research@redgraphs.com

The Value of Timely Supply Chain Data

Martin Geissmann and Eric McGill

16 September 2020

Abstract

Using an extensive supply chain dataset provided by RedGraphs Inc., we derive various trading signals based on the network connections between companies. We form sorted quintile portfolios and find many alpha-generating long-short strategies involving customer and supplier momentum, fundamental profitability factors, unlisted companies, and complex interactions. While we could not replicate the results of the simplest strategies previously investigated by others, it is likely that these simple strategies have merely been incorporated into the market in the intervening years. Because these basic signals are now included in the market, we find a clear advantage in having complete and timely available data, in which low-latency supply chain data generally results in better performance. We also find that anything beyond 1st degree customer relationships are still not incorporated to the market and remain good candidates for alpha-generation going forward.

1 Introduction

If American Airlines (*AAL*) reports an unforeseen drop in their number of passengers, market participants update their expectations for AAL's future profitability. If the downturn turns out to be persistent, AAL is likely to order fewer jets in the time coming. To formulate that in a probabilistic way, we could say that the likelihood of a decreased order in the future has increased as a result of the fewer air traffic passengers today.

This of course will put obvious pressure on Boeing (*BA*), the main supplier of AAL's planes. Further up the supply chain, Kaiser Aluminium (*KALU*) is likely to be affected, too.

It is no novelty that economic actors are heavily interconnected. A shock (positive or negative) in one company's business will be felt by its suppliers (and to some extend also to its customers) by some factor. Having the knowledge of the supply chain allows investors to trace back inter-company spillover effects.

In general, sector analysts are well aware of the 1st degree outward connections, i.e. they know about a company's direct customers. Direct supplier information may many times be omitted in the public filings. Further out in the supply chain, 2nd degree relations might only be partially known of or considered, while higher degree ones are likely to be ignored altogether.

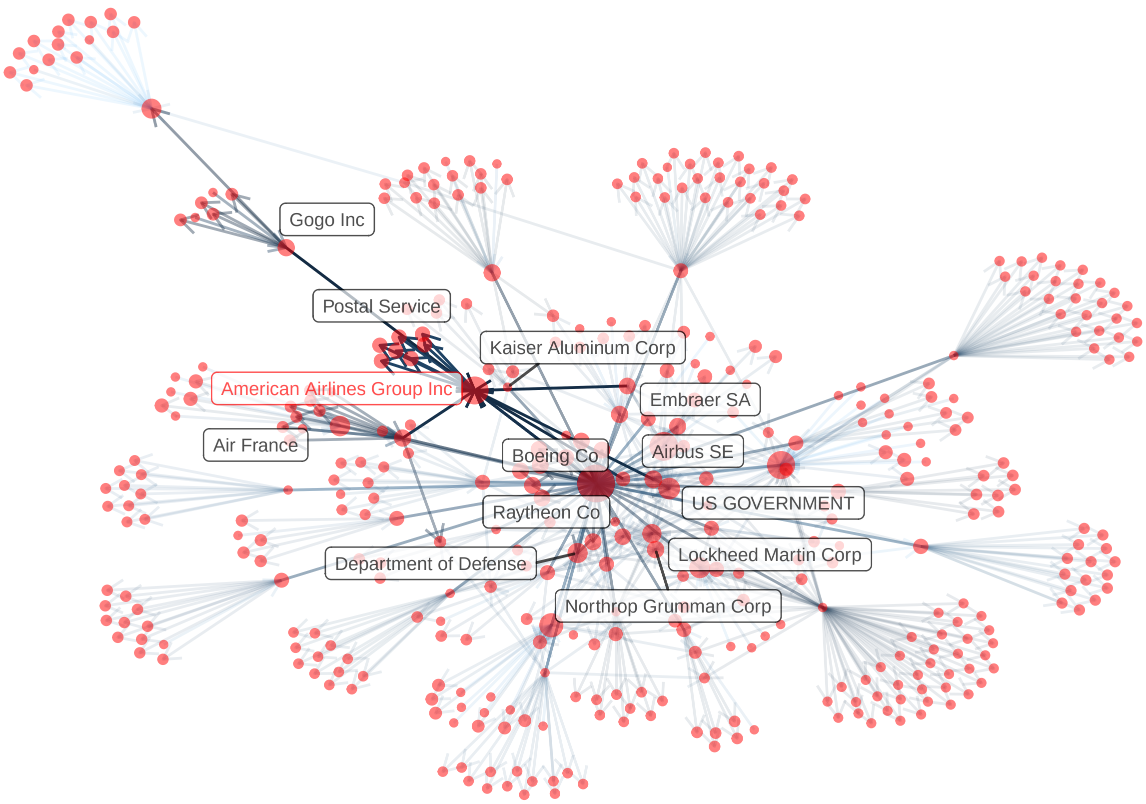


Figure 1: Network of AAL, up to the 3rd degree, as of 2018-01-01

Supply chain information of public companies generally comes incomplete and unstructured. An external collaboration might be mentioned once in a report or on an investor call, yet not be released in an official *customers and suppliers* list. While still many relations stay unknown, US listed companies are required to disclose their major customers in their 10-K reporting.[1] Sometimes additional information is provided at the discretion of the filing company, as well.

In the parts that follow, we assess the value in (structured) supply chain data for investors. We do so by building simple sorted quintile portfolios using metrics from the company's customers or suppliers. The long/short (*L/S*) strategies are formed on an equal weighted with a short leg of stocks from the first quintile and a long leg of stocks from the fifth quintile portfolio. We evaluate them below.

2 Method and Data

We look at three main categories of sorting approaches as suggested by Zhao et al. (2015).

First, we look at what has been called **return momentum** strategies. Those suggest that the stock performance of related companies (customers or suppliers) to some extend forecast the performance of a target company's stock. It assumes that the information efficiency on company-specific news is higher for the company itself and lower for related companies. For example, a company reports lower revenues, this is expected to be immediately priced into their stock (high information efficiency). However, it might however take a moment for the related companies' stocks to react accordingly.

Return momentum strategies look for momentum signals from the peers' stock performance. The strategy seeks to capitalize on the information that is already priced in one company, but not or only partially in its customers or suppliers, or the customers of the customers, or suppliers of suppliers.

Second, we focus on companies' fundamental metrics instead of using their peers' stock returns. This category of strategies are called **fundamental flow factor** strategies. Factors computed from related companies' fundamentals such as the return on equity (*ROE*), return on assets (*ROA*), or the earning yield are used to predict the performance of the target company's stock performance.

Third, **link interaction factor** strategies seek to find performance based on features directly induced from the number of suppliers or customers or from the supply chain network as a whole. The general idea is that a company with more customers and suppliers, or a company that is more prominently connected to others, is more diversified and independently profitable than a company whose revenue is largely dependent on a few large customers. Serving a single customer will amplify that customer's pricing power over the supplier, which for the target company is likely to lead to smaller profit margins. Simultaneously, if that single customer is idiosyncratically facing a rougher

[1] *Regulation S-K334 and U.S. GAAP335 both require disclosures about major customers. (...) Regulation S-K requires disclosure if loss of a customer, or a few customers, would have a material adverse effect on a segment. (...) U.S. GAAP requires disclosure of each customer that comprises 10 percent or more of total revenue.*, see SEC Disclosure Update and Simplification, Section III.D.3. available on https://www.sec.gov/rules/final/2018/33-10532.pdf

course of business, the supplier might be less flexible to generate its revenue from others, for which it must first establish new channels of sale.

The general approach followed in this paper is to form sorted quintile portfolios based on signals deviated from the aforementioned strategies. That is, for each point in time t, we compute a signal for each company c available in our dataset.

As an example, for a *return momentum* strategy, this would mean to be looking at the mean stock performance of all of customers of every company in time $t-1$ (we use the 1-month, 6-month, and 12-month lagged cumulative stock performance). This computed figure is the signal which we store in a (t, r) matrix. For each time t (each row) we sort the stocks according to the signal small to large and allocate them to five quantiles.[2]

Subsequently, we compute the return series of each of the five portfolios in which the stocks are equally weighted. Ideally, the sorting leads to visible differences and steadily increasing performance figures (average return or Sharpe ratio) between the quintile portfolios, i.e. the (risk-adjusted) return of the first quintile portfolio is smaller than second, which is smaller than the third, etc.[3]

For the evaluation, the performance of a L/S portfolio is calculated as the return difference between the fifth and the first quintile portfolios. To assess the profitability, we focus on the Sharpe ratio (*SR*), besides also testing against common risk factors (CAPM/Fama French).

The following data has been used:

- The main supply chain dataset was provided by RedGraphs Inc.,[4] a company that extracts supply chain information from various sources using a state-of-the-art machine learning approach (i.e. natural language processing, *NLP*). They provide structured data with minimal delay after publishment. An example of the main fields used can be found in Table 1. We use data covering a full four year span from 2016 to 2019.

Table 1: Sample data (RedGraphs Inc.)

VALID_FROM	**LAST_SEEN**	**TARGET_COMPANY**	**SOURCE_COMPANY**	**REL_TYPE**	**_**
2017-11-15	2018-11-26	Boeing Co	Rockwell Collins Inc	customer	...
2016-02-12	2019-02-07	Boeing Co	United Aircraft Corp. JSC	customer	...
2017-02-24	2018-02-22	Boeing Co	Iridium Communications Inc	customer	...
2018-11-16	2018-11-15	Boeing Co	Cubic Corp	customer	...
...	...	...	...	...	...

- Stock (log) returns are calculated from adjusted daily closing prices gathered from Alpha Vantage.[5] We exclude penny stocks if their closing price is $0.005 or

[2] We chose end-of-month as also the risk factors are provided on Kenneth R. French's Data Library (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html) are computed end-of-month.

[3] $r_{PF1} < r_{PF2} < r_{PF3} < r_{PF4} < r_{PF5}$

[4] See https://www.redgraphs.com/

[5] See https://www.alphavantage.co/

below. We winsorize the log returns at the 1% and at the 99% level as a measure to remove obvious erroneous changes in price.

- We use a yearly fundamentals extracted from the Alpha Dataset provided by Vhinny.com.[6] While higher frequency fundamentals would be beneficial, it is relatively hard to come by such data in a historical time series format.

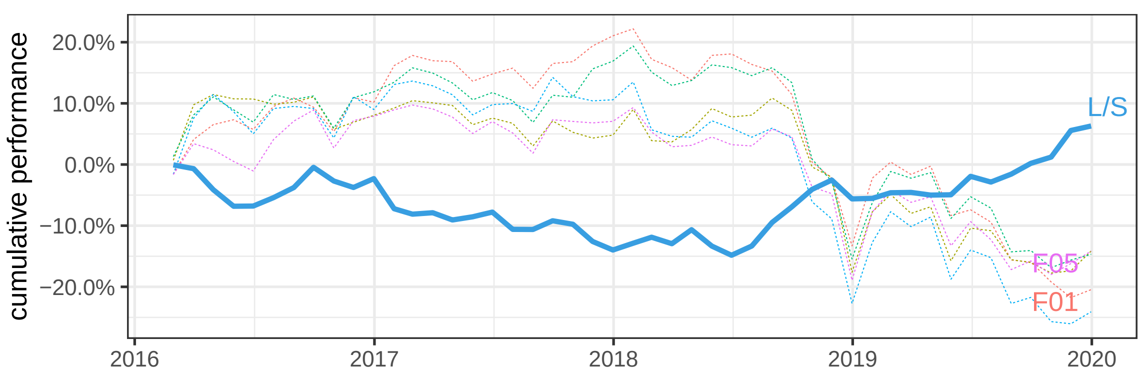
3 Results

3.1 Return momentum

In return momentum strategies we look first at customers' and then at suppliers' past stock performance. The notion of looking at customers is straightforward: a company should be doing well if its customers are doing well (which should be reflected in an increase in everyone's market valuation). One step up the supply chain, the suppliers of that one company should also be doing well.

The same mechanism for suppliers is a bit less intuitive (i.e. a company's stock is doing well if its suppliers are doing well), but could reflect suppliers providing better goods and services to their customers that improve the customers' productivity (or inversely dropping goods or services that require replacement). For it to work, the stock of a customer of a company should rise after the rise of the stock of its supplier. As reported in the following, we actually find patterns in both directions, yet less clear for the supplier side. This is reasonable because suppliers changing their offerings is less frequent than customers simply purchasing more or less of the same offerings.

Figure **??** shows the performance of a return momentum strategy with a return lag of one month on all direct customers. The strategy delivers a SR[7] of 1.07, and a significant CAPM alpha of 6.7% (annualized). Positive returns could also be found for the six and twelve months lagged direct customer momentum.



[6] See https://vhinny.com/download
[7] Annualized Sharpe ratio, 0% risk-free rate

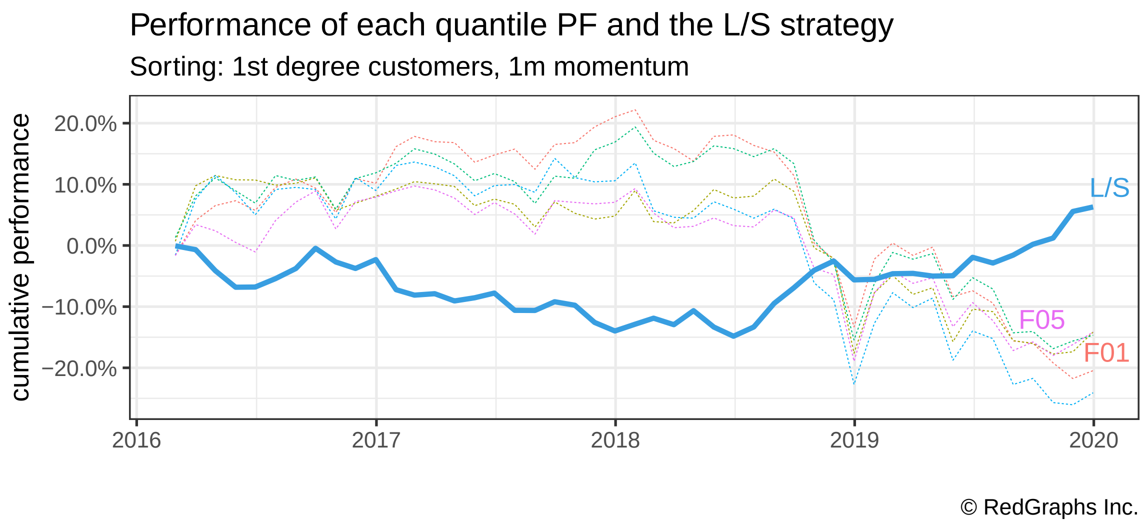


Table 2 gives the achieved SR from various momentum strategies tested. We report the SR as a measure to compare L/S performances with different volatility profiles.

Table 2: Sharpe ratio of momentum strategies

Strategy	**1m lag**	**6m lag**	**12m lag**
1st degree customers, equal weighted	0.290	0.146	0.218
1st degree suppliers, equal weighted	-0.045	0.540	0.678
2st degree customers, equal weighted	1.318	1.257	0.829
2st degree suppliers, equal weighted	0.670	0.228	0.330

Momentum strategies on direct suppliers could also be identified to have delivered positive (yet weaker) returns. While their CAPM alpha was mostly insignificant, the loading on the Fama French and momentum factor was small and insignificant, which shows that their performance is not driven by systematic market momentum. The same observation could also be made for the momentum strategies on the customer side, and implies that neither is fully incorporated into the stock market.

Looking at second degree customers, we also find the L/S strategy to have delivered positive abnormal returns, see Figure 2. The strategies on all momentum lags resulted in positively performing L/S portfolios, with the six month lagged signal on the second degree customers performance delivering the highest SR of 0.92.

The momentum strategy on first degree customers performed best with the shortest lag (1 month) but declined significantly with longer lags. The same momentum strategy on second degree customers delivered better results with a six months lag. This supports the argument of delayed spread of information across the network. On the company itself, information is quickly priced in, whereas it takes some time to spread to other companies up the supply chain as analysts slowly incorporate the information.

Also as expected, the less intuitive and less frequent supplier side momentum performance is indeed lower. A company's value does not seem to be influenced much by

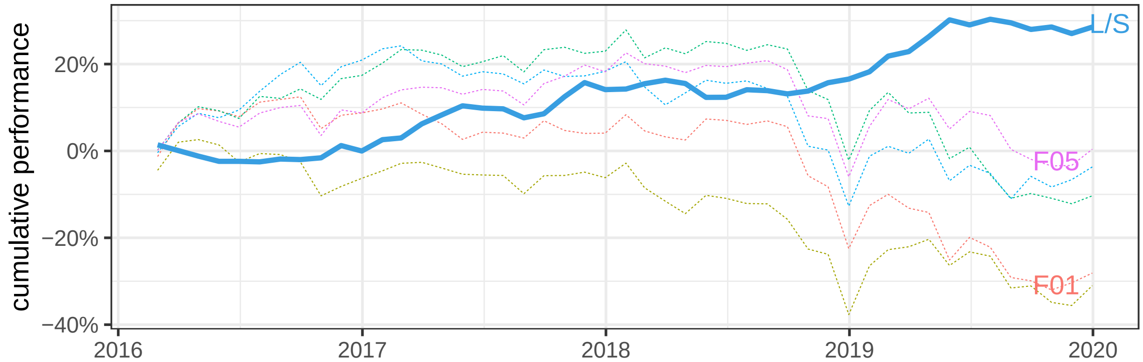


Figure 2: Return momentum: all 2nd degree customers equal, 6m lag

an idiosyncratic move in its suppliers' stock as opposed to its customers' stock.

One possible explanation for our signals' performance is that smaller capitalization stocks are generally covered by fewer analysts. Therefore, one might expect that those stocks exhibit higher market inefficiencies and would be more profitable in a supply chain momentum strategy. To evaluate such patterns, we form conditional double-sorted portfolios, i.e. first sorting the equities according to their market capitalization (where we use three quantiles low/medium/high), and second sorting for the momentum signal as before.

Contrary to the hypothesis, we do not find better performance with small-caps. This indicates that profound knowledge on the supply chain is relevant for all company sizes, small to large. For example, the large cap (top 1/3) subset shows a clear pattern in quantile portfolio returns according to the second degree customers (see Figure 3). We reason that while large caps are more covered by analysts, those do not necessarily take into account the elevated complexity of those companies' supply chain connections. This bodes well for the long-term success of a supply chain trading strategy.

3.2 Fundamental flow factors

Using the related companies recently reported fundamentals (and especially the change in fundamentals) is an appealing approach to forming portfolios. The metrics focused on in this paper reflect the financial performance of companies (*earning yield*, the *gross profit margin*, the *ROE*, and the *ROA*).

For financial metrics, we wanted to invert the portfolio construction method used above. Instead of building the L/S portfolio with a long position from the highest quantile and a short position from the lowest quantile, we choose the opposite. The long exposure would be on stocks whose peers recently had a decrease in a metric and a short exposure on stocks whose peers recently had an increase. The logic is that if suppliers (or customers) of a company report increasing profit margins, ROE, ROA, or earning yields, this indicates that they siphon off a larger share of the total value created as opposed to just sending more funds to their suppliers. Stock performance can come from increasing revenues or increasing margins, and it is important to check both.

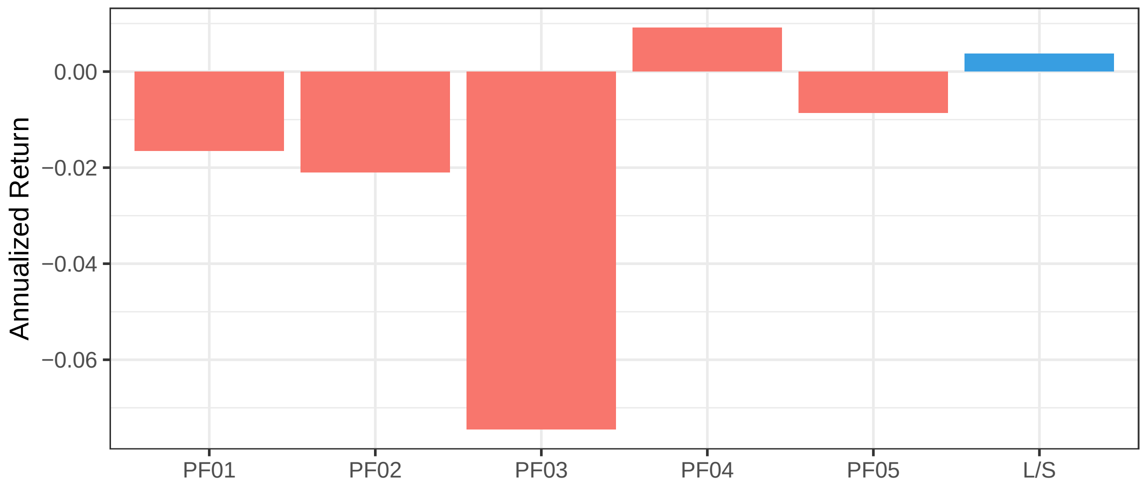


Figure 3: Return momentum: all 2nd degree customers equal, 6m lag, top 1/3 market cap only

We find positive performances especially in strategies based on signals from the supplier ROE and earning yield (see Table 3). This backs the hypothesis that if further up the supply chain increased profit margins are achieved, the company in focus is likely to be contributing to that increased profitability by overpaying. As a result, the company tends to perform worse than its peers.

Table 3: Sharpe ratio of fundamental flow strategies

Metric	on customer	on suppliers
Mean change in ROE	-0.227	-0.847
Mean change in ROA	-0.192	-0.891
Mean change in earning yield	-0.063	0.588
Mean change in gross profit margin	-0.149	-0.094

The same L/S strategies formed on customers' metrics still lead to positive returns, however with considerably lower SR.

We do not report the results on the portfolios constructed from the absolute financial metrics figures (as opposed to the changes), mainly due to them giving ambiguous results. It has to be assumed that the bare figures do not mean much as a static snapshot of a supply chain in temporary equilibrium. In other words, there is nothing inherently special about any given level of ROE, ROA, etc. for generating subsequent returns in the stock market. Rather, it is the changes in these metrics and their associated impact on the supply chain that matters. Markets can also be driven by surprises, i.e. deviations from the consensus' expectations, but we lack meaningful historical data to perform such a backtest.

3.3 Interaction factors

We initially used the mere number of first degrees inward, outward, or total, as signals, but the result in a L/S strategy was inconclusive. While similar to Zhao et al. (2015), some degree ratios (e.g. number of degrees divided by market capitalization) deliver better results, but such signals are seemingly heavily synthesized (*data dredging*) and don't reflect an underlying base truth.

In contrast, using changes in the number of degrees give better results. As shown in Table 4, we get positive returns (yet comparably small SR) for the change in number of outward degrees. That is to say, those companies that recently reported additional outward connections (i.e. new customers) outperformed, which makes sense.

We also report the change in *unlisted* degrees. Those are trading partners that are not listed on a stock exchange, which includes private companies but also several government or government-near entities (US Government, Department of Defense, United Nations etc.). This is motivated by the idea that such customers tend to be less prone to economic cycles and might tend to pay above market prices as they focus on a longer-term horizon. We observe that an increase in connections with unlisted companies resulted in a better performance (SR above 0.7 for all inward/outward/total), which clearly reflects the value of gathering data beyond the standard filings.

Table 4: Sharpe ratio of number of degrees strategies

Metric	**total**	**inward**	**outward**
Change in number of 1st degrees	-0.285	-0.107	0.282
Change in number of unlisted 1st degrees	0.330	0.452	-0.259

To incorporate graph characteristics beyond the first degree, we have to look at centrality measures. Centrality gives a score of importance to each node in a network. While several approaches to measuring the centrality have been proposed,[8] we chose the *eigenvector centrality* as it most closely follows the path of money flowing through the economy.

As expected, we find better performances in companies with higher centrality scores. More interconnected companies on average outperform on average. Interestingly, we do not observe a significant loading of the L/S portfolio's return series on the the Fama/French's SMB factor (nor on their other factors), indicating that the centrality incorporates further characteristics beyond a company just being small or large. Since complex network analytics are difficult to intuitively process, this is a logical finding.

We also find value in the change of the centrality score. Companies whose centrality has recently increased tend to outperform. The L/S strategy delivers a SR of 1.11, see Table 5.

3.4 The value from timeliness

While a company's direct connections are probably partially known by heart by its sector analysts, it would seem that having access to a data feed with minimal delay

[8]See Bonacich and Lloyd (2001)

Table 5: Sharpe ratio of centrality strategies

Signal	SR
Eigenvector centrality	0.782
Change in eigenvector centrality	0.768

after initial publication would be especially valuable in today's highly efficient news-driven markets.

To back this hypothesis, we compare the results of the L/S portfolios built on the most recent data points versus those built on data with a two month delay. The delayed signals are derived from supply chain data that was current two months prior.

The results given in Table 6 shows that most strategies performed considerably worse or even negatively on outdated signals. All momentum strategies did not deliver the exceptional results on the delayed data, suggesting that having the most recent data at hand is crucial. The same conclusion can be drawn for the interaction factors, and especially the eigenvector centrality.

Table 6: Sharpe ratio when delayed

Strategy	SR	SR (delayed 2m)	SR difference
Momentum/1st degree customers, equal w., 1m lag	0.290	-0.276	-0.567
Momentum/1st degree suppliers, equal w., 1m lag	-0.045	0.124	0.169
Momentum/2st degree customers, equal w., 1m lag	1.318	-0.539	-1.857
Momentum/2st degree customers, equal w., 6m lag	1.257	-0.114	-1.371
Momentum/2st degree suppliers, equal w., 1m lag	0.670	-1.002	-1.673
Momentum/2st degree suppliers, equal w., 6m lag	0.228	-1.178	-1.407
Change in number of 1st degrees, total	-0.285	0.161	0.446
Change in number of 1st degrees, inward	-0.107	-0.783	-0.676
Change in number of 1st degrees, outward	0.282	-0.359	-0.641
Change in number of unlisted 1st degrees, outward	-0.259	0.611	0.870
Change in number of unlisted 1st degrees, inward	0.452	0.663	0.211
Change in number of unlisted 1st degrees, total	0.330	1.042	0.712
Eigenvector centrality	0.782	-0.216	-0.998
Change in eigenvector centrality	0.768	0.739	-0.028

We conclude that timely data is not just a nice-to-have benefit to traders, but crucial in today's efficient markets.

4 Conclusion

In this paper, we have constructed a variety of portfolios based on supply chain data metrics and its timeliness. While some of the L/S portfolios achieved exceptional performance for the time window evaluated, there are also some strategies that do not lead to the same level of results. In general, the strategies that make intuitive sense and have clear reasoning behind them have worked well. From the analysis, the conclusions are straightforward:

- Investors still appear to focus primarily on customer relationships, as first degree customer momentum remains a strong near-term factor and is especially affected by timeliness of the data; second degree customer momentum shows its effect after a significant delay and is ripe for acceleration with a timely data source.

- Supplier momentum is a weaker effect than customer momentum, but is still not fully incorporated into the market.
- Changes in profitability metrics can push or pull value up and down the supply chain, affecting the future performance of customers and suppliers; this is likely separate from volume-driven propagation because of its opposite directional effect.
- Information on unlisted company relationships is not fully incorporated in the market.
- Complex metrics like eigenvector centrality are not fully incorporated in the market, either statically or dynamically.
- Timeliness of the data matters significantly.

Contrary to the results in previous publications (Zhao et al. 2015), we found less significant performance for simple strategies. We believe that as supply chain data has become more readily available, the benefit to an individual investor has decreased, especially if the information is outdated and only used to the first degree.

We believe that the following points might have also caused the alpha to be different from prior measurements:

- The equity market is highly efficient.[9] Simple signals quickly become incorporated, and spillovers are priced in with little to no delay. Thus, basic news from one company quickly results in a price reaction on the related customers and suppliers, at least on the first degree, making basic supply chain indicators now a requirement.
- Refinements of the strategies may be needed to fully replicate the measured alpha, such as filters for recent volatility that could prevent binning a stock in the improper quantile, higher frequency fundamental data to capture recent changes, or adjustments for expectations and surprises in forecasts.
- Coverage differences undoubtedly exist between the RedGraphs and the data used in previous publications because of the reporting time period and the sources used, which could easily drive measurement differences.

It is notable that the strategies presented in this paper are of relatively simple nature. While many of them are profitable in our backtests, we are aware that this is insufficient for a trader to start replicating as-is. The number of stocks in both the long and the short leg, as well as the monthly turnover[10] need to be considered as appropriate for the investor's risk and trading tolerances.

[9] See Fama (1991)

[10] On average, a quantile was composed of over 300 stocks, which means that 600+ stocks are involved in building a L/S portfolio. The monthly turnover was as high as 75% for some strategies.

Regardless, the main goal was to show that even relatively straightforward signals from timely supply chain data contain alpha, and that is the case.

Disclaimer

References

Bonacich, Phillip, and Paulette Lloyd. 2001. "Eigenvector-Like Measures of Centrality for Asymmetric Relations." *Social Networks* 23 (3): 191–201.

Fama, Eugene F. 1991. "Efficient Capital Markets: II." *The Journal of Finance* 46 (5): 1575–1617.

Zhao, George, Kevin Webster, CFA Yin Luo, Sheng Wang, David Elledge, Miguel-A Alvarez, and Allen Wang. 2015. "The Logistics of Supply Chain Alpha." *Deutsche Bank Market Research*.