borgatti 2005 reviews these measures and classifies them based on characterization of network flows. First, network traffic could be assumed to follow a walk (both nodes and links can be repeated), a trail (a sequence in which no link is repeated), a path (a sequence in which no node is repeated), or a geodesic path (the shortest path between two nodes).

Second, network traffic can be assumed to spread serially (through only one path at a time), or in parallel (through multiple paths at the same time).

Though making generalizations about firm level shocks is problematic, we can provide some reasoning about how shocks may be transmitted from one firm to another. First, firm level shock is unlikely to follow a geodesic path, i.e. the shortest distance, because firm level shocks that transmit across a supply chain network do not have final recipients and are unlikely to follow the shortest path between firms. According to Borgatti 2005, this means that closeness and betweenness centrality are inappropriate for economic shocks since they implicitly assume that traffic follows geodesic paths. Second, economic shocks are likely to have feedback effects. A supply shock in one firm could affect the supply of downstream firms, which eventually could transmit back to the original firm through the purchase orders or the reserve sales. For instance, a shock to a microchip plant may affect the downstream device manufacturer's fulfillment, which may result in future reduced orders to the microchip plant due to goodwill loss; and a shock to an oil firm could affect the cost of gasoline, which affects the costs of a transportation firm, which could then affect the oil firm itself. Just because a shock originated in a firm does not imply that it is immune from a subsequent feedback shock. Thus, supply chain network shocks are unlikely to be restricted to follow paths or trails, in which nodes and links are not repeated. Based on these assumptions, the most appropriate centrality metric for economic links is eigenvector centrality. As discussed in Bonacich 1972, eigenvector centrality is the principal eigenvector of the network's adjacency matrix. Nodes are more central if they are connected to other nodes that are themselves more central[11]. The linear relationship in eigenvector centrality also corresponds to the linear relationships of shock propagation shown in the first-order effect. Since eigenvector centrality cannot always be applied to asymmetric adjacency matrices (Bonacich and Lloyd 2001), for simplicity, we make our sparse adjacency matrix symmetric by taking the maximum value of the upper and lower triangular components.

Since the eigenvector centrality measure is skewed, we take the log of centrality in these statistics. Figure 7 presents the histogram of log eigenvector centrality for all firms in SPLC, versus all manufacturing firms (NAICS 3) in SPLC, and all logistics firms (NAICS 4). The mean of centrality for SPLC is about 0.04%, while the mean for manufacturing and logistics are both about 0.05%. This means that the manufacturing and the logistics firms are relatively more central in the supply chain network than other firms on average, and a random shock that propagates through the network is likely to hit such a firm about 0.05% of the time. The histograms are slightly skewed negatively, reflecting the asymmetric nature of the network as discussed in Acemoglu et al. 2012.

[11] In matrix notation, this is $Wc = \lambda c$, c is the principal eigenvector of the adjacency matrix.

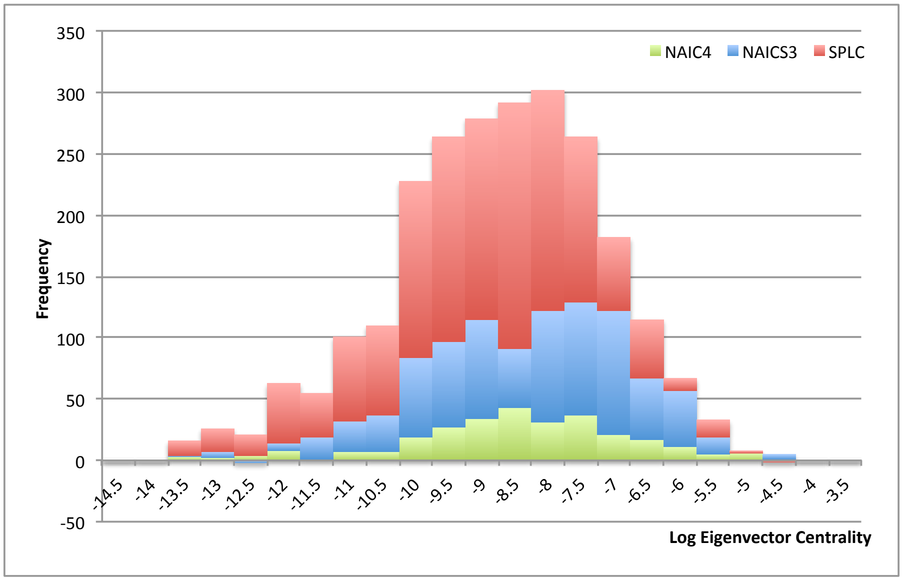

Figure 7 **Distribution of Eigenvector Centrality**

Apart from eigenvector centrality, and the closeness and betweenness centrality we have excluded, we test in-degree centrality (i.e., the number of suppliers) and out-degree centrality (i.e., the number of customers), which are proxies to the supplier multiplicity and the customer multiplicity. We are able to find a salient significant trend for in-degree centrality, but not for out-degree centrality, implying that the number of suppliers is associated with a firm's exposure to systematic risks, but not necessarily the number of customers. Due to data limitations, higher order network importance measures such as Herfindahl concentration may be misleading. After testing using the supplier Herfindahl concentration and customer Herfindahl concentration from our data, we do not find significant results in trends, and, therefore, only present results for eigenvector centrality and in-degree centrality.

In our dataset, 1041 firms fall into the manufacturing industry (NAICS code 31-33). As in the last section, we delete stocks with a price less than 5 dollars per share to avoid large liquidity effects and then sort firms into five quintiles based on the chosen centrality. The availability of the stock prices limits the sample size to 716 firms; therefore, each portfolio in the manufacturing industry group contains 143 firms.

For logistics firms (NAICS code 42-49), we find 284 firms that fall into this sector. After price selection, 238 firms remain. We construct portfolios as above so that each quintile contains 47 firms. We note that all of the results in this section are presented with the first-digit NAICS classification of 3 and 4; in unreported tables, we replicate our findings at the two-digit level and find results that are qualitatively identical.

We do not examine other industries due to data limitation, i.e., other industries do not have enough firms to form quintile portfolios to test the statistic significance. Figure 8 shows the industry

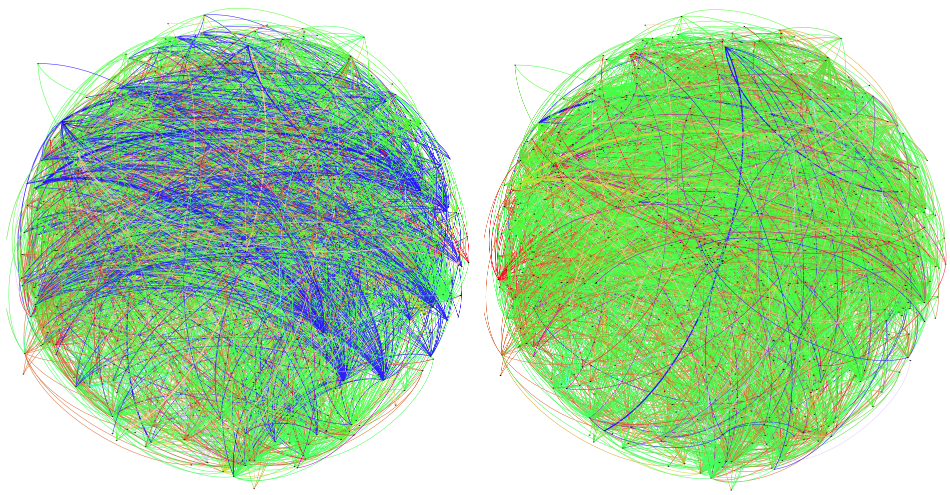

Figure 8 **Who are my customers (left) and who are my suppliers (right)**

Note. This figure plots all SPLC relationships subject to S&P 500 universe. Manufacturing firm relationships are colored in green, wholesaler and retailer are colored in blue, transportation are colored in red, and other industries use other colors. The left figure uses customer firm's color, while right figure uses supplier firm's color. The width of the link comes from the log sales.

of the firm on both ends of the supply chain. For the customer map on the left, most of the links this time are blue and green, meaning manufacturing and logistic firms are the major customers in the whole economy. For the supply map on the right, most of the links are green, meaning only manufacturing firms are the major suppliers in the whole economy. Therefore, the manufacturing and logistics firm we analyzed are actually the most central sectors in the supply chain network.

4.2.1. Firm Characteristics Sorted by Centrality In Table 9, we statistically verify that more central firms in manufacturing have lower stock returns than less central firms, while more central firms in logistics have higher stock returns than less central firms. We use eigenvector centrality for Table 9.

For the manufacturing industry, the lowest (value-weighted) quintile portfolio has an average monthly return of 1.77%, compared to only 1.07% per month for firms in the highest quintile, a statistically significant difference. The value-weighted portfolios are rebalanced monthly. We also present results based on equally-weighted portfolios, which again show a strong negative relationship between centrality and average returns. The economic magnitude of the relationship between centrality and average returns is substantial. For the value-weighted portfolios, the difference in returns between the highest and the lowest quintiles of eigenvector centrality is roughly -0.7% per month, or approximately -8.5% per year.

We examine other possible variables that may be related to centrality such as the size effect and the value effect. Using the log-scaled average size of firms in each quintile, we find a significant relationship between centrality and firm size. As eigenvector centrality increases, firm sizes are larger on average. For the value effect, however, the average ratio of book value to market value shows no salient trend as centrality increases.

For the logistics industry, the trend in average value-weighted returns as centrality increases seems to be opposite to that found in the manufacturing industry. We find statistical significance between two extreme quintile returns for the equal-weighted portfolios and observe an increasing relationship between centrality and average returns for both the value-weighted and the equal-weighted portfolios. Similar to the observations for manufacturing firms, firm sizes are larger as centrality increases, and no clear trend appears in the differences across quintiles for firm book-to-market ratios.

We use in-degree centrality for Table 10, which shows similar results to those in Table 10. We note that eigenvector centrality and in-degree centrality have, however, different foci. Since eigenvector centrality treats the network as undirected, it does not differentiate between suppliers and customers, but it captures the indirect information of how central a firm's linked partners are. Therefore, it gives more global information on a firm's centrality in the network. Since in-degree centrality omits customer information, it focuses on local supplier information. In-degree centrality also does not capture the indirect centrality information inherent in the firm's linked partners.

Overall, from this analysis, we observe that average stock returns have a positive relationship with both eigenvector centrality and in-degree centrality for manufacturing, while this relationship

Table 9 Firm Characteristics Sorted by Eigenvector Centrality

	1(High)	2	3	4	5(Low)	High-Low	t-stat
Manufacturing (31-33)							
(Ave.) Eigenvector centrality 10^{-3}	2.316	0.733	0.410	0.207	0.087	2.229***	(16.04)
Value weighted returns %	1.07	1.25	1.59	1.53	1.77	-0.70*	(-1.97)
Equal weighted returns %	0.41	0.63	0.71	0.82	0.98	-0.57*	(-1.90)
Log(average size)	6.972	6.359	6.082	6.139	6.022	0.952***	(5.18)
Average BE/ME	0.477	0.552	0.553	0.488	0.471	0.006	(0.08)
Logistics (42-49)							
(Ave.) Eigenvector centrality 10^{-3}	2.189	0.511	0.257	0.137	0.067	2.122***	(8.88)
Value weighted returns %	2.03	1.54	1.74	1.53	1.50	0.53	(0.93)
Equal weighted returns %	1.67	1.87	1.27	1.68	1.01	0.66*	(1.93)
Log(average size)	6.871	6.299	6.291	6.119	6.152	0.719***	(3.77)
Average BE/ME	0.552	0.530	0.500	0.649	0.536	0.016	(0.20)

*p-value<10%, **p-value<5%, ***p-value<1%

Note. This table summarizes the firm characteristics across five quintiles of eigenvector centrality, including average value-weighted monthly returns, average equal-weighted monthly return, log-scaled average firm size, and average book-to-market ratio. For manufacturing, each portfolio has 143 firms. For logistics, each portfolio has 47 firms. The value-weighted portfolios are rebalanced monthly.

is reversed for logistics firms. Given that common factors explain cross-sectional return variation, we additionally control for these factors in the subsequent factor regression tests.

4.2.2. 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 following CAPM and the four-factor model.

$$R_{it} - R_{ft} = \alpha_i + b_i\left(R_{mt} - R_{ft}\right) + \epsilon_{it}. \tag{5}$$

$$R_{it} - R_{ft} = \alpha_i + b_i\left(R_{mt} - R_{ft}\right) + s_i SMB_t + h_i HML_t + u_i MOM_t + \epsilon_{it}. \tag{6}$$

It is reasonable to imagine that network characteristics may be related to these market-wide factors. For example, high returns associated with centrality may be explained by exposure to market-wide excess stock returns. Or, centrality could be related to SMB, the size factor, where central firms behave more like large firms than small firms. Correlations with HML, the value factor, and MOM, the momentum factor, may also reflect the concentration of suppliers and customers.

Table 11 presents estimates of the time-series factor regressions on five sorted value-weighted centrality portfolios for firms in manufacturing, using the eigenvector centrality measure. The estimates reveal a clear pattern of decreasing excess returns (alphas) as centrality increases. The alpha in the lowest centrality quintile is 0.51% in the CAPM model and 0.93% in the four-factor

Table 10 Firm Characteristics Sorted by In-degree Centrality

	1(High)	2	3	4	1(Low)	High-Low	t-stat
Manufacturing (31-33)							
(Ave.) In-degree centrality	65.156	15.444	8.422	5.067	3.822	61.333***	(10.30)
Value weighted returns %	1.09	1.12	1.34	1.32	1.78	-0.70*	(-1.77)
Equal weighted returns %	0.67	0.96	0.87	0.87	1.22	-0.55*	(-2.06)
Log(Average Size)	7.194	6.512	6.231	5.962	5.928	1.266***	(8.32)
Average BE/ME	0.4386	0.4719	0.5116	0.5334	0.5145	-0.0759	(-0.95)
Logistics (42-49)							
(Ave.) In-degree centrality	68.037	17.259	9.444	5.3333	3.148	64.889***	(-7.01)
Value weighted returns %	1.67	1.67	1.55	1.52	1.34	0.33	(0.89)
Equal weighted returns %	1.98	1.56	1.39	1.29	1.37	0.61	(1.36)
Log(Average Size)	6.862	6.328	60402	6.079	6.253	0.610***	(3.24)
Average BE/ME	0.524	0.616	0.487	0.535	0.605	-0.081	(-0.93)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the firm characteristics across five quintiles of in-degree centrality, including average value-weighted monthly returns, average equal-weighted monthly return, log-scaled average firm size, and average book-to-market ratio. For manufacturing, each portfolio has 143 firms. For logistics, each portfolio has 47 firms. The value-weighted portfolios are rebalanced monthly.

model. In the highest centrality portfolio, the alpha estimate is 0.24% in the CAPM model and 0.11% in the four-factor model. The alpha difference between the two extreme quintiles is 0.27% for CAPM and 0.82% for the four factor model. The differences in alphas between the highest and lowest centrality portfolios are again statistically significant and economically meaningful. The explanatory power of centrality for stock returns is then not diminished even after controlling for known factors related to firm size, leverage, and momentum.

Among the other factors, trends seem present within the SMB factor. For the size effect in the four-factor regression, the portfolio's exposure to small stocks becomes lower as centrality goes higher. The lowest centrality quintile shows a coefficient of 0.78 on SMB, compared to a coefficient of -0.35 for the highest centrality quintile. To further control for firm sizes, we perform double-sorting quintile portfolios based on both centrality and firm size. We first sort all firms into five quintiles based on their sizes, and then, for each quintile, we sort firms into five sub-quintiles based on their centralities. We then construct value-weighted portfolios in each sub-quintile. In Table 12, we find the same results that excess returns decrease as centrality increases for manufacturing firms.

No clear trend, however, appears for market premium, HML, and momentum since the loadings are not statistically different between centrality quintiles. This reinforces the point that known common risk factors do not explain the role of supply chain network centrality for firm returns.

Table 13 presents estimates of the time-series factor regressions on five sorted value-weighted centrality portfolios for firms in logistics using the eigenvector centrality measure. Opposite to the results in Table 11 for manufacturing firms, the estimates reveal a clear pattern of increasing alphas as centrality increases. The alpha in the lowest centrality quintile is 0.76% in the CAPM model and 0.49% in the four-factor model. In the highest centrality portfolio, the alpha is 1.31% in the CAPM model and 1.43% in the four-factor model. The alpha difference between the two extreme quintiles is 0.56% for CAPM and 0.98% for the four-factor model. Other factors do not show statistically clear trends.

Tables 14 and 15 repeat the tests in Tables 11 and 13 using the in-degree centrality measure. The results are similar. Above all, the difference in alpha across extreme quintiles is statistically significant and economically important. This suggests that the standard pricing models do not explain all the cross-sectional variation in returns across centrality quintiles and that other factors related to network centrality might be included.

We also perform similar tests for other industries. In Tables 16 and 17, we find Mining (NAICS code 21), Utilities (NAICS code 22), and Construction (NAICS code 23) have the same eigenvector centrality result as observed for manufacturing firms (NAICS code 31-33). Comparing Table 15 with Table 11, we can see that mining, utilities, and construction earn significantly lower excess returns than manufacturing, as their alphas are all negative. This implies that firms in these industries are exposed to less residual systematic risk than manufacturing firms after controlling for common factors. Since mining, utilities, and construction are typically upstream industries compared to manufacturing, this evidence implies that upstream firms may be exposed to less systematic risk due to supply chain network structure. One possible reason for the difference here is that many manufacturing firms may have closer-to-linear supply chains with less diversification and higher systematic risks than firms with more connections. On the contrary, utilities, construction, and mining firms may be part of more networked chains, providing inputs to many other firms in different industries, which may mitigate risk exposure beyond the common factors. This observa-

Table 11 Factor Sensitivities by Eigenvector Centrality for Manufacturing Firms

N3	Factor Loadings					
Portfolio	α (%)	$R_{mt} - R_{ft}$	SMB	HML	MOM	$Adj.\ R^2$(%)
1(High)	0.235	0.888***				90.85
	(1.50)	(15.47)				
	0.114	0.894***	-0.347*	0.018	0.084	90.01
	(0.49)	(12.23)	(-2.07)	(0.119)	(1.025)	
2	0.295*	0.773***				88.74
	(1.78)	(13.79)				
	0.277	0.938***	-0.184	-0.453***	-0.061	93.77
	(1.34)	(14.28)	(-1.22)	(-3.29)	(-0.83)	
3	0.328	1.060***				92.78
	(1.33)	(17.60)				
	0.482*	0.953***	0.363*	-0.005	-0.008	93.04
	(1.86)	(11.63)	(1.93)	(-0.03)	(-0.09)	
4	0.356	1.256***				87.45
	(0.89)	(12.97)				
	0.571	1.087***	0.446	0.130	-0.142	87.82
	(1.36)	(8.22)	(1.47)	(0.47)	(-0.96)	
5(Low)	0.507	1.410***				85.54
	(1.55)	(11.96)				
	0.934*	1.157***	0.780**	-0.257	-0.132	87.53
	(1.95)	(7.63)	(2.24)	(-0.80)	(-0.78)	
High-Low	-0.272*	-0.522				
	(-1.72)	(-3.92)				
	-0.820*	-0.263	-1.127**	0.275	0.216	
	(-1.96)	(-1.28)	(-2.40)	(0.64)	(0.94)	

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table reports estimates of the Market Model (3) and the Fama-French model including a fourth momentum factor (4). Firms are chosen in the manufacturing industry sectors according to NAICS standard, including NAICS code 31-33. Portfolio returns are value-weighted firms returns formed for five quintiles of eigenvector centrality based on our supply chain data. Factor data are from Kenneth French's website. Observations are monthly returns from July 2011 to June 2013. 'High-Low' reports the difference between coefficient estimates from the first and the fifth centrality quintiles.

tion supplements the empirical bullwhip effects such as those observed in Cachon et al. 2007 and Osadchiy el al. 2011.

For service industries (NAICS code 51-56), we do not find clear trends. As discussed in Section 3, this may be due to the data limitations, as service industries do not have the rich data structure present in manufacturing and logistics. For example, household consumption may be the primary customer segment for firms in Arts, Entertainment, and Recreation (NAICS code 71) or Accommodation and Food Service (NAICS code 72), but that segment is not captured in our data. It is also possible that other factors drive the pattern in the service industries. For example, financial ownership may be a more dominant factor than supply chain relationship for the Finance and Insurance Industry (NAICS code 52).

The abnormal returns we find may be compensation for both financial and operating risk. Since capital structure should be uncorrelated with firm returns in an efficient market, we argue that adding controls for leverage, such as using unlevered returns, would only change the multipliers of the factor model but not the significance of our results. We, therefore, conclude that the abnormal returns reflect that, in addition to the common factors, additional variation in systematic risk effects can be explained by supply chain network structure.

4.2.3. Robustness test on industry concentration Literature such as Hou and Robinson 2006 find that firm's returns are related to industry concentration. The

Table 12 **25 Portfolios Formed by Double-soring on Firm Size and Eigenvector Centrality for Manufacturing**

N3	Firm Size									
Centrality	1(Large)	2	3	4	5(Small)	1(Large)	2	3	4	5(Small)
	CAPM α					$t(\alpha)$				
1(High)	0.207	-0.015	-1.423*	-0.874	-0.773	0.51	-0.02	-1.77	-1.22	-1.15
2	0.268	-0.082	-0.121	-0.730	0.148	0.94	-0.16	-0.26	-1.34	0.21
3	0.634**	0.005	-0.157	-0.008	-0.915	2.12	0.19	-0.35	-0.01	-1.21
4	0.561	0.050	0.445	-0.120	0.302	1.46	0.09	0.97	-0.14	0.50
5 (Low)	0.579*	0.911*	0.459	0.348	-0.415	1.72	1.74	1.35	0.66	-0.60
	Fama French & Momemtum factors α					$t(\alpha)$				
1(High)	0.047	-0.030	-0.298	-0.171	-0.131	0.16	-0.06	-0.58	-0.30	-0.21
2	0.348	0.608	0.181	-0.179	0.613	0.96	0.88	0.40	-0.41	0.96
3	0.356	0.547	0.363	0.692	-0.279	1.43	1.13	0.95	1.39	-0.39
4	0.829**	0.337	0.944**	0.721	0.664	2.09	0.58	2.30	0.98	1.03
5 (Low)	0.641*	1.106*	0.600	1.041**	0.029	1.81	1.90	1.66	2.69	0.05

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table reports 25 constant term α_i from the estimates of the Market Model (3) and the Fama-French model including a fourth momentum factor (4). The 25 portfolios are first sorted on sizes, then on eigenvector centrality. Portfolio returns are value-weighted. Firms are chosen in the manufacturing industry sectors according to NAICS standard, including NAICS code 31-33.

negative relationship for manufacturing firms and the positive relationship for logistics firms between centrality and returns could be influenced by other firm characteristics, such as greater concentrations of customer and supplier firms. Therefore, we further investigate whether the abnormal returns in manufacturing firms and logistics firms are relevant to their supplier or customer industry concentration, thus their supplier or customer industry competition. We measure industry j's concentration using the Herfindahl index, which is defined as

Table 13 Factor Sensitivities by Eigenvector Centrality for Logistics Firms

N4		Factor Loadings				
Portfolio	Alpha(%)	$R_{mt} - R_{ft}$	SMB	HML	MOM	$Adj.\ R^2$(%)
1(High)	1.314***	0.747***				84.93
	(3.26)	(7.62)				
	1.428***	0.768***	0.006	-0.589	0.024	86.43
	(3.44)	(5.85)	(0.02)	(-2.14)	(-0.16)	
2	0.894***	0.671***				70.41
	(3.78)	(11.67)				
	0.916***	0.976***	0.034	-0.502	0.031	72.32
	(2.41)	(8.13)	(0.13)	(-1.99)	(0.23)	
3	0.812**	0.964***				83.05
	(2.23)	(10.89)				
	0.801**	0.758***	-0.140	-0.152	0.164	83.75
	(3.36)	(10.03)	(-0.81)	(-0.96)	(1.93)	
4	0.708**	0.857***				86.41
	(2.50)	(12.40)				
	0.669**	0.916***	-0.171	-0.190	0.019	85.49
	(2.14)	(9.26)	(-0.75)	(-0.92)	(0.17)	
5(Low)	0.759	0.776***				69.60
	(1.44)	(6.03)				
	0.485	0.942***	-0.548	0.141	0.048	67.70
	(0.84)	(5.17)	(-1.31)	(0.37)	(0.23)	
High-Low	0.556	-0.029				
	(1.53)	(-0.20)				
	0.975*	-0.175	0.553	-0.730	-0.024	
	(1.93)	(-0.90)	(1.24)	(-1.69)	(-0.11)	

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table reports estimates of the Market Model (3) and the Fama-French model including a fourth momentum factor (4). Portfolio returns are value-weighted. Firms are chosen in the logistics industry sectors according to NAICS standard, including wholesale trade (NAICS code 42), retail trade (NAICS code 44-45), Transportation & Warehousing (NAICS code 48-49). Portfolio returns are value-weighted firms returns formed for five quintiles of eigenvector centrality based on our supply chain data. Factor data are from Kenneth French's website. Observations are monthly returns from July 2011 to June 2013. Stocks with a price less than five dollars are excluded to avoid liquidity effect. 'High-Low' reports the difference between coefficient estimates from the first and the fifth centrality quintiles.

$$H_j = \sum_{i=1}^{I} s_{ij}^2,$$

where s_{ij} is the market share of firm i in industry j. Market share can be computed using revenue or market equity. Both measures are only imperfectly correlated with true market share. We use both revenue and the market capitalization to construct the Herfindahl indexes. For firm i's supplier / customer industry concentration, denoted by SH_i and CH_j respectively, we use sales weighted average Herfindahl index, which is defined (by ourselves) as

Table 14 **Factor Sensitivities by In-degree Centrality for Manufacturing Firms**

N3	Factor Loadings					
Portfolio	α (%)	$R_{mt} - R_{ft}$	SMB	HML	MOM	$Adj.\ R^2$(%)
1(High)	0.100	1.274***				84.32
	(0.21)	(11.41)				
	0.260	0.895***	-0.297*	-0.298*	-0.029	91.64
	(1.18)	(12.86)	(-1.87)	(-2.04)	(-0.37)	
2	0.309	0.808***				89.73
	(1.35)	(14.51)				
	0.234	1.104***	-0.054	-0.310*	0.028	92.92
	(0.89)	(13.20)	(-0.28)	(-1.77)	(0.30)	
3	0.204	1.073***				92.92
	(0.82)	(17.64)				
	0.401	1.021***	0.236	0.106	-0.205*	91.66
	(1.30)	(10.42)	(1.05)	(0.52)	(-1.87)	
4	0.243	1.146***				90.82
	(0.80)	(15.45)				
	0.476	0.942***	0.101	-0.369	-0.040	87.12
	(1.03)	(8.94)	(0.42)	(-1.67)	(-0.34)	
5(Low)	0.851**	0.972***				87.04
	(2.72)	(12.73)				
	0.984***	1.083***	0.558	-0.398	-0.160	86.03
	(2.96)	(7.41)	(1.67)	(-1.30)	(-0.97)	
High-Low	-0.751*	0.302**				
	(-1.73)	(2.70)				
	-0.724*	-0.188	-0.855*	0.100	0.130	
	(-1.79)	(-1.01)	(-2.01)	(0.26)	(0.63)	

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table reports estimates of the Market Model (3) and the Fama-French model including a fourth momentum factor (4). Firms are chosen in the manufacturing industry sectors according to NAICS standard, including NAICS code 31-33. Portfolio returns are value-weighted firms returns formed for five quintiles of indegree centrality based on our supply chain data. Factor data are from Kenneth French's website. Observations are monthly returns from July 2011 to June 2013. 'High-Low' reports the difference between coefficient estimates from the first and the fifth centrality quintiles.

$$SH_i = \sum_{i=1}^{I} w_{ji} H_j, CH_i = \sum_{i=1}^{I} w_{ij} H_j$$

.

Similarly to the previous double-sorting method, we first sort all firms into five quintiles based on their supplier concentration or customer concentration; then, for each quintile, we sort firms into five sub-quintiles based on their centralities. In unreported tables, the trends in abnormal returns

Table 15 Factor Sensitivities by In-degree Centrality for Logistics Firms

N4		Factor Loadings				
Portfolio	Alpha(%)	$R_{mt} - R_{ft}$	SMB	HML	MOM	$Adj.\ R^2$(%)
1(High)	1.137***	0.553***				76.05
	(4.40)	(8.79)				
	1.02***	0.677***	-0.277	-0.245	0.143	80.04
	(4.04)	(8.50)	(-1.52)	(-1.47)	(1.61)	
2	0.828*	0.882***				85.01
	(2.06)	(9.00)				
	0.738*	0.963***	0.521*	-0.390	0.091	86.58
	(1.93)	(7.97)	(1.88)	(-1.54)	(0.67)	
3	0.577	1.015***				80.11
	(1.37)	(9.88)				
	0.495	1.104***	-0.178	-0.198	0.123	78.54
	(1.06)	(7.47)	(-0.52)	(-0.64)	(0.75)	
4	0.480	1.079***				76.91
	(1.27)	(11.71)				
	0.541	1.054***	-0.520	0.150	0.102	77.81
	(1.29)	(7.92)	(-1.71)	(0.54)	(0.68)	
5(Low)	0.339	1.046***				75.62
	(0.69)	(8.69)				
	0.054	1.221***	-0.496	0.131	0.135	75.11
	(0.10)	(7.24)	(-1.28)	(0.37)	(0.71)	
High-Low	0.798*	-0.493***				
	(1.92)	(-4.13)				
	0.962*	-0.544***	0.544	-0.646	0.096	
	(2.07)	(-3.17)	(0.56)	(-1.04)	(0.04)	

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table reports estimates of the Market Model (3) and the Fama-French model including a fourth momentum factor (4). Portfolio returns are value-weighted. Firms are chosen in the logistics industry sectors according to NAICS standard, including wholesale trade (NAICS code 42), retail trade (NAICS code 44-45), Transportation & Warehousing (NAICS code 48-49). Portfolio returns are value-weighted firms returns formed for five quintiles of indegree centrality based on our supply chain data. Factor data are from Kenneth French's website. Observations are monthly returns from July 2011 to June 2013. Stocks with a price less than five dollars are excluded to avoid liquidity effect. 'High-Low' reports the difference between coefficient estimates from the first and the fifth centrality quintiles.

still hold, meaning that the second-order centrality effect is robust after controlling for supplier and customer industry concentration.

5. Conclusion

In this paper, we find evidence that supply chain structure and firm returns are closely connected and that firms' supply chain relationships can explain this measure of supply chain performance, assuming that the supply chain structure is fixed in the short run. First, firm returns are influenced by the first-order effect of their supply chain partners' performance. With a network model of firm returns, we find that concurrent returns of both suppliers and customers are significant in explaining a firm's returns. We also observe significant lead-lag relationships from the firm's own lagged effect and the suppliers' lagged effect, but not from a customer lagged effect. A long-short equity strategy based on the supplier lagged effect yields monthly abnormal returns of 56 basis points. The cross lagged effect results have several important implications for returns information diffusion in supply chain networks.

From the financial market perspective, this result may indicate investors' limited attention to suppliers. Another possible reason is that supplier information is generally harder to obtain than customer information, since firms are more reluctant to disclose suppliers than customers, perhaps to protect proprietary suppliers from competitor firms.

From the operations management perspective, this result may indicate that the supply chain generally coordinates better in the upstream than in the downstream direction. Customer firms may not know all of their supplier's information until the one-month lag has elapsed. The result may also indicate larger market power for supplier firms in the supply chain than customer firms. Another possible explanation is that customer firms may order less foreseeing a demand shock, causing supplier firms to show a decrease in revenue ahead of customer firms due to input delivery lead time.

Table 16 Firm Characteristics Sorted by Eigenvector Centrality

	1(High)	2	3	4	5(Low)	High-Low	t-stat
NAICS code (21-23)							
(Ave.) Eigenvector centrality 10^{-3}	0.266	0.082	0.047	0.027	0.006	0.260***	11.31
Value weighted returns %	0.49	0.69	0.85	1.01	1.10	-0.62	-1.49
Equal weighted returns %	0.11	0.32	0.78	0.76	0.86	-0.75*	1.79
Log(average size)	6.921	6.730	6.350	6.550	6.233	0.688***	4.08
Average BE/ME	0.614	0.756	0.694	0.597	0.551	0.062	1.06

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the firm characteristics across five quintiles of eigenvector centrality, including average value-weighted monthly returns, average equal-weighted monthly return, log-scaled average firm size, and average book-to-market ratio. Each portfolio has 31 firms. The value-weighted portfolios are rebalanced monthly.

We observe some variation in the results across different sectors. Possible reasons for these observations may be that some industries have better supply chain coordination than others or that investors may pay more attention to supply chain information in certain industries so that those industries only have significant concurrent supply and customer effects. For firms with insignificant concurrent cross-firm effects, their major suppliers and customers may reside in economic sectors beyond the scope of this paper, i.e., private firms, government, household, or the foreign sector.

Our second main finding concerns the second-order impact of a firm's network position, which explains part of its systematic risk. From the fundamental theory of idiosyncratic shock transmission leading to aggregate risks, we argue that the capability of risk diversification by incorporating

Table 17 Factor Sensitivities by Eigenvector Centrality for Mining, Utilities and Construction Firms

N2		Factor Loadings				
Portfolio	α (%)	$R_{mt}-R_{ft}$	SMB	HML	MOM	$Adj.\ R^2(\%)$
1(High)	-1.153*	1.399***				79.54
	(-1.74)	(9.09)				
	-1.179	1.458***	-0.091	-0.330	0.114	76.83
	(-1.52)	(5.80)	(-0.15)	(-0.68)	(0.45)	
2	-0.897	1.512***				79.42
	(-1.25)	(9.06)				
	-1.023	1.583***	-0.329	0.092	-0.103	76.28
	(-1.21)	(5.76)	(-0.48)	(0.17)	(-0.37)	
3	-0.346	0.762***				61.90
	(-0.62)	(5.93)				
	-0.680	0.935***	-0.458	0.262	0.155	59.96
	(-1.09)	(4.63)	(-0.92)	(0.67)	(0.76)	
4	-0.374	1.129***				72.88
	(-0.58)	(7.58)				
	-0.598	1.213***	-0.071	-0.376	0.261	71.72
	(-0.83)	(5.20)	(-0.12)	(-0.84)	(1.11)	
5(Low)	-0.479	1.339***				78.22
	(-0.72)	(8.74)				
	-0.626	1.456***	-0.201	-0.215	0.221	75.94
	(-0.82)	(5.90)	(-0.33)	(-0.45)	(0.89)	
High-Low	-0.674*	0.060				
	(-1.95)	(1.25)				
	-0.553	0.002	0.110	-0.115	-0.107	
	(-1.51)	(0.02)	(0.51)	(-0.68)	(-1.20)	

*p-value<10%, **p-value<5%, ***p-value<1% *Notes.* This table reports estimates of the Market Model and the Fama-French model including a fourth momentum factor. Portfolio returns are value-weighted. Firms are chosen in the NAICS standard starting with first digit "2", including Mining (NAICS code 21), Utilities (NAICS code 22) and Construction (NAICS code 23). Portfolio returns are value-weighted firms returns formed for five quintiles of eigenvector centrality based on our supply chain data. Factor data are from Kenneth French's website. Observations are monthly returns from July 2011 to June 2013. 'High-Low' reports the difference between coefficient estimates from the first and the fifth centrality quintiles.

more supply chain partners actually depends on the correlation of the idiosyncratic shocks. For manufacturing industries, firms can choose multiple less correlated partners to diversify idiosyncratic risks so that more central firms are exposed to less systematic risks and earn lower returns on average. For logistics industries, it may be difficult or costly for firms to hedge idiosyncratic risks from their partners, as their supply chain partners are more likely to be correlated due to geographical or industry proximity. As a result, more central firms in the logistics industries are exposed to higher systematic risks and thus earn higher returns on average. We also find that firms in mining, utilities, and construction industries share similar results to manufacturing firms while our limited data for service industries do not yield a clear pattern. Fundamentally different from the industry level results and underlying economic support from Ahern 2013, which argues more central industries earn higher expected returns monotonously, we find non-monotonous opposite systematic risk effects for firms in different industries.

Our results hold for both the eigenvector centrality measure and the in-degree centrality measure. We do not find significant results for out-degree centrality. This result implies that, from the systematic risk perspective, supplier relationships are more important than customer relationships. Other centrality measures including in-degree Herfindahl and out-degree Herfindahl are difficult to use due to our data limitations. In general, our finding improves on ex-post statistical measures of well known common risk factors and provides new evidence to support the view that firm-specific shocks may aggregate to form economy-wide volatility. It also demonstrates that firms' decisions on supply chain structure may form part of their economic fundamentals as an ex-ante determinant of systematic risk.

For managerial implications, our results suggest that managers should be aware of both the concurrent effects from the direct connections to their customers and suppliers on their firm's returns performance, as well as the suppliers' previous performance. We also suggest managers in different industries should adopt different supply chain strategies towards the control of systematic risk due to the nature of the industry. For manafacturing, our results reinforce the support for operational hedging of supply, such as the form used by Nokia, although managers should also be aware that decreasing exposures to systematic risk in this way can also lead to lower future long-term average returns. For logistics, our results suggest that operational hedging may be costly.

For future work, we plan to further investigate the relationship between supply chain structure and firm returns using event studies or different granularity levels and to test for the effects of other characteristics, such as higher order centrality measures, for which our current data set is not sufficient. We also plan to build a panel over time to detect causes of the lagged effect, to identify the risk mediation effect in manufacturing and the opposite effect in logistics, to examine the micro-foundational implications, and to build and test a normative model of supply chain formation that includes these observed linear and nonlinear effects. We may also address other interesting questions by analyzing available supply chain data not used in this paper, such as the un-quantified sales relationships, the cost of good sold (COGS), the competitor information, as well as relationships involving government, household, and foreign sectors.

Acknowledgments

This work was support by the University of Chicago Booth School of Business. We thank Bryan Kelly for suggestions and support on the Bloomberg data, Lauren Cohen for his comments and for sharing the Compustat data. We also thank Eugene Fama, Bora Keskin, Serguei Netassine, Vishal Gaur, and Alireza Tahbaz-Salehi for their valuable feedback, and seminar participants from the Chicago Booth Operations group for their comments.

References

[Acemoglu et al. 2012] Acemoglu D, Carvalho VM, Ozdaglar A, Tahbaz-Salehi A (2012) The network origins of aggregate fluctuations. *Econometrica* 80(5):1977-2016.

[Ahern 2013] Ahern KR (2013) Network centrality and the cross section of stock returns. Working paper, Marshall School of Business, University of Southern California, Los Angelos, CA.

[Anupindi and Akella 1993] Anupindi R, Akella R (1993) Diversification under supply uncertainty. *Management Sci.* 39(8):944-963.

[Atalay et al. 2011] Atalay E, Hortaçsu A, Roberts J, Syverson C (2011) Network structure of production. *Proc. Natl. Acad. Sci. USA* 108(13):5199-5202.

[Atalay et al. 2013] Atalay E, Hortasu A, Syverson C (2013) Vertical integration and input flows. *Amer. Econom. Rev.* Forthcoming.

[Balch 2013] Balch T (2013) Assessment of a supply chain-based indicator for stock trading. http://augmentedtrader.wordpress.com/2013/07/26/assessment-of-a-supply-chain-based-indicator-for-stock-trading/

[Banz 1981] Banz RW (1981) The relationship between return and market value of common stocks. *J. Financial Econom.* 9(1):3-18.

[Barberis et al. 1998] Barberis N, Shleifer A, Vishny R (1998) A model of investor sentiment. *J. Financial Econom.* 49(3):307-343.

[Berk et al. 1999] Berk JB, Green RC, Naik V (1999) Optimal investment, growth options, and security returns. *J. Finance* 54(5):1553-1607.

[Badrinath el at. 1995] Badrinath S, Kale J, Noe T (1995) Of shepherds, sheep, and the cross-autocorrelations in equity returns. *Rev. of Financial Stud.* 8, 401-430.

[Brennan el at. 1993] Brennan M, Jegadeesh N, Swaminathan B (1993) Investment analysis and the adjustment of stock prices to common information. *Rev. of Financial Stud.* 6, 799-824.

[Bonacich 1972] Bonacich P (1972) Factoring and weighting approaches to status scores and clique identification. *J. Math. Socio.* 2:113-120.

[Bonacich and Lloyd 2001] Bonacich P, Lloyd P (2001) Eigenvector-like measures of centrality for asymmetric relations. *Social Networks* 22(3):191-201.

[Bondt and Thaler 1985] Bondt WFM, Thaler R (1985) Does the stock market overreact. *J. Finance* 40(3):793-805.

[Borgatti 2005] Borgatti SP, (2005) Centrality and Network Flow. *Social Networks* 27, 55-71.

[Cachon et al. 2007] Cachon GP, Randall T, Schmidt GM (2007) In search of the bullwhip effect. *Manufacturing Service Oper. Management* 9(4):457-479.

[Cahan et al. 2013] Cahan R, Chen Z, Wang S, Luo Y, Alvarez M, Jussa J (2013) Signal processing: uncovering hidden economic links. *Deutsche Bank Markets Research* March 28, 2013

[Carhart 1997] Carhart MM (1997) On persistence in mutual fund performance. *J. Finance* 52(1):57-82.

[Carvalho and Gabaix 2013] Carvalho VM, Gabaix X (2013) The great diversification and its undoing. *Amer. Econom. Rev.* 103(5):1697-1727.

[Chordia and Swaminathan 2000] Chordia T, Swaminathan B (2000) Trading volume and cross-autocorrelations in stock returns. *J. Finance* 52, 1035-1058.

[Clauset et al. 2009] Clauset A, Shalizi CR, Newman MEJ (2009) Power-law distribution in empirical data. *SIAM Review* 51(4):661-703.

[Cohen and Frazzini 2008] Cohen L, Frazzini A (2008) Economic links and predictable returns. *J. Finance* 63(4):1977-2011.

[Ellis et al. 2012] Ellis JA, Fee CE, Thomas SE (2012) Proprietary costs and the disclosure of information about customer. *J. Accounting Res.* 50(3):685-727.

[Fama and French 1988] Fama EF, French KR (1988) Permanent and temporary components of stock prices. *J. Political Econom.* 96(2):246-273.

[Fama and French 1993] Fama EF, French KR (1993) Common risk factors in the returns on bonds and stocks. *J. Financial Econom.* 33(1):3-56.

[Fama and MacBeth 1973] Fama EF, MacBeth JD (1973) Risk, return, and equilibrium: empirical tests. *J. Political Econom.* 81(3):607-636.

[Fruin et al. 2012] Fruin P, Osiol J, Wang K (2012) Supply chain interactions part 1: industries. *S&P Capital IQ*

[Fruchterman and Reingold 1991] Fruchterman T, Reingold E (1991) Graph drawing by force-directed placement. *J. Software: Practice and Experience* 21(11):1129-1164.

[Hou 2006] Hou K (2006) Industry information diffusion and the lead-lag effect in stock returns. *Rev. of Financial Stud.* 20, 1113-1138.

[Hou and Moskowitz 2005] Hou K, Moskowitz T (2005) Market frictions, price delay, and the cross-section of expected returns. *Rev of Financial Stud.* 18, 981-1020.

[Hendricks and Singhal 2003] Hendricks KB, Singhal VR (2003) The effect of supply chain glitches on shareholder wealth. *J. Oper. Management* 21(5):501-522.

[Hou and Robinson 2006] Hou K, Robinson D (2006) Industry concentration and average stock returns. *J. Finance* 71(4):1927-1956.

[Jegadeesh 1990] Jegadeesh N (1990) Evidence of predictable behavior of security returns. *J. Finance* 45(3):881-898.

[Jegadeesh and Titman 1993] Jegadeesh N, Titman S (1993) Returns to buying winners and selling losers: implications for stock market efficiency. *J. Finance* 48(1):65-91.

[Johnson 2002] Johnson T (2002) Rational momentum effects. *J. Finance* 57(2):585-608.

[Kelly el at. 2013] Kelly B, Lustig HN, Nieuwerburgh SV (2013) Firm volatility in granular networks. Working paper, Booth School of Business, University of Chicago, Chicago, IL.

[Lintner 1965] Lintner J (1965) The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Rev. Econom. Statist.* 47(1):13-37.

[Lo and MacKinlay 1988] Lo AW, MacKinlay AC (1988) Stock market prices do not follow random walks: evidence from a simple specification test. *Rev. Financial Stud.* 27(3):41-66.

[Lo and MacKinlay 1990] Lo AW, MacKinlay AC (1990) When are contrarian profits due to stock market overreaction? *Rev. Financial Stud.* 3, 175-205.

[Menzly and Ozbas 2010] Menzly L, Ozbas O (2010) Market segmentation and cross-predictability of returns. J. Finance 65(4):1555-1580.

[Moskowitz and Grinblatt 1999] Moskowitz TJ, Grinblatt M (1999) Do industries explain momentum? *J. Finance* 54(4):1249-1290.

[Motif Investing 2014] Motif Investing (2014) AAPL suppliers. https://trader.motifinvesting.com/motifs/aapl-suppliers-Te8hfgRR#/overview

[Osadchiy el al. 2011] Osadchiy N, Gaur V, Seshadri S (2011) Systematic risk and the bullwhip effect in supply chains. Working paper, Emory University Goizueta Business School, Atlanta, GA.

[Pástor and Stambaugh 2003] Pástor L, Stambaugh RF (2003) Liquidity risk and expected stock returns. *J. Political Econom.* 111:642-685.

[Petersen 2009] Petersen M (2009) Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies* 22, 435- 480.

[Rosernberg et al. 1985] Rosernberg B, Reid K, Lanstein R (1985) Persuasive evidence of market inefficiency. *J. Portfolio Management* 11(3):9-17.

[Salvini et al. 2012] Salvini M, Mesomeris S, Avettand-Fenoel JR, Wang Y (2012) Quantitative musing: sales segment momentum. *Deutsche Bank Quantitative Strategy* July 31, 2012.

[Sagi and Seasholes 2007] Sagi JS, Seasholes MS (2007) Firm-specific attributes and the cross-section of momentum. *J. Financial Econom.* 84(2):389-434.

[Shahrur et al. 2010] Shahrur H, Becker Y, Rosenfeld D (2010) Return predictability along the supply chain: the international evidence. *Financial Analysts J.* 66(3):60-77.

[Sharpe 1964] Sharpe WF (1964) Capital asset prices: a theory of market equilibrium under conditions of risk. *J. Finance* 19(3):425-442.

[Wall Street Journal 2001] Wall Street Journal (2001) A fire in albuquerque sparks crisis for european cell-phone giants. http://online.wsj.com/news/articles/SB980720939804883010

[Yahoo Finance 2013] Yahoo Finance (2013) A trading strategy involving Apple suppliers has apparently been way more profitable than investing in Apple itself. http://finance.yahoo.com/news/trading-strategy-involving-apple-suppliers-180543300.html

Appendix of "Supply Chain Network Structure and Firm Returns"

Jing Wu
Booth School of Business, University of Chicago, Chicago, IL 60637, wujing@chicagobooth.edu

John R. Birge
Booth School of Business, University of Chicago, Chicago, IL 60637, john.birge@chicagobooth.edu

History: This document supplements the paper "Supply Chain Network Structure and Firm Returns". It provides discussion on Fama-MacBeth regression and Network Centrality, as well as robustness test results mentioned in the paper.

---

1. The example in Subsection 2.2

Suppose an economy with 2 regions (A and B) and 3 potential future states with equal probability ($Prob(S = S_i) = \frac{1}{3}$, $\forall i \in \{1, 2, 3\}$):

S_1: both A and B function;

S_2: A cannot produce and B can;

S_3: B cannot produce and A can.

Next, suppose we have 4 firms in the economy, 3 manufacturers and 1 distributer. For the manufacturers, it is limited in production capacity, and it produces a payoff of 1 (due to fixed production capacity) as long as one of their input region function. Firm 1, 2, and 3 are manufacturers. Firm 1 only sources input from region A, Firm 2 only sources input from region B, and Firm 3 sources from both regions. Firm 4 is the distributor, it connects to both region A and region B with a fixed cost of 1 in all states. Therefore, in each of the states mentioned above, the payoff for these 4 firms are below:

$\Pi_1 = \{1, 0, 1\}$, $\Pi_2 = \{1, 1, 0\}$, $\Pi_3 = \{1, 1, 1\}$, $\Pi_4 = \{1, 0, 0\}$.

Let Ω denote the covaraince matrix for the firms' payoffs. Then we have

$$\Omega = \begin{bmatrix} \frac{1}{3} & -\frac{1}{6} & 0 & \frac{1}{6} \\ -\frac{1}{6} & \frac{1}{3} & 0 & \frac{1}{6} \\ 0 & 0 & 0 & 0 \\ \frac{1}{6} & \frac{1}{6} & 0 & \frac{1}{3} \end{bmatrix} \tag{1}$$

Suppose we have a representative mean-variance investor, and let $\mu = [\mu_1, \mu_2, \mu_3, \mu_4]$ denote firms expected return. Then for any feasible returns $\tilde{\mu}$ the investor targets, the investor find the portfolio weights $w = [w_1, w_2, w_3, w_4]$ by solving

$$\begin{aligned} &\underset{w}{min} w^{'}\Omega w \\ &s.t. w^{'}\mu = \tilde{\mu}, w^{'}1 = 1 \end{aligned}$$

By differentiating the Lagrangian with respect to w we get $\Omega w - \lambda_1\mu - \lambda_2 1 = 0$. The symmetry of firm 1 and firm 2 gives $w_1 = w_2$ and $\mu_1 = \mu_2$. By plugging in the values, we have

$$\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix} = \frac{1}{\lambda_1} \begin{bmatrix} \frac{1}{6}w_1 + \frac{1}{6}w_4 \\ \frac{1}{6}w_1 + \frac{1}{6}w_4 \\ 0 \\ \frac{1}{3}w_1 + \frac{1}{3}w_4 \end{bmatrix} + \frac{\lambda_2}{\lambda_1}$$

Therefore, it is clear that $\mu_3 < \mu_1 = \mu_2 < \mu_4$, i.e. the manufacturers have lower risk than the distributor, and the dual sourcing manufacturer is less risky than the single sourcing manufacturer. This relationship is shown in our empirical result of the second order effect.

2. Fama-MacBeth Regression

OLS standard errors are uncorrelated when the residuals are independently and identically distributed (i.i.d.). When the residuals are correlated across observations, OLS standard errors can be biased and either over or underestimate the true variability of the coefficient estimates. The residuals of a given firm may have time series dependence for a given firm, which is called unobserved firm effect. Alternatively, the residuals of a given year may have cross-sectional dependence, which is called unobserved time effect. In the model specification of the first-order effect, we have defined $r_{i,t}$ as the return of firm i in month t, which is a linear combination of its own one-month lagged effect, weighted sum of supplier and customer one-month lagged effect, weighted sum of supplier and customer returns, as well as its own idiosyncratic shocks:

$$r_{i,t} = \alpha + \beta_1 r_{i,t-1} + \beta_2 \sum_j w_{ij}^{in} r_{j,t-1} + \beta_3 \sum_j w_{ij}^{out} r_{j,t-1} + \beta_4 \sum_j w_{ij}^{in} r_{j,t} + \beta_5 \sum_j w_{ij}^{out} r_{j,t} + \epsilon_{i,t}. \tag{2}$$

Since most time series correlation has been captured by the one-month lagged effects, and we have found trailing horizons of more than two months have insignificant effect on current returns, the model specification have little unobserved firm effect in the residuals after controlling for the one-month lagged effects. Therefore, we should focus on the unobserved time effect, thus we choose Fama-MacBeth regression to correct the possible biased estimate in OLS. Fama-MacBeth regression is first proposed by Fama and MacBeth 1973, and it is the most commonly used solution to the time effect in asset pricing literature. A detailed discussion of Fama-MacBeth regression versus other solutions such as clustered standard errors is given in Petersen 2009.

The Fama-MacBeth method estimates the loadings on risk factors in two steps to avoid problems of correlation across contemporaneous residuals in panel data. The first step runs T cross sectional regressions to get T estimates coefficients for assets, while the second step uses the average of the T estimated coefficients to find the loading estimates, which is below.

$$\hat{\beta}_{FM} = \sum_{t=1}^{T} \frac{\hat{\beta}_t}{T} \tag{3}$$

$$= \frac{1}{T} \sum_{t=1}^{T} \left(\frac{\sum_{t=1}^{N} X_{it} Y_{it}}{\sum_{t=1}^{N} X_{it}^2} \right) \tag{4}$$

$$= \beta + \frac{1}{T} \sum_{t=1}^{T} \left(\frac{\sum_{i=1}^{N} X_{it} \epsilon_{it}}{\sum_{i=1}^{N} X_{it}^2} \right) \tag{5}$$

and the estimated variance of the Fama-MacBeth estimate is calculated as

$$S^2 \left(\hat{\beta}_{FM} \right) = \frac{1}{T} \sum_{t=1}^{T} \frac{\left(\hat{\beta}_t - \hat{\beta}_{FM} \right)^2}{T-1} \tag{6}$$

The variance formula requires that cross sectional estimates of the coefficients are independent of each other, i.e. there is no firm effect. Since our model specification have little unobserved firm effect, Fama-MacBeth regression is a good solution to treat the unobserved time effect in the model, and should yield unbiased estimate.

Fama-MacBeth regression is used in the empirical tests of the first-order effects. Below is the complete Table 5 in the paper, including loadings on the common factors.

Table 1 Fama-MacBeth Regression after Controlling for Common Factors.

This table summarizes the Fama-Macbeth results after controlling for common asset pricing factors. It is the complete table 5 in the paper. †.

$$r_{i,t} = \alpha + \beta_1 r_{i,t-1} + \beta_2 \sum_j w_{ij}^{in} r_{j,t-1} + \beta_3 \sum_j w_{ij}^{out} r_{j,t-1} + \beta_4 \sum_j w_{ij}^{in} r_{j,t} + \beta_5 \sum_j w_{ij}^{out} r_{j,t}$$
$$+ b_i \left(R_{mt} - R_{ft}\right) + s_i SMB_t + h_i HML_t + u_i MOM_t + \epsilon_{it}$$

	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$	$R_{mt} - R_{ft}$	SMB_t	HML_t	MOM_t
Ave. Coef	-0.000	-0.086***	0.063***	0.010	0.111***	0.503*	0.007 ***	0.006***	-0.001	-0.002***
(T-Stat)	(-0.45)	(-9.16)	(3.42)	(0.23)	(4.28)	(1.78)	(14.71)	(7.06)	(-1.41)	(-4.22)
Ave. Coef	-0.001	-0.091***	0.050***	0.029			0.011 ***	0.007***	-0.000	-0.002***
(T-Stat)	(-1.09)	(-10.43)	(3.02)	(0.70)			(33.17)	(8.58)	(-0.68)	(-6.19)
Ave. Coef	-0.002*	-0.054***					0.011***	0.006***	-0.000	-0.002***
(T-Stat)	(-1.80)	(-7.93)					(37.58)	(8.11)	(-0.74)	(-5.22)
Ave. Coef	-0.001		0.029**				0.011***	0.007***	-0.001	-0.002***
(T-Stat)	(-1.60)		(2.29)				(34.70)	(9.30)	(-1.57)	(-6.84)
Ave. Coef	-0.002**			0.034*			0.011***	0.006***	-0.001	-0.002***
(T-Stat)	(-2.50)			(2.05)			(36.27)	(7.98)	(-1.16)	(-6.42)
Ave. Coef	-0.001**				0.126***		0.010***	0.005***	-0.001	-0.002***
(T-Stat)	(-1.75)				(6.24)		(28.19)	(7.30)	(-1.38)	(-4.94)
Ave. Coef	-0.002***					0.501*	0.009***	0.006***	-0.001	-0.002***
(T-Stat)	(-2.83)					(1.69)	(23.57)	(7.72)	(-1.08)	(-5.72)
Ave. Coef	-0.001		0.029**		0.130***		0.009***	0.006***	-0.001	-0.002***
(T-Stat)	(-0.886)		(2.14)		(5.93)		(24.89)	(8.12)	(-1.80)	(-5.36)
Ave. Coef	-0.003***			0.041		0.492**	0.009***	0.006***	-0.001	-0.002***
(T-Stat)	(-2.91)			(1.66)		(2.19)	(22.53)	(7.17)	(-1.20)	(-5.67)
Ave. Coef	-0.002***				0.114***	0.485*	0.008***	0.005***	-0.001	-0.001***
(T-Stat)	(-2.16)				(5.45)	(1.79)	(18.30)	(6.57)	(-1.37)	(-4.36)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the Fama-MacBeth results after controlling for common asset pricing factors. We have similar results to those in Table 4. The results are consistent for both univariate and multivariate cases. All factors are defined by self-financing portfolio. Factor data is from the Kenneth French data library.

3. Robustness Test on Investor Inattention

Although the results such as the supplier lagged effect are consistent with the investor's limited attention hypothesis, there are a number of other plausible explanations of the data. This section shows results for a series of robustness tests for investor inattention.

A number of papers find that larger firms, or firms with higher levels of analyst coverage, institutional ownership, and trading volume, lead smaller firms or firms with lower levels of analyst coverage, institutional ownership, and trading volume (e.g., Lo and MacKinlay 1990, Brennan el at. 1993 Badrinath el at. 1995, Chordia and Swaminathan 2000, Hou and Moskowitz 2005, Hou 2006). The supplier lag effect results could be caused by firms of different size, analyst coverage, institutional ownership, and trading volume. To ensure that our results are not driven by those alternative explanations, we conduct the following robustness tests.

To control for the firm size difference, we only pick the firms that have their market capitalization larger than the input supplier weighted firms' market capitalization, i.e.

$$ME_i > \sum_j w_{ij}^{in} ME_j \tag{7}$$

In other words, the firms we pick are all larger firms compared to their average supplier firms weighted by their purchase orders. Since smaller supplier firms are less noticeable to investors, then if we still see supplier lagged effect this should due to other reasons than the firm sizes. In Table 2, we see the supplier lag effect still exists. Actually, since what are left after the filtering are relatively larger firms, their supply chain relationships captured by SPLC are more likely to represent their actual supply chain position, the lag effect becomes even stronger by comparing the t-statistics with those without filtering out any firms.

To control for the institution ownership, we only pick those firms that have their institution ownership ratio larger than the input supplier weighted institution ownership ratio, i.e.

$$\left(\frac{InstitutionOwnedShares}{TotalShareOutstanding}\right)_i > \sum_j w_{ij}^{in} \left(\frac{InstitutionOwnedShares}{TotalShareOutstanding}\right)_j \tag{8}$$

In other words, the firms we pick are owned less than their average supplier firms by the institutions. Institution ownership data is from Thomson-Reuters Institutional Holdings (13F) Database [1]. The result is shown in Table 3, still the supplier lagged effect persists.

[1] http://www.whartonwrds.com/archive-pages/our-datasets/thomson-reuters-2/#sthash.V7aCJYVw.dpuf

Table 2 Fama-MacBeth Regression Controlling Market Capitalization

	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$
Ave. Coef	0.065***	-0.091***	0.070**	0.025	0.391***	0.370***
(T-Stat)	(6.29)	(-5.99)	(2.77)	(0.88)	(15.18)	(12.92)
Ave. Coef	0.014***	-0.103***	0.105***	0.045		
(T-Stat)	(14.87)	(-5.21)	(3.14)	(1.27)		
Ave. Coef	0.014***	-0.031**				
(T-Stat)	(18.15)	(-2.56)				
Ave. Coef	0.014***		0.047***			
(T-Stat)	(18.45)		(3.07)			
Ave. Coef	0.013***			0.031*		
(T-Stat)	(18.27)			(1.97)		
Ave. Coef	0.008***				0.589***	
(T-Stat)	(10.40)				(24.38)	
Ave. Coef	0.006***					0.650***
(T-Stat)	(6.09)					(22.14)
Ave. Coef	0.009***		0.032**		0.593***	
(T-Stat)	(10.52)		(2.14)		(24.32)	
Ave. Coef	0.006***			0.021		0.656***
(T-Stat)	(5.68)			(1.321)		(21.951)
Ave. Coef	0.006***				0.388***	0.373***
(T-Stat)	(6.42)				(15.89)	(13.19)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the Fama-MacBeth results of the regression (1) using concurrent returns and one-month lagged effect as independent variables. Since we want to test whether the robustness of whether larger suppliers affect the supplier lag effect, the firms are chosen so that supplier's ME < firm's ME.

To control for the analyst coverage, we only pick those firms that have their number of analyst forecast larger than the input supplier weighted number of analyst forecast, i.e.

$$AnalystForecastCount_i > \sum_j w_{ij}^{in} AnalystForecastCount_j \tag{9}$$

In other words, the firms we pick have higher analyst coverage than their average supplier firms. Analyst coverage data is from the IBES dataset. The average number of analyst forecast as of June 30, 2013 is 7.84, with Apple and Intel have largest number of analyst forecasts, 56 and 45 respectively. About 49.49% firms in the SPLC universe are not covered by any analyst forecast at all. The result is shown in Table 4, again the supplier lagged effect persists.

Lastly, to control for the trading volume, we only pick those firms that have their trading volume turnover rate larger than the input supplier weighted turnover rate, i.e.

$$\left(\frac{TradingVolume}{TotalShareOutstanding}\right)_i > \sum_j w_{ij}^{in} \left(\frac{TradingVolume}{TotalShareOutstanding}\right)_j \tag{10}$$

Table 3 Fama-MacBeth Regression Controlling Institution Ownership

	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$
Ave. Coef	0.002*	-0.090***	0.084***	0.027	0.414***	0.566***
(T-Stat)	(1.77)	(-6.84)	(3.45)	(0.79)	(13.14)	(14.07)
Ave. Coef	0.013***	-0.101***	0.119***	-0.003		
(T-Stat)	(10.09)	(-5.71)	(3.89)	(-0.08)		
Ave. Coef	0.013***	-0.041**				
(T-Stat)	(10.85)	(-3.71)				
Ave. Coef	0.013***		0.048**			
(T-Stat)	(11.57)		(2.53)			
Ave. Coef	0.012***			0.029		
(T-Stat)	(11.36)			(1.15)		
Ave. Coef	0.006***				0.631***	
(T-Stat)	(5.67)				(23.87)	
Ave. Coef	0.001					0.857***
(T-Stat)	(0.98)					(24.96)
Ave. Coef	0.008***		0.046**		0.636***	
(T-Stat)	(6.44)		(2.54)		(23.57)	
Ave. Coef	0.001			0.031		0.861***
(T-Stat)	(0.98)			(1.23)		(24.65)
Ave. Coef	0.002				0.414***	0.560***
(T-Stat)	(1.56)				(13.84)	(15.49)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the Fama-MacBeth results of the regression (1) using concurrent returns and one-month lagged effect as independent variables. Since we want to test whether the robustness of whether suppliers of higher institution ownership affect the supplier lag effect, the firms are chosen so that $\left(\frac{InstitutionOwnedShares}{TotalShareOutstanding}\right)_i > \sum_j w_{ij}^{in} \left(\frac{InstitutionOwnedShares}{TotalShareOutstanding}\right)_j$.

In other words, the firms we pick are traded more frequently than their average supplier firms. Share trading volume data comes from the CRSP dataset. The supplier lagged effect does not disappear based on the results in Table 5.

References

[Badrinath el at. 1995] Badrinath S, Kale J, Noe T (1995) Of shepherds, sheep, and the cross-autocorrelations in equity returns. *Rev. of Financial Stud.* 8, 401-430.

[Brennan el at. 1993] Brennan M, Jegadeesh N, Swaminathan B (1993) Investment analysis and the adjustment of stock prices to common information. *Rev. of Financial Stud.* 6, 799-824.

[Chordia and Swaminathan 2000] Chordia T, Swaminathan B (2000) Trading volume and cross-autocorrelations in stock returns. *J. Finance* 52, 1035-1058.

[Fama and MacBeth 1973] Fama EF, MacBeth JD (1973) Risk, return, and equilibrium: empirical tests. *J. Political Econom.* 81(3):607-636.

[Hou 2006] Hou K (2006) Industry information diffusion and the lead-lag effect in stock returns. *Rev. of Financial Stud.* 20, 1113-1138.

Table 4 Fama-MacBeth Regression Controlling Analyst Coverage

	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$
Ave. Coef	-0.000	-0.081***	0.047**	-0.007	0.377***	1.024*
(T-Stat)	(-0.11)	(-6.49)	(2.25)	(-0.17)	(16.15)	(1.89)
Ave. Coef	0.008***	-0.077***	0.071***	-0.067		
(T-Stat)	(7.94)	(-4.95)	(2.90)	(-0.75)		
Ave. Coef	0.008***	-0.035**				
(T-Stat)	(8.48)	(-3.48)				
Ave. Coef	0.008***		0.032**			
(T-Stat)	(8.70)		(2.19)			
Ave. Coef	0.008***			-0.12		
(T-Stat)	(8.34)			(-0.88)		
Ave. Coef	0.003***				0.590***	
(T-Stat)	(3.46)				(28.88)	
Ave. Coef	-0.001					1.230**
(T-Stat)	(-0.73)					(2.51)
Ave. Coef	0.004***		0.027*		0.595***	
(T-Stat)	(4.07)		(1.88)		(28.73)	
Ave. Coef	-0.001			-0.016		1.243***
(T-Stat)	(-0.74)			(-0.55)		(2.50)
Ave. Coef	-0.000				0.374***	1.001*
(T-Stat)	(-0.457)				(17.03)	(1.92)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the Fama-MacBeth results of the regression (1) using concurrent returns and one-month lagged effect as independent variables. Since we want to test whether the robustness of whether suppliers covered more by analysts affect on the supplier lag effect, the firms are chosen so that $AnalystForecastCount_i > \sum_j w_{ij}^{in} AnalystForecastCount_j$.

[Hou and Moskowitz 2005] Hou K, Moskowitz T (2005) Market frictions, price delay, and the cross-section of expected returns. *Rev of Financial Stud.* 18, 981-1020.

[Lo and MacKinlay 1990] Lo AW, MacKinlay AC (1990) When are contrarian profits due to stock market overreaction? *Rev. Financial Stud.* 3, 175-205.

[Petersen 2009] Petersen M (2009) Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies* 22, 435- 480.

Table 5 Fama-MacBeth Regression Controlling Trading Volume

	α	$r_{i,t-1}$	$\sum_j w_{ij}^{in} r_{j,t-1}$	$\sum_j w_{ij}^{out} r_{j,t-1}$	$\sum_j w_{ij}^{in} r_{j,t}$	$\sum_j w_{ij}^{out} r_{j,t}$
Ave. Coef	0.000	-0.081***	0.054**	0.060	0.429**	0.887***
(T-Stat)	(0.30)	(-7.94)	(2.53)	(0.17)	(17.39)	(2.41)
Ave. Coef	0.010***	-0.081***	0.087***	-0.045		
(T-Stat)	(8.56)	(-6.32)	(3.45)	(-0.69)		
Ave. Coef	0.009***	-0.038**				
(T-Stat)	(8.70)	(-4.42)				
Ave. Coef	0.009***		0.031*			
(T-Stat)	(9.13)		(1.96)			
Ave. Coef	0.009***			-0.08		
(T-Stat)	(9.05)			(-0.85)		
Ave. Coef	0.004***				0.653***	
(T-Stat)	(3.54)				(29.66)	
Ave. Coef	-0.002					1.158***
(T-Stat)	(-1.54)					(3.48)
Ave. Coef	0.005***		0.658***		0.026*	
(T-Stat)	(4.40)		(29.06)		(1.77)	
Ave. Coef	-0.001			1.166***		-0.009
(T-Stat)	(-1.20)			(3.45)		(-0.36)
Ave. Coef	-0.001				0.424***	0.882**
(T-Stat)	(-0.909)				(19.05)	(2.49)

*p-value<10%, **p-value<5%, ***p-value<1%

Notes. This table summarizes the Fama-MacBeth results of the regression (1) using concurrent returns and one-month lagged effect as independent variables. Since we want to test whether the robustness of whether suppliers traded more frequently affect on the supplier lag effect, the firms are chosen so that $\left(\frac{TradingVolume}{TotalShareOutstanding}\right)_i > \sum_j w_{ij}^{in}\left(\frac{TradingVolume}{TotalShareOutstanding}\right)_j$.