BI Norwegian Business School - campus Oslo

GRA 19502

Master Thesis

Component of continuous assessment: Thesis Master of Science

Economically Linked Firms: An Opportunity for Abnormal Returns has Disappeared, but the Predictability Remain

Navn: Alexander Vestgård, Jørgen Hem Toftevaag

Start: 02.03.2018 09.00

Finish: 03.09.2018 12.00

Economically Linked Firms:

An Opportunity for Abnormal Returns has Disappeared, but the Predictability Remain

ALEXANDER VESTGÅRD and JØRGEN HEM TOFTEVAAG*
BI Norwegian Business School
Master of Science in Business
Major in Finance

ABSTRACT

In this master thesis, we investigate whether a customer-momentum strategy, which previously has been shown to yield high abnormal returns performs equally well today. We also examine if there is evidence for bi-directional Granger causality (predictability) between customer and supplier returns, by utilizing customer-supplier links in a panel-data-setup. We find that the abnormal returns from the customer-momentum strategy is no longer present in the market, although predictability remains. Further, we find evidence of increased trading activity for the stocks utilized in the portfolio and increased investor attention toward customer-supplier links. Our results are partly consistent towards market efficiency and consistent with greater investor attention towards strategies of abnormal returns, following the publication of a paper demonstrating the effectiveness of this strategy.

Submitted: 02.08.2018

* Both Alexander and Jørgen are students at BI Norwegian Business School, Master of Science in Business with a Major in Finance. We would like to thank our supervisor, Patrick Konermann at BI Norwegian Business School, for the extraordinary supervision and cooperation during the work on this thesis. We would also like to thank Lauren H. Cohen at Harvard Business School for providing his dataset and insight and Andrea Frazzini at NYU Stern for the dataset, code and help.

Tables of Contents

List of Graphs and Figures

List of Tables

1 Introduction

Cohen and Frazzini (2008) find 18.6% abnormal[1] annualized returns using a long-short portfolio strategy, from now on referred to as the customer-momentum strategy. We investigate whether there is evidence for these abnormal returns in the market today and for what Cohen and Frazzini argue is evidence of investor limited attention. In their paper, Cohen and Frazzini argue that these returns are the effect of predictability across economically linked firms. Specifically, in this setting, the links are between customers and suppliers, where announcements and news in one firm affect the other. Cohen and Frazzini argue that investors display limited attention towards these links, and therefore, news from one firm are not immediately reflected in economically linked firms. Specifically, bad (good) news from the customer has a negative (positive) effect on the supplier and the supplier's stock price does not immediately incorporate this. Therefore, the customer-momentum strategy consists of sorting monthly customer returns into quintiles and then buying (shorting) the top (bottom) quintile of corresponding suppliers. In an efficient market one should earn no abnormal returns using this strategy as prices should reflect all available information (Fama 1991; 1998). However, as it takes time for the good or bad customer news to be incorporated in the suppliers stock price, there is a predictable lagged effect in the stock market, which ultimately yields the aforementioned 18.6% annualized abnormal return. Figure 1 shows the cumulative value-weighted customer-momentum strategy versus the CRSP all-stocks value-weighted portfolio, for 1981-2004, which provides visual evidence that Cohen and Frazzini's strategy throughout the years beat the market. It is also worth noting that the CRSP returns represents an investment of own funds, while in the strategy, funds are obtained through shorting.

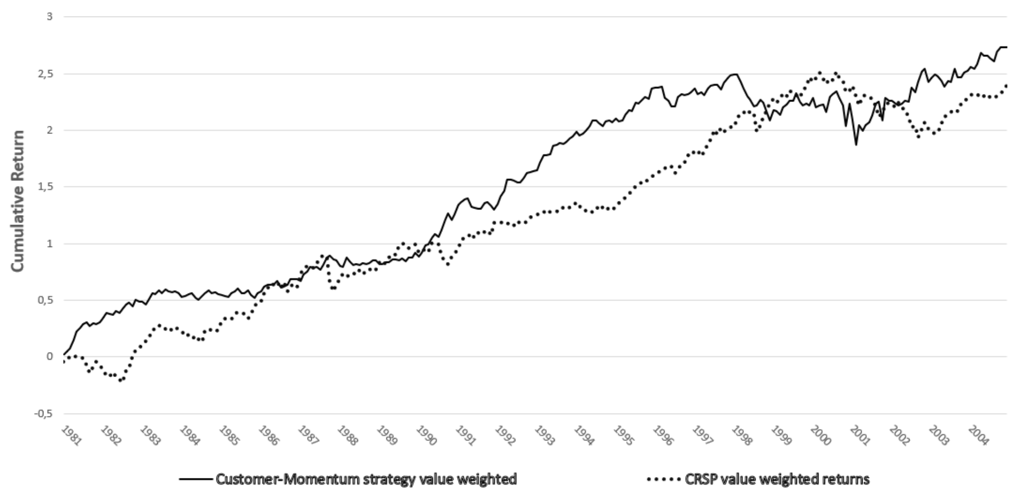

Figure 1: Cohen and Frazzini's Strategy versus CRSP all-stocks value weighted portfolio

Mclean and Pontiff (2016) find that abnormal returns have drastically declined or disappeared both post-sample and post-publication for long-short portfolios. However, they did not investigate Cohen and Frazzini's customer-momentum-strategy. We, therefore, argue it is interesting to investigate whether the high abnormal returns still exist for post-publication

[1] Abnormal return is here the alpha of a Fama-French 3-factor model. In general, throughout the thesis, abnormal return is defined as the alpha of a particular regression.

of “Economic Links and Predictable Returns” (2008), as this would be contradictory to both market efficiency and the results of Mclean and Pontiff (2016). Mclean and Pontiff (2016) argue “If return predictability reflects mispricing and publication leads sophisticated investors to learn about and trade against the mispricing, then we expect the returns associated with a predictor should disappear or at least decay after the paper is published.” This is further equal to increased investor attention towards strategies yielding abnormal returns, such as the customer-momentum strategy of Cohen and Frazzini (2008). We create an extended dataset using the same procedure as Cohen and Frazzini (2008), as this gives us the possibility to investigate post-sample and post-publication abnormal returns as done in Mclean and Pontiff (2016). Since value-weighted returns have the interesting interpretation of buy-and-hold returns and bad-model problems are more severe in inferences from equal-weighted returns (Fama, 1998), we mainly focus on the value-weighted returns. However, we also show the equal-weighted returns for the remainder of this thesis, as we argue that the evidence produced by our investigation regarding these returns provides additional insights.

In addition, we extend the paper by introducing panel Vector Auto Regressions (panel-VAR) to this setting as we treat the different customer-supplier relations as individual entities, which we argue provides new or additional evidence. As Fama (1998) argues, long-term return anomalies tend to disappear with reasonable changes in methodology. Additionally, using panel-VAR on stock returns of customers and suppliers could possibly explain more of the variation in the stock prices, as it is by nature quite free of specification and the data determines the relationship. Billio et al. (2012) postulates that in an informationally efficient financial market, short-term asset-price changes should not be related to other lagged variables and we, therefore, argue that the panel-VAR will provide evidence for or against market efficiency. Lastly, through the Granger Causality test[2] we can provide evidence for uni- or bi-directional predictive causality, which in turn could provide evidence for or against the Limited Attention Hypothesis, as it shows if customer returns Granger-causes supplier returns and vice versa. In addition, we are able to review how much of future variation in supplier (customer) returns can be explained by own returns and related firm returns when exposed to a shock directly or through the linked firm, which is possible through forecast error variance decomposition (FEVD). To our knowledge neither using panel-VAR in this setting nor reviewing Cohen and Frazzini’s results on an extended dataset has been done before.

The main results of our thesis are shown in chapter 5. Our findings are, firstly, that we find no evidence of abnormal returns in the period 2009-2017. This the is post-publication period as Cohen and Frazzini published in 2008. Further, Cohen and Frazzini have a cleaned version of their dataset online, for which we find similar results as they did for 1981-2004, however, with marginally lower returns. Additionally, Cohen and Frazzini provided us with a re-cleaned and extended dataset (1980-2009). In the period of 1999-2009 for this dataset we find significantly decreased abnormal returns, nonetheless, the returns are still economically and statistically significant. Further, we run Mclean and Pontiff regressions on the extended and re-cleaned data from Cohen and Frazzini from 1981-1997 and our dataset from 1998-2017. For the value-weighted returns, we find a statistically significant reduction for the post-sample and post-publication indicator, which might imply that the returns decrease or disappear due to statistical biases or information leakage post-sample and sophisticated

[2] Clive Granger (1969)

investors post-publication, as discussed in Mclean and Pontiff (2016). However, there is no evidence for reduced returns for the Mclean and Pontiff regression on equal-weighted returns, and we argue this could be due to investors focusing on the value-weighted returns, as this is the focus of Cohen and Frazzini's paper. We also find evidence for extraordinarily increased trading activity for the out of sample and post-publication periods, which also could explain the reduction in returns for the strategy.

Further, we investigate predictability in returns both upstream, a suppliers effect on customers, and downstream, a customers effect on suppliers, through vector autoregressions. We find that customer returns hold higher predictability towards supplier returns, however, suppliers also significantly predict customer returns. In addition, we check predictive causality (Granger-causality) of the customer-supplier relationships based on the vector autoregression results. The results from this test for the samples of customer-supplier relationships implies that there is bi-directional predictive causality between the customer and supplier, which backs the results from the VAR. Lastly, we utilize forecast error variance decomposition to review how much of future variation in returns can be explained by own returns and the downstream/upstream firm returns, which is motivated by the findings of Menzly and Ozbas (2010). The results are quite symmetric, i.e. supplier returns explain equally well future variation in customer returns as vice versa when the other firm is exposed to a shock. We also find evidence that customers and suppliers are more dependent on each other, as future variation in returns explained by the other firm is larger for the period of 2009-2017 relative to 1981-2004.

2 Related Literature

The following section is divided into three parts, where we first review the literature related to customer-supplier and economically linked firms. We then move to investor limited attention and lastly industry and production networks.

Concerning *customer-supplier links*, Cohen and Frazzini (2008) show that customer returns predict supplier returns, as investors do not immediately react to this information. They were the first to do this and there have been many researchers who have followed in their footsteps, as a starting point for their papers. Further, Müller (2017) finds that firms with the same stock-characteristics omit a higher cross-predictability than other firms. Pandit et al. (2011) find that suppliers experience information externalities at the time of their customer's earnings announcements, which could be used to predict returns. Lastly, Cen et al. (2017) shows that their information diffusion speed measure helps investors generate "sharper" customer-momentum strategies, regarding slow information diffusion.

All of the above-mentioned papers are related to *limited attention*. However, one of the first who modelled the concept of investor inattention was Merton (1987). Merton's model is based on investors who obtain information, and trade, on a small number of stocks. Stocks with fewer traders sell at a discount stemming from the inability to share risks, which in turn causes return predictability. Huberman and Regev (2001) study investor inattention to salient news about a firm. In their study, a firm's stock price soars on the re-release of information in the New York Times that had been published in Nature[3] 5 months earlier. This leads us towards investor limited attention, where Kahneman (1973) and Peng and Xiong (2005) concentrate on investors' learning behaviour given limited attention, and they find that attention is a scarce cognitive resource and attention to one task necessarily requires a substitution of cognitive resources from other tasks. Peng (2005) shows that information capacity constraints can cause a delay in asset price responses to news, which is what we are investigating in this thesis. Hong and Stein (2007) explains limited attention mechanisms more closely and Johnston and Pashler (1998) have a summary of the literature on attention limitations.

In the years after the publication of Cohen and Frazzini (2008), researchers within the field of return prediction from economic links seem to focus on inter-*industry networks* and not customer-supplier links. As Müller (2017) says, given the obvious economic links between firms within the same industry or along the industry supply chain, the current literature focus is not surprising. Menzly and Ozbas (2010) investigate market segmentation and finds that stocks in economically related supplier and customer industries cross-predict each others return and that the magnitude of return predictability declines with informed investors. Hong et al. (2007), Aobdia et al. (2014) and Rapach et al. (2015) look at inter-industry networks and cross-predictability of returns. Herskovic (2017) examines the asset pricing implications of input-output networks. Most of these papers use a similar dataset to the one we are using[4] and the main idea behind their research is the same, which stems from Cohen and Frazzini (2008), Menzly and Ozbas (2006) and Fama and French (1997).

[3] `https://www.nature.com`
[4] Gathered from COMPUSTAT/CRSP, while some uses BEA surveys

In addition to the customer-supplier links literature, our thesis also contributes to a broader literature on behavioural finance, asset pricing and market efficiency. We contribute to this literature by reviewing whether the abnormal returns are still there or not, by using an extended dataset to Cohen and Frazzini (2008), which is a check for if the market has removed the strategy's potential returns after it is published, or market efficiency (Mclean and Pontiff (2016); Fama (1998)). Further, the extension of implementing vector autoregressions presents a different method of investigating investor inattention towards customer-supplier relationships.

3 Data

We obtain a cleaned version of Cohen and Frazzini (2008)'s original data, from 1980-2004, from Andrea Frazzini's homepage.[5] Cohen and Frazzini also provide us with a dataset for 1979-2009, a version which is re-cleaned and extended. These datasets contain customer sales, names and CRSP (customer) PERMNO, date of relation, together with total supplier sales and CRSP (supplier) PERMNO. Monthly stock returns are found in the CRSP database and matched with the mentioned datasets by date and CRSP PERMNO number. We extend the dataset, in which we mimic Cohen and Frazzini (2008)'s data collection process. As our goals include to review whether the abnormal returns are still present today and reviewing investor limited attention in a different way, it is important that we do not change the data collection process relative to the datasets of which we are comparing results. If the datasets somehow systematically differ from each other it greatly weakens the results, as it could be a symptom of contrasting datasets and thereby possibly attributable to underlying differences. Hence, we follow all restrictions Cohen and Frazzini (2008) used. Our full data collection process is explained to a greater extent and in a more detailed manner in Appendix A, and we show summary statistics for overlapping years between our and Cohen and Frazzini' extended and re-cleaned dataset for robustness in Appendix B.

We extract customer-supplier relationships through the COMPUSTAT customer segment file for the time period 1998-2017. The reason we overlap with Cohen and Frazzini's extended version is to obtain robustness across datasets. Ideally, we would collect data for 1979-2017, however, before 1998 customer names were reported as abbreviations which greatly increases the time needed to build a sufficient and correct dataset, hence, we rely on Cohen and Frazzini's dataset for pre-1998 data. The COMPUSTAT customer segment file contains information about customers who represent more than 10% of a supplier's sales which suppliers report in their financial statements.[6] In practice, a firm can also voluntarily disclose customers that account for less than 10% of total revenues.[7]

Unfortunately, the customer segment file only reports the company name of the customer. As such, we match the company names with a dataset containing both CRSP PERMNO-number and company names through an algorithm which utilizes vectoral decomposition of text.[8] The algorithm generates a list of potential matches for each unique customer company name, for which we manually check non-exact matches to review whether it is the same company or not. Further, as we specify the algorithm to use tokens (whole words) and report only matches above a certain score, we manually control and review, for a second time, the entire dataset of 300,000 [9] observations and conservatively adjust if segment, public information, industry and name suggest a match is warranted. This procedure is similar to Cohen and Frazzini's hand-matching of pre-1998 relations. For suppliers, we extract PERMNO-number through matching by GVkey, as this is included in the COMPUSTAT customer segment file. Further, we only include companies with non-missing book

[5] `http://people.stern.nyu.edu/afrazzin/data_library.htm`

[6] Statement of Financial Accounting Standards (SFAS) No. 14 and No. 131 states that suppliers must report all customers with more than 10% of total sales.

[7] I.e. an unknown supplier might gain credibility by voluntarily disclosing a well-known company as a customer as a form of signalling.

[8] Matchit algorithm in Stata, written by Julio D. Raffo

[9] Many customers are reported as ”Not Reported” or contains regions, countries etc, as such the process becomes slightly less tedious and a lot of observations are removed

equity and market equity at fiscal year-end, a restriction Cohen and Frazzini also imposed. As the re-cleaned and extended data of Cohen and Frazzini was not cleaned for non-missing values of book and market values, we have done this process on both our own and their re-cleaned set. In accordance with Cohen and Frazzini's method, we impose a 6-month gap between fiscal year end of the accounting data and stock prices to ensure that investors are aware of the customer-supplier relation. Further, Cohen and Frazzini focus their analysis exclusively on common stocks.[10] We find that share code assignment has not changed since Cohen and Frazzini's publication, and therefore, only common stocks with share codes 10 or 11 are included for the analyses. Share codes, returns and price are gathered from CRSP through SAS-Studio which is linked to WRDS. Andrea Frazzini provided us with their SAS-code which matches financial accounting data with the mentioned datasets containing the customer-supplier relations with the unique PERMNO-numbers. The final sample is based on 25,867 customer-supplier relations representing 7,272 unique customer-supplier relations between 1998-2017. After the six-month restriction from the reported date of a customer-supplier relation is imposed, the monthly returns for one year are gathered.[11] It is from the collection of these returns we provide summary statistics on. Whereas the construction of portfolios and VAR have the additional restriction of common stocks only and a price above five dollars. These specific steps are based on the SAS-code provided by Andrea Frazzini, hence we are very confident that the construction and following output of summary statistics and abnormal returns are the same as in their paper and thus comparable.

Table 1 show the main summary statistics of our thesis. Panel A consists of the sample coverage in relation to the CRSP common stock universe. As Panel B shows, the suppliers mimic the regular size distribution of the CRSP common stock universe. In contrast, the customers in our sample are tilted towards large cap companies. This is a characteristic due to the reporting from which we base our sample on. A larger company is naturally more likely to represent a greater portion of a suppliers sales. Furthermore, we have on average covered 63% of the value in the CRSP universe with the customers representing the majority. In terms of percentage of companies covered the suppliers hold the majority, while total sample coverage is at 23%. On average, customer and suppliers are 80% of the time in different industries[12], and therefore, our analysis is mostly based on return predictability between firms not in the same industry. Finally, the average number of customers per supplier varies greatly with a mean of 1.82.

Our dataset is highly comparable to the original dataset of Cohen and Frazzini (2008) even though we are mostly covering a different time period. There are a couple of interesting differences however, the average link duration has increased from 2.7 to 3.2, implying that customers and suppliers are in a relationship for a longer time. As firms are linked for a longer period, we argue it should be easier for investors to obtain information about these links, which then should decrease investor limited attention. Additionally, there is a reduced percentage of sales to customers, which was 19.8%, while it for our dataset have dropped to 16.54%. Further, when reviewing our dataset we find that the percentage is decreasing at a steady rate from 20% in 1999 to around 11% in 2017, implying that the average customer

[10] Share code 10 and 11

[11] this is the reason we compare datasets for 2000-2009, as some observed customer-supplier relations from 1997 will give return observations in 1999 creating a bad comparison

[12] We assign stocks to 49 industries based on their SIC code. The industry definitions are based on Fama and French (1997) and are obtained from CRSP

Table 1: Summary Statistics: 2000-2017

This table shows the summary statistics for
our dataset for the years 2000-2017.
"Stock universe" is all stocks in CRSP.
Link duration is number of years the firms are connected without breaks.
Same industry is based on the Fama & French industry definitions
Size percentiles is based on the size
of a customer or supplier in regards to the CRSP stock universe.
Percentage of sales to customer is the average sales
from a supplier the customers in the dataset count for.

	Min	Max	Mean	SD	Median
Panel A: Time Series (Annual Observations)					
Number of firms	714	1077	905	117	920
Number of customers	401	497	441	28	434
Full sample % coverage of stock universe (EW)	17.1	25.4	22.7	2.1	23.5
Full sample % coverage of stock universe (VW)	56.8	67.8	63.3	2.9	63.4
Supplier % coverage of stock universe (EW)	14.6	20.0	18.7	1.4	19.1
Supplier % coverage of stock universe (VW)	9.8	19.1	15.6	2.2	16.2
Customer % coverage of stock universe (EW)	5.4	12.9	9.3	1.8	9.4
Customer % coverage of stock universe (VW)	52.0	62.2	57.9	2.6	58.0
% of customer-supplier in the same industry	18.0	21.7	20.0	1.2	20.4
Link duration (Years)	1.0	18.0	3.2	3.1	2.0
Panel B: Firms (Pooled Firm-Year Observations)					
Supplier size percentile	0.00	0.99	0.51	0.28	0.52
Customer size percentile	0.01	0.99	0.81	0.20	0.89
Number of customers per firm	1.00	23.00	1.82	1.69	1.00
Percentage of sales to customer	0.00	100	16.38	15.93	13.00

is less important to a supplier, which further should reduce the importance of a shock from the customer to the supplier. One could speculate if this is due to increased globalization as it will intuitively lead to a wider customer base or/and increased supplier competition as customer demand would then be spread among more suppliers. It is also worth to point out that while the median customer remains around the 90th percentile of the CRSP stock universe, the customer is on average in the 80th percentile, down from the average of 90th percentile in Cohen and Frazzini's dataset. Lastly, our dataset[13] contains an average of 905

[13] We also show summary statistics for robustness across datasets and for the relevant sub-samples used in the thesis, in Appendix B.

yearly observations for suppliers and 441 customers, compared to Cohen and Frazzini's average of 918 yearly observations for suppliers and 433 for customers in their paper.

As discussed earlier, it is essential that our data collection procedure closely mimics that of Cohen and Frazzini (2008). Therefore, the following section provides summary statistics for the overlapping years of 2000-2009. We choose to compare our dataset to the re-cleaned and extended version provided by Andrea Frazzini. As it has been cleaned two times and is extended it is arguably the best version and it allows for comparison for a longer period of time. Figure 2 shows the yearly number of customers and suppliers in our dataset and the re-cleaned dataset of Cohen and Frazzini. This provides evidence that our data collection method is sufficient, as the number of observations for both suppliers and customers is relatively similar and quite stable throughout the period. The difference steadily shrinks throughout the period, which is likely a symptom of better reporting standards which enable us to more easily match customer names with accounting and financial data. I.e. spending additional time on cleaning and finding additional observations for the dataset becomes futile due to improved reporting. We also provide summary statistics for both our dataset and Cohen and Frazzini's re-cleaned version in Table A.9, table A.8 and Table A.7 in Appendix B. The tables indicate that we cover somewhat less of the CRSP universe in terms of total companies and value, which is consistent with the fewer observations, while the customer and supplier specific information are close to identical in all regards, which supports the notion of our collection method mimicking that of Cohen and Frazzini (2008). In addition Table A.10 can be found in Appendix B, it is based on our observations from 1998-2017 and Cohen and Frazzini (2008)'s observations from 1980-1997 and it is this dataset which is used in the Mclean and Pontiff regressions.

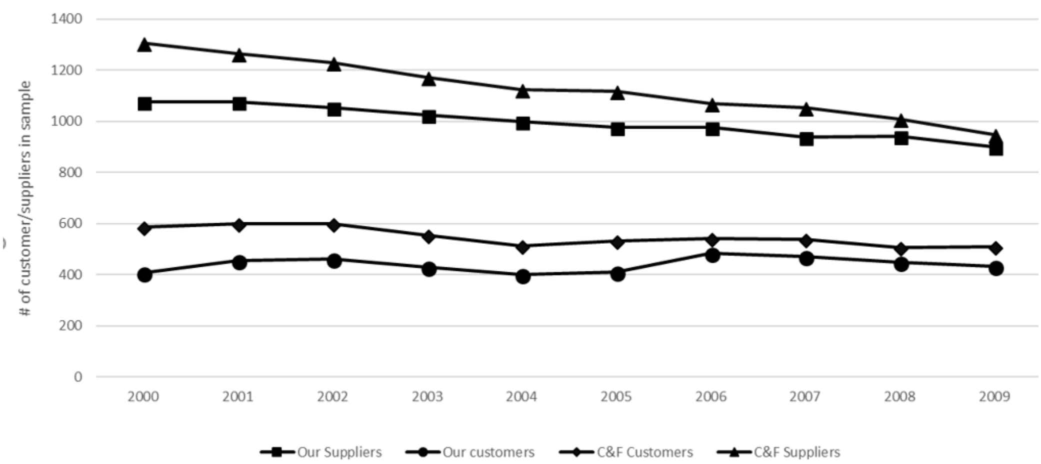

Figure 2: Yearly number of customer and supplier observations

4 Methodology

The following section starts with an introduction of the limited attention hypothesis. Further, we postulate three hypotheses regarding the customer-momentum strategy, predictability and bi-directional predictive causality. Following postulation, we describe our method of implementation and testing in the same order. Finally, we introduce some additional investigations which are interesting to review.

The basic theory and intuition behind Cohen and Frazzini (2008)'s paper is based on the fact that when two firms are economically linked, actions or announcements in one firm, should affect the other. For instance, if the customers share price drops due to negative news regarding future prospects, it is natural to assume this will alter the customers current demand towards goods provided by suppliers. This, in turn, should reduce the suppliers share price as they have been negatively affected by their customer. Additionally, the decline in share price for the supplier should by intuition be related to the percentage of sales they have to that specific customer, something which Pandit et al. (2011) finds evidence of. Which again, should happen effectively right after the customers share price drop if the markets are efficient. However, as shown by Cohen and Frazzini (2008), this has not been the case. The supplier's stock price does not instantly incorporate the information about their customers, which in turn generates return predictability, as Figure 3 shows through the supplier Coastcast and their customer Callaway who represented 50% of their sales.

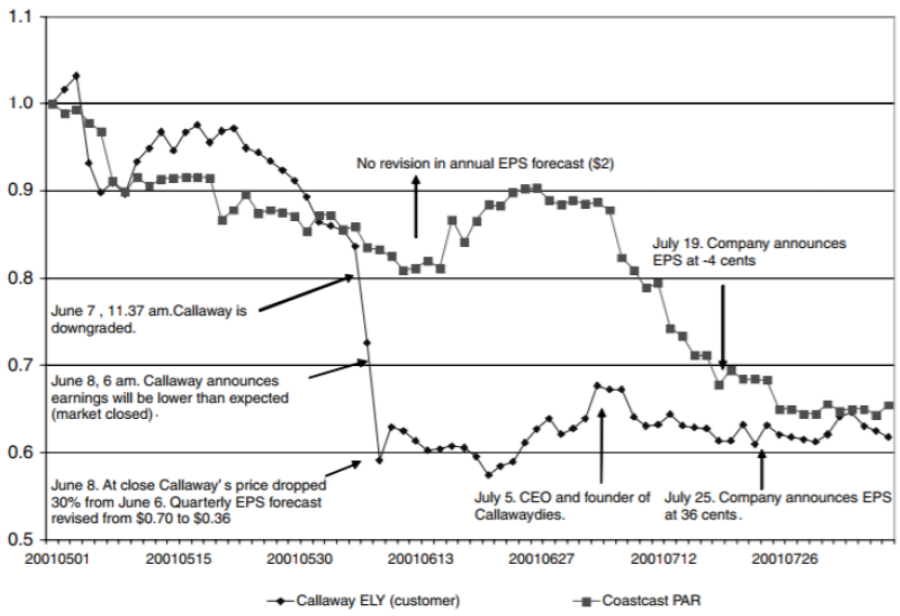

Figure 1. Coastcast Corporation and Callaway Golf Corporation. This figure plots the stock prices of Coastcast Corporation (ticker = PAR) and Callaway Golf Corporation (ticker = ELY) between May and August 2001. Prices are normalized (05/01/2001 = 1).

Figure 3: Coastcast and Callaway example

Furthermore, direct bilateral exposures between financial institutions and other market participants can be used to investigate connectivity, which is the case between suppliers and customers (Chan-Lau (2017); Diebold and Yilmaz (2014)). Specifically, in this setting, we argue it is reasonable to expect customers affect their suppliers to a greater degree than vice versa. As mentioned, the data contains information about customers who can be regarded as major stakeholders for the suppliers, whereas the importance of the supplier to the customer is not clearly defined in the same way. Nonetheless, we argue it is likely and quite intuitive that negative (positive) news should impact the related firm in a negative (positive) way. In general, we argue the impact should follow the degree of dependence between the related firms.

Further, this brings us towards the limited attention hypothesis. The limited attention hypothesis is based on underreaction in stock prices regarding firm-specific information that induces changes in the valuation of related firms, generating return predictability across assets. In particular, stock prices underreact to negative (positive) news involving related firms, and in turn generate negative (positive) subsequent price drift (Cohen and Frazzini, 2008). Hence, in the presence of investors subjected to attention constraints, stock prices do not rightly incorporate news about related firms, and thereby generate stock price predictability across assets. The aforementioned arguments are based on immediate news incorporation being correct. One could argue that the lagged effect is due to a lagged cash flow effect. In essence, it takes time for the decline in a firm's cash flow to affect the related firm. While this can certainly be true and is likely, it does not the change how an investor should respond to the news, whether the lagged cash flow effect happens today or in six months, negative or positive news should be incorporated in the prevailing price.

In relation to the *customer-momentum strategy*, consider The Efficient Market Hypothesis (Fama (1991); Fama (1998)) states that it should not be possible to beat the market over time and that anomalies should disappear over time. Fama (1998) argues that "Consistent with the market efficiency hypothesis that the anomalies are chance results, apparent overreaction to information is about as common as underreaction, and post-event continuation of pre-event abnormal returns is about as frequent as post-event reversal". A zero-cost strategy with an annualized abnormal return of 18.6% percent for the period between 1981 and 2004 is not in line with the statement above. In essence, the success of the strategy is effectively a result of temporary mispricing in the market, due to investor limited attention. We also put the strategy in the setting of Mclean and Pontiff who argues that sophisticated investors will swarm towards opportunities of abnormal return. As such, hypothesis (i) postulates that the abnormal returns of Cohen and Frazzini (2008)'s customer-momentum strategy is not present in the market today, considering the findings of Mclean and Pontiff (2016) and Fama (1991).

Hypothesis (i)

H_0 = Customer-momentum strategy generates no abnormal returns

H_A = Customer-momentum strategy generates abnormal returns

Further, we want to check if the *predictability* found in Cohen and Frazzini (2008) is still present, by utilizing an econometric method that has not been used in this setting before. In the event where the abnormal returns of the zero-cost strategy have disappeared, one cannot state whether the return predictability between linked firms has disappeared. Hypothesis (ii) is therefore solely towards the notion of predictability. Through using panel-VAR, we are able to investigate if there exists predictability between customer and supplier returns in a different way. This puts hypothesis (ii) in the setting of Fama (1998), who argues that most anomalies should disappear with the introduction of new or other econometric techniques. Furthermore, the results for this hypothesis will also provide additional evidence for or against the limited attention hypothesis. Additionally, according to market efficiency, past stock values should hold no predictive power, therefore, the results will also provide additional evidence for or against market efficiency, as argued by Billio et al. (2012).

Hypothesis (ii)
H_0 = No predictability between customer and supplier returns
H_A = Predictability between customer and supplier returns

Additionally, we postulate hypothesis (iii) that there is *bi-directional predictive causality* and not exclusively uni-directional predictive causality from the customer to the supplier. This hypothesis is motivated in part by Menzly and Ozbas (2010), who finds that stocks that are in economically related supplier and customer industries cross-predict each others returns. As such, we argue it is interesting to investigate whether the same relationship exist among firms. As the dataset by construction is tilted toward customers being relatively more important for suppliers than vice versa, Cohen and Frazzini constructs a measure of "important supplier" and find evidence that only the lagged return of the important suppliers have predictive power towards customer returns when running monthly cross-sectional regressions. Thus, they find evidence for important suppliers to have predictive power towards customer returns. Our method is somewhat similar, however, we utilize panel-VAR to review whether we can find bi-directional causality among customer-supplier relations, without imposing a specific term towards which suppliers are regarded as important. We argue that using Granger Causality test can provide evidence for which entity Granger-causes the other, and in so doing, we are also able to provide evidence for or against bi-directional predictability in a different manner than showed by Cohen and Frazzini.

Hypothesis (iii)
H_0 = Uni-directional predictive causality from customer to supplier
H_A = Bi-directional predictive causality between customer and supplier

In the next section, we will touch upon how we are going to *implement and test* these hypotheses, as well as possible interpretations of these investigations.

We follow Cohen and Frazzini's customer-momentum strategy. In which, each month customers are assigned to five different quintiles according to last month's return in an ascending order. We go short (long) in the respective quintile of suppliers whose customers are in the bottom (top) quintile. The supplier stocks in these quintiles are then equal- or value-weighted[14]. As earlier discussed, a strategy such as this should earn zero abnormal returns in an efficient market, as it is based on information available to all investors[15] (Fama, 1991). As such, the equation below is utilized to find evidence for abnormal returns, where the intercept a_i is the abnormal return. R_{it} is the portfolio return for the equal or value-weighted customer-momentum strategy. We first run the regression with only market return (MKT), which is the market model. Further, we then include the Small Minus Big (SMB) and High Minus Low (HML) factors, which is then the Fama and French's 3-factor

[14] Cohen and Frazzini (2008) show that there are no differences between equal-weighted and value-weighted portfolios

[15] Lagged returns and customer-supplier links

model. Finally, we include Carhart's momentum factor (UMD), which is an extension of the Fama and French 3-factor model, known as the Carhart 4-factor model. The controls of the customer-momentum portfolios through Fama and French 3-factors (Fama and French, 1993) and Carhart's momentum factor (Carhart, 1997)[16] are necessary steps in order to conclude if the portfolios yield abnormal returns. As the returns of the customer-momentum portfolio could be related to the aforementioned factors, the inclusion of them allows us to interpret the abnormal returns as the effect of the customer-momentum strategy, given that there are no other underlying factors related to the customer-momentum strategy. The included factors follow the methodology of Cohen and Frazzini (2008).

$$R_{it} = \alpha_i + \beta_1\, MKT_{i,t} + \beta_2\, SMB_{i,t} + \beta_3\, HML_{i,t} + \beta_4\, UMD_{i,t} + e_{it} \qquad (1)$$

First, we investigate whether there is evidence for abnormal returns in the market today. We, therefore, run the aforementioned strategies and collect monthly returns for the period 2009-2017, we then review whether there exist abnormal returns through the regression above. Although Cohen and Frazzini review the strategy until 2004, we choose to begin in 2009 as there is a possibility that the returns between 2004 and 2009 are influenced by information leakage to investors and/or statistical biases, which we will explain shortly, and a shift in 2009 due to publication. Finally, we are then able to conclude with regard to hypothesis (i). If the alpha of equation (1) is statistically significantly different from zero, we have evidence for abnormal returns and we reject the null. In the case where we do not find significant results, we do not reject the null and conclude that the customer-momentum strategy no longer generates abnormal returns. Furthermore, as Mclean and Pontiff (2016) find evidence for declining returns in long-short portfolios following publication, we investigate the customer-momentum strategy following publication of Cohen and Frazzini (2008). Hence, we run a linear regression throughout the in-sample, post-sample and post-publication period with the portfolio returns as the dependent variable with a dummy for the post-publication and post-sample periods on the right-hand side, as done in Mclean and Pontiff (2016). This enables us to check if there is a significant negative change in portfolio returns post-sample and post-publication represented by the coefficient of the respective dummy variables. The post-sample period is between the end of data-sample and publication, which is 2005 and June 2008. The reason for the post-sample indicator is to test if the return predictability in a published study results from statistical biases, as the predictability then should disappear out of sample (Mclean and Pontiff, 2016). Additionally, investors could learn about the predictor while the study is still a working paper, as such the coefficient is interpreted as a combination of the two. For post-publication returns, Mclean and Pontiff's findings are consistent with the idea that academic research draws attention to return predictors. As mentioned, they argue that sophisticated investors swarm towards strategies that yield abnormal returns, and thus, returns decline after publication. However, Mclean and Pontiff (2016) make use of a large panel of predictors, which mitigates the issue of coinciding factors affecting the dummies, as they review the regression-results on an aggregate level. However, we argue that the results from our regression regarding the customer-momentum returns and possible changes are still interesting to look at as a way to review which periods of time affects the results of the customer-momentum strategy, however, we must be careful in regards to causal interpretations.

[16] We obtain the Fama-French factors and Carhart's momentum factor from WRDS.

$$R_{it} = \alpha_i + \beta_1 Post\ Sample\ Dummy_{i,t} + \beta_2\ Post\ Publication\ Dummy_{i,t} + e_{it} \quad (2)$$

In addition to the linear regression on the returns, we test if the portfolio returns are affected by a change in trading activity dynamics. We construct the measures of trading activity following Mclean and Pontiff (2016), which are measured as the logged trading volume of shares traded and logged dollar value of shares traded. By checking for trading activity dynamics we are able to assert to a greater degree whether investors change their behaviour following publication through increased trading of suppliers included in the customer-momentum strategy.

The aforementioned methodology focuses on the customer-momentum strategy and abnormal returns. However, as obtaining abnormal returns tells a story of mispricing and predictability in the market, the lack of abnormal returns from a pre-specified strategy does not necessarily tell the opposing story of unpredictability and market efficiency. Therefore, we now switch our focus to the notion of predictability. Hence, we relieve ourselves of portfolio specification and investigate the customer-supplier relation through a Vector Autoregression Model. In a VAR, all variables are endogenous, as such, the data itself determines whether or not there is evidence of predictability through past values. Our sample and setting is somewhat special, and as such, we argue our data can be treated as panel-data i.e all consecutive customer-supplier relations are treated as individual entities. Through the panel-VAR model, we can review the average explanatory power of lagged returns and investigate bi-directional causality through a Granger causality test. Further, if we were to impose the traditional VAR on the relationships, all customer-supplier relations would be treated as homogeneous relationships. In other words, we would implicitly specify lagged returns of the linked firms to impact each other returns identically. We argue there are strong reasons to reject this notion. The intuition behind this thesis and that of Cohen and Frazzini are based on economically linked firms and the strength of these links. It follows that the predictive power of lagged returns is likely influenced by the strength of the relationship and thus varies throughout the different customer-supplier relationship. Hence, we treat each customer-supplier relation on a pairwise basis, this allows us to model our VAR with a fixed term, which in practice allows for heterogeneous effects. In other words, the customer and supplier are allowed to affect each other to a varying degree for each pairwise link. We argue this is a more realistic assumption than homogeneous relationships.

The exact relationship between related companies are naturally varying and quite complex, i.e it would not be unreasonable to create a more complex model to include effects above and beyond that of heterogeneous effects. Schweitzer et al. (2009) provides an insightful description of the challenges of financial network modelling: “In the complex-network context, ‘links’ are not binary (existing or not existing), but are weighted according to the economic interaction under consideration.... Furthermore, links represent traded volumes, invested capital, and so on, and their weight can change over time”. However, we argue that in our context of reviewing if there exists a predictive relationship, the somewhat binary nature of our setup should be sufficient to capture this. More specifically, we argue that a model which considers time-varying degree of dependence and other factors which influences the degree of inter-connectivity might be reasonable additions. However, we argue our simpler set-up should be sufficient to answer whether cross-predictability between supplier and customer returns exists. Based on the discussion above, we use the following model of a k-variate panel of order 1 with panel-specific fixed effects.

$$y_{it} = \beta_1\, y_{i,t-1} + \beta_2\, x_{i,t-1} + f_i + e_{it} \tag{3}$$

Above, t denotes time and i index the relation of k total pairwise customer-supplier relations. y_{it} is a (1xk) vector of dependent variables (either customer or supplier return). β_1 is a (kxk) matrix of coefficients to be estimated for the dependents lagged return $y_{i,t-1}$. β_2 is a (kxk) matrix of coefficients to be estimated for the lagged return of the related firms return, where $x_{i,t-1}$ represents the related firms lagged return. e_{it} is the (1xk) vector of error terms of the individual regressions.

Without adjustments, the fixed term f_i is correlated with the lagged regressors on the right-hand side, which left unaddressed will cause the coefficients to be biased and inconsistent. To solve this problem, we use a "Helmert transformation", as done in Love and Zicchino (2006) and explained in Arellano and Bover (1995). The procedure transforms the lags by removing the mean of future observation. This preserves the orthogonality between transformed variables and lagged regressors, and the lagged regressors can then be used as instruments and the coefficients can be estimated by system GMM.[17]

The customer-momentum portfolio makes use of the predictability of last month's customer return towards their supplier. As such, for some of the questions we are trying to answer, a PVAR(1) would be sufficient. Nonetheless, as VAR in the traditional setting and panel setting is a-theoretical by nature it is common practice to use information criteria to decide which model best fits the data, as such we base the lag length on the coefficient of determination (CD) or R-squared. In the VAR-setting, one typically uses the information criteria of multivariate AIC, BIC and HQIC[18] to decide the number of lags. The moment model selection criteria developed by Andrews and Lu (2001) allows for estimation of MAIC, MBIC and MQIC. As in VAR-modelling, the model which minimizes MAIC, MBIC and MQIC is the preferred one. However, Andrews and Lu base their selection criteria on Hansen's J Statistic (Hansen, 1982) which requires the number of moment conditions to be greater than the number of endogenous variables in the model Love and Zicchino (2006). As such, we would need to specify the number of lags used as instruments to be higher than the number of lags we are trying to estimate, which would be a model of the "over-identification" type. As one of our intentions by utilizing VAR is to keep the model as simple as possible, we stick to the simpler lag-selection criteria of CD to find the most parsimonious model. Therefore, to best model the predictive power of past returns, we opt for a PVAR(1) as this produces among the highest coefficients of determination (CD). When reviewing higher lag orders the CD increases with only around 0.05 for each new lag added. Therefore, we argue that a PVAR(1) is the most parsimonious.

The model will enable us to answer hypothesis (ii) regarding predictability between customer and supplier returns. The interpretation is straightforward if customer or supplier lagged returns are statistically significant towards the related firm, then we have evidence for predictability and reject the null of no predictability, which again would be evidence against market efficiency.

[17] Our model is "just-identified", meaning it has the same number of regressors as instruments. Therefore, system GMM is numerically equivalent to equation-by-equation 2SLS. (Love and Zicchino, 2006)

[18] Multivariate Akaike, (Schwarz)-Bayesian and Hannan-Quinn Information Criteria are the most commonly used in VAR's (Brooks, 2015)

Further, an interesting test to run after the Panel-VAR is the Granger Causality test, to investigate if there is a uni- or bi-directional predictive causality between customers and suppliers (Granger-causation), as Diebold and Yilmaz (2014) argue this is a way to investigate pairwise directionality. Holtz-Eakin et al. (1988) comment on the importance of estimating the appropriate lag-length prior to causality testing, which we do, in short panels, as no inferences concerning causality could be drawn in absence of such tests. Further, Billio et al. (2012) argue that in the presence of costs of gathering and processing information, there may be Granger causality among price changes of financial assets and that these returns are hard to arbitrage away because of the said frictions, which would imply that we cannot reject hypothesis (iii). The Granger Causality test does not account for confounding effects, such as macroeconomic inference, and further, it does not consider non-linear relationships, as the vector autoregression considers linear relationships (Brooks, 2015). Further, Diebold and Yilmaz (2014) argue that Granger causality tests are appealing, as there is no need for identifying assumptions. The results from the Granger causality test will, in turn, enable us to answer hypothesis (iii), as well as providing additional evidence for hypothesis (ii). Towards hypothesis (iii), the null of uni-directional predictive causality is rejected in the case of Granger-causality from both customer and supplier towards the related firm, we will then have evidence for bi-directional predictive causality. Further, the coefficients of the VAR will tell a story of the degree of predictive power the customer and supplier have towards each other.

Additionally, we propose to check if there are any differences between sub-samples grouped by the percentage of sales customers contribute to their suppliers. We split the sample into above and below median groups in regards to sales percentage, where median is based on the average sale percentage between the different relations. The intuition is that one would expect lagged customer returns to hold greater predictive power towards the above-median group of suppliers. Further, as an additional check for connectivity between customers and suppliers, we argue that investigating Forecast Error Variance Decomposition (FEVD (Brooks, 2015)) results of the panel-VAR, will provide evidence for how shocks in one firm influences the variation in returns of the other part of the link in future months. It is particularly interesting to view how this may have changed between time periods, as greater connectivity between firms should be accompanied by a decline in investor inattention as the customer-supplier link can be viewed as relatively more important.

5 Results and Analysis

We start this chapter by examining the outcome of the customer-momentum strategy for the years 2009 to 2017. We then investigate the strategy in a Mclean and Pontiff setting. The section then moves towards the Panel-VAR setting for periods 1981-2004 and 2009-2017 and concludes with some additional results regarding FEVD.

Table 2: Customer-Momentum Strategy, Abnormal Returns: 2009-2017

In this table we show the abnormal returns from running the Cohen and Frazzini strategy for 2009-2017. The results comes from Equation 1 and adjustments of it. Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *. The explanatory variables are the Fama and French 3-factors, as well as Carharts Momentum-variable. t-value in parentheses.

Panel A: Value-Weights	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	L/S
Excess returns	0.180	-0.012	-0.097	0.271	0.029	-0.151
	[0.64]	[-0.06]	[-0.37]	[1.16]	[0.12]	[-0.40]
3-factor alpha	0.141	-0.028	-0.191	0.227	0.036	-0.104
	[0.50]	[-0.13]	[-0.72]	[0.95]	[0.15]	[-0.27]
4-factor alpha	0.146	-0.031	-0.195	0.223	0.032	-0.116
	[0.52]	[-0.15]	[-0.74]	[0.94]	[0.13]	[-0.30]
Panel B: Equal-Weights	**Q1 (Low)**	**Q2**	**Q3**	**Q4**	**Q5 (High)**	**L/S**
Excess returns	-0.275	-0.107	-0.289	-0.081	0.124	0.399
	[-1.02]	[-0.44]	[-1.20]	[-0.33]	[0.51]	[1.45]
3-factor alpha	-0.123	-0.025	-0.146	0.079	0.250	0.373
	[-0.61]	[-0.16]	[-0.79]	[0.43]	[1.48]	[1.33]
4-factor alpha	-0.123	-0.023	-0.155	0.067	0.243	0.366
	[-0.60]	[-0.15]	[-0.94]	[0.43]	[1.52]	[1.33]

As is evident from Table 2, neither value- nor equal-weighted customer-momentum strategies yield statistically significant abnormal returns for the market model, the 3-factor model or the 4-factor model. In regards to economic significance, the equal-weighted portfolio does yield 0.37% monthly abnormal return when controlling with the Fama and French 3-factor model, which we deem as slightly economically significant.[19] However, it is important to note that in Cohen and Frazzini's original results this portfolio yielded monthly abnormal returns of 1.3%[20], as such, the decline is quite dramatic. Furthermore, the decline for the value-weighted portfolio is even larger, as the customer-momentum portfolio has in recent years yielded negative monthly abnormal returns of 0.1% compared to a positive return of 1.55% in Cohen and Frazzini's paper.[21] In general, value-weighted results are more important than equal-weighted, due to the buy-and-hold interpretation (Fama, 1998). We argue that the low or no abnormal returns, together with the lack of statistical significance,

[19] A total of 4.44% annualized abnormal returns - (0.37*12)

[20] For the years 81-04 for the cleaned and re-cleaned versions of data, this portfolio yields abnormal returns of 1.25% and 0.99% respectively, as shown in Appendix B

[21] For the years 1981-2004 for the cleaned and re-cleaned versions of data, this portfolio yield abnormal returns of 1.48% and 0.96% respectively, as shown in Appendix B

is in line with the efficient market theory, as a long-short portfolio is theorized to yield no abnormal returns in an efficient market (Fama, 1991). In addition, the results are possibly consistent with Mclean and Pontiff (2016) who find that returns decrease after publication, a subject we will return to shortly. Further, the results are quite clear in regard to hypothesis (i) of the customer-momentum strategy yielding abnormal returns, for which we do not reject the null. Which means, that there are no abnormal returns to be found from this strategy in the market today, for either the equal- or value-weighted customer-momentum strategy. We believe this is a very reasonable conclusion, as the comparison of our dataset with that of Cohen and Frazzini exhibits very similar sample characteristics. Furthermore, in terms of methodology, the construction of monthly returns and implementation of regressions are based on code provided by Andrea Frazzini, which greatly strengthens our argument that our results are directly comparable to that of Cohen and Frazzini and thus differences in results are caused by market pricing and subsequent performance of the strategy.

Furthermore, it is interesting to note that Cohen and Frazzini (2008) found alphas to be rising monotonically across quintiles, as the portfolio goes from low to high returns. I.e, high (low) customer returns are followed by high (low) supplier returns. The returns of the customer-momentum strategy were also asymmetric, as they were largely driven by slow diffusion of negative news. Something which exhibits a pattern consistent with short-sale constraints according to Cohen and Frazzini (2008). This monotonic relationship is also found in their cleaned and re-cleaned data in Appendix B. However, the asymmetric relationship towards the quintile of low returns are only found in the cleaned version, as the re-cleaned version exhibits customer-momentum returns largely driven by asymmetry towards the high quintile.[22] Our results, on the other hand, show neither a rising monotonic relationship in the alphas nor any clear asymmetry towards the low or high quintile. This is consistent with our conclusion and results towards the customer-momentum strategy, i.e. that there is no longer a clear pattern of slow diffusion for any of the quintiles, with the reasonable interpretation that investors seemingly act immediately to firm-related news, unlike the period of 1981-2004.

The results of the *Mclean and Pontiff regressions* are shown below in Table 3. As we have now shown the disappearance of the customer-momentum strategy's abnormal returns, we review how the portfolio returns are affected in the post-sample and -publication period. We run the regressions in the period of 1981-2017[23] as explained in chapter 4.b, methodology. We utilize the entire period to attain an in-sample period similar to the average of Mclean and Pontiff (2016). For the value-weighted returns, the post-sample returns are 1.4% lower, while the post-publication returns are 0.9% lower each month, both of which are significant at the 10% level. While this is in line with what Mclean and Pontiff (2016) finds towards statistical biases and/or leakages post-sample and decaying returns post-publication, we must be careful in interpreting our results as evidence of this. The dummies essentially show a significant decline in these periods, however, this could be caused by other effects as well.[24] We, therefore, investigate how the trading activity has

[22] If one follows the reasoning of Cohen and Frazzini in regard to short-sale constraints, which we agree with, and as the difference in results are exclusively driven by the data at hand, we argue that the re-cleaned version consists of fewer companies subject to short-sale constraints.

[23] The dataset is a combination of Cohen and Frazzini's data from 1980-1997, while we use our data from 1998-2017. As argued in chapter 3, data, the datasets are largely similar, which justifies the combination to attain a larger sample which is also evident from Table A.10

[24] E.g. both dummies includes years containing the recent financial crisis, a period where most

changed for the value-weighted portfolio, to provide additional evidence whether there is a sample and publication effect. As we find no significance for the dummies regarding equal-weighted returns, a puzzling question which we discuss later, we argue it would be inappropriate to further examine these returns.

Table 3: McLean and Pontiff Regressions on returns: 1981-2017

In this table, we test for a reduction in the abnormal returns from Cohen and Frazzini's strategy. We run Equation 2 to obtain the results. Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *. t-values in parantheses.

Variables	Value-Weighted Returns	Equal-Weighted Returns
Post-sample	**-0.014***	-0.005
dummy	**[-1.89]**	[-1.02]
Post-publication	**-0.009***	-0.003
dummy	**[-1.72]**	[-0.76]
Constant	0.010	0.009
	[3.61]	[4.47]
Observations	444	444
R-squared	0.013	0.003

Table 4: McLean and Pontiff regressions on Trading Activity Dynamics for robustness: 1981-2017

The regressors are Trading Volume, which is measured as shares traded, and Dollar Volume, which is measured as shares traded multiplied by price. The Equation is similar to Equation 2, with dynamics instead of returns as regressors. Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *. t-values in parentheses.

Variables	Trading Volume	Dollar Volume
Post-sample	**2.22*** **	**2.24*** **
dummy	**[12.95]**	**[11.96]**
Post-publication	**2.31*** **	**2.50*** **
dummy	**[19.77]**	**[19.53]**
Observations	444	444

We construct variables of logs of monthly averages of trading volume and dollar volume of all the stocks in the long and short side of the customer-momentum portfolio. We then run regressions on these variables on the left-hand side, equal to equation 2. The results in Table 4 show that both trading and dollar volume increase drastically, both post-sample and post-publication. As the dependent variables are in logs, the coefficients imply an increase in percent for each of the post-indicators. Our results show that trading and dollar volume increase with 222% and 224% post-sample. For the post-publication period, trading and

trading strategies would arguably suffer.

dollar volume increases by 231% and 250% respectively for the stocks in the customer-momentum portfolio. However, the dollar volume in North America increased by 82%[25] and the trading volume increased by 164%[26] during the post-sample period. This is likely to influence our results, which reduces our ability to interpret them, given that the portfolio stocks are representative of the general market, which we argue is a quite realistic assumption.[27] However, as the significant increase in trading activity is above that of the general market for the post-sample period, we argue it is weak evidence of increased trading activity in portfolio stocks in the post-sample period, which could be due to statistical biases and/or leakages, as argued by Mclean and Pontiff (2016). In regard to the post-publication period, the dollar volume in North America increased by 15%[28] and the trading volume decreased by 48% [29]. We, therefore, regard the issue of noise from the total market in our regression towards the post-publication dummy as quite small. Hence, we have some evidence for Mclean and Pontiff (2016)'s arguments that publication enlightened investors about the abnormal returns of the customer-momentum strategy, which leads to the following disappearance of these returns.

On a clarifying note, we have now argued that there is some evidence for increased attention towards customer-supplier links following publication. However, there is no reason to assume that increased trading and dollar volume would or should be limited to the quintiles of the portfolio alone. I.e. that attention to the customer-supplier link will only be given if returns are large or small. Hence, we review the quintiles not included in the portfolio in Table A.6 in Appendix B, which shows similar results, that trading both in terms of total and dollar values have had a large increase for companies regarded as suppliers, which supports our notion.

As mentioned earlier, we will now briefly discuss why there is no evidence of reduced equal-weighted returns post-sample and -publication, along with some concluding remarks. As the equal-weighted returns do not show a statistically significant decline, neither for the post-sample, nor post-publication indicator, we argue it weakens the argument that publication caused a decline and that the abnormal returns are driven by statistical biases. Further, it is difficult to clearly state a reason why the value-weighted returns are more affected than the equal-weighted return. Possible reasons are that the value-weighted returns initially generated larger abnormal returns and as this was the focus of Cohen and Frazzini (2008) it might have generated more attention. In regards to our overall assessment of whether abnormal returns have disappeared due to publication or biases/leakages, there is some evidence and arguments that the disappearance of the abnormal returns is also driven by a general decline. When running the customer-momentum strategy for the period 81-04, 00-09 and 09-17, for both equal- and value-weighted portfolios, the abnormal returns are steadily declining. Furthermore, it has in the period of 1981-2017 become arguably easier to both obtain and analyze data, which by intuition should "create" more sophisticated investors, which leave fewer opportunities for abnormal returns. Thus, we argue it is likely that the customer-momentum strategy has experienced a general decline throughout time, combined with our evidence of increased trading post-publication we argue attention

[25] (48.75/26.25) - 1. Numbers in $ trillion (TheWorldBank, 2018a)

[26] (383.27/145.13) - 1. Numbers in $ trillion (TheWorldBank, 2018b)

[27] Our restrictions towards price above five dollars and common stocks somewhat weakens this assumption

[28] (41.07/35.64) - 1. Numbers in $ trillion (TheWorldBank, 2018a)

[29] (111.83/213.85) - 1. Numbers in $ trillion (TheWorldBank, 2018b)

towards customer-supplier links have increased throughout the years and that the publication of Cohen and Frazzini contributed to this effect. Furthermore, as the dummies exhibit very similar coefficients and together represent the period of 2005-2017, the dummies could show an increase in trading which is not necessarily attributable to publication or leakage. This logic also supports a general rise in attention towards customer-supplier links. The only explicit reason we have come up with, beyond increased attention to the customer-supplier link, regards investors favouring more frequent trading of firms which do not sell to the end-consumer, as these companies are by nature not in our sample, however, we deem this line of reasoning implausible.

We now move to the results and analysis in the *panel-VAR setting*, in order to review the customer-supplier relation in a new light. All analysis is based on the same returns used to form the customer-momentum strategy. However, while the customer-momentum strategy focuses on supplier returns based on the extreme quintiles of monthly customer returns, we now analyze all applicable data[30] from both customers and suppliers. Further, where there are gaps in the customer-supplier relations, we treat the link as a new one after the gap. Gaps are induced by either missing return for a month[31], due to a ended customer-supplier relation returning after a year/years or because the price drops below five dollars. Further, we impose a 6-month gap to these relations, to ensure that investors are aware of the customer-supplier links. For this data, we also impose the 5$ trading price restriction to the suppliers and customers, as this is done in Cohen and Frazzini (2008) and in their online Appendix, to relieve us from the possibility that the predictability is driven by micro capitalization illiquid stocks.

Table 5: Panel Vector Autoregression on Customer-Supplier links: 1981-2004

We utilize C&F's data in this table.
All customer-supplier links are treated as Panel-data,
where we run Equation 3 to obtain these results.
Coefficients are multiplied by 100.
Statistical significance at the 1, 5 and 10% level is indicated
by a ***, ** and *. z-values in parentheses.

Variables	Customer Returns	Supplier Returns
Customer Returns t-1	**-0.7****	**6.7*****
	[-2.15]	**[14.83]**
Supplier Returns t-1	**2.1*****	**1.5*****
	[10.79]	**[4.10]**
Average # of Months of the links (panels)	20	20
# of Observations	210026	210026
# of Panels	10469	10469

In Table 5, we show the Panel-VAR results for Cohen and Frazzini's re-cleaned dataset from 1981-2004. As both Cohen and Frazzini (2008) and we find statistically significant

30 By applicable data, we mean monthly returns that follow the same assumptions which are used in portfolio formation.

31 This could have been corrected for by using Holtz-Eakin et al. (1988)'s method of setting missing observations to 0, however, we argue our large sample does not warrant this method

evidence for return predictability from the customer-momentum strategy in this period, we review if the Panel-VAR arrives at the same conclusion. As we find significant predictability for the customers first lag towards suppliers, we argue that this is evidence for the predictability Cohen and Frazzini finds. Further, we also find significant evidence of predictability stemming from the suppliers first lag towards the customer. Additionally, one would in the presence of limited attention towards customer-supplier links expect the past returns of suppliers to hold less predictive power towards customer as the dependence by intuition should be lower, which is what we find. It is also worth noting that Cohen and Frazzini (2008) in their online Appendix run a cross-sectional regression and find significant evidence that lagged supplier returns have predictive power over customer returns with a coefficient of 1.7-2.0. Although this is towards their measure of important suppliers, the results are strikingly similar. In conclusion, we reject the null of no predictability between customer and supplier returns of hypothesis (ii), for the period 1981-2004, since the lags are significant. Thereby, our results are consistent with the performance of the customer-momentum strategy. Lastly, it also seems to be a small, but significant predictability for both suppliers and customers based on their own lag, indicating a momentum-effect, however, as this is not the topic of our thesis we do not investigate this further.

Table 6: Granger Causality Test on Customer-Supplier links: 1981-2004

In this table, we report P-values of the Granger causality test. The test is done on the panel-VAR results from Cohen and Frazzinis data, where the dependent variables are customer and supplier returns.

	Caused Variables	
	Customer Returns	Supplier Returns
Customer Returns		**0.000**
Supplier Returns	**0.000**	
Degress of freedom	1	1

Next, we use the Panel-VAR results to run a Granger causality test to test for predictive causality of the customer-supplier relationships. Firstly, as above, we review Cohen and Frazzini's re-cleaned dataset, in the time period of 1981-2004, as shown in Table 6. We find significant evidence of bi-directional predictive causality between customers and suppliers. This implies that we reject the null hypothesis of uni-directional predictive causality of Hypothesis (iii), for 1981-2004. As we find bi-directional predictive causation, we argue that these results provide additional evidence for the limited attention hypothesis, as Cohen and Frazzini (2008) finds evidence for predictability in their online Appendix as well. Both results regarding hypothesis (ii) and (iii) for the period of 1981-2004 are, in our setting, evidence against the idea that new econometric methods remove return anomalies, as Fama (1998) argues.

Further, we extend the Panel-VAR analysis to our dataset, for the period from 2009-2017. We show these results in Table 7. The results show that there is still strongly significant predictability for the customers first lag towards the supplier. One implication of this, as we find no significance for abnormal returns in the customer-momentum strategy, is that the market is only partly efficient, which is what Fama (1991) postulates. The market has arbitraged away the abnormal returns, but the predictability is still present, but lower. This could also be evidence for Mclean and Pontiff (2016)'s argument that publication of return-predictability papers make sophisticated investors swarm towards the strategy, and thereby arbitraging the returns away. However, for the lagged supplier returns, which we find significant evidence for towards the customer, the coefficient is now negative in contrast to the period of 81-04 in Table 5, which is difficult to interpret in a coherent manner. However, towards our hypothesis (ii) the results are nonetheless clear for the period of 2009-2017 as well, we reject the null of no predictability between customer and supplier returns.

Table 7: Panel Vector Autoregression on Customer-Supplier links: 2009-2017

We utilize our data in this table.
All customer-supplier links are treated as Panel-data,
where we run Equation 3 to obtain these results.
Coefficients are multiplied by 100 for readability.
Statistical significance at the 1, 5 and 10% level is indicated
by a ***, ** and *. z-values in parentheses.

Variables	Customer Returns	Supplier Returns
Customer Returns t-1	0.3	**5.6*** **
	[0.54]	**[8.37]**
Supplier Returns t-1	**-0.8*** **	**-3.5*** **
	[-2.8]	**[-6.41]**
Average # of Months of the links (panels)	27	27
# of Observations	71459	71459
# of Panels	2639	2639

Table 8: Granger Causality Test on Customer-Supplier links: 2009-2017

In this table, we report P-values of the Granger causality test. The test is
done on the panel-VAR results from our data,
where the dependent variables are customer and supplier returns.

	Caused Variables	
	Customer Returns	Supplier Returns
Customer Returns		**0.000**
Supplier Returns	**0.005**	
Degress of freedom	1	1

Furthermore, in Table 8 we show the Granger Causality test results for our dataset. As in the period of 1981-2004, we find evidence of bi-directional causality. We reject the null hypothesis of uni-directional predictive causality of hypothesis (iii). Again, this is evidence for investor limited attention.

Next, we split the customer-supplier sample into above and below median percentage of total sales to a customer, to investigate differences in predictability for different magnitudes of firm linkages. We argue that this measure should imply an increase in predictability, as both the supplier and customer is more reliant on the other part of the link when the percentage is higher. We find evidence that the above median group has a higher predictability for

1981-2004, which is shown in Table A.11 in Appendix B. This is especially for customers first lag towards the supplier. For our extended dataset of 2009-2017, shown in Table A.12 in Appendix B, we find evidence for lower predictability for the above median firms from the customers first lag towards the supplier, than for the below median firms. This is extremely puzzling, and we have no clear intuition or answer to why. As such, we are unable to draw any conclusions from this. A better measure for degree of dependence in the relationship is likely needed in order to find evidence for or against this notion.

Our final investigation towards the customer-supplier links is to utilize *forecast error variance decomposition (FEVD)*. In Table A.13 we show the results for the Cohen and Frazzini (2008) re-cleaned dataset from 1981-2004. We review an initial shock to the supplier on the left side and an initial shock to the customer on the right side of Table A.13. We review both shocks following the logic of our presented thoughts towards bi-directional dependence. Through the FEVD, we are able to examine how much of future variation in the returns are explained by own returns and the corresponding customer/supplier when exposed to a shock. As evident from the table, around 6% of the variation in future return is explained by the supplier (customer) return, but only if the shock is delivered to the customer (supplier). Hence, there seems to exist a quite symmetric relationship in terms of how future variation in returns is explained when exposed to a shock from the related company. Further, when reviewing the period of 2009-2017 in Table A.14 we find the same relationship, however, it is a bit stronger, as close to 9% of the variation in future returns can be explained by the related company when it is exposed to a shock. We argue that this increased explanatory effect can be interpreted as an increased connectivity between customers and suppliers. It is then natural to expect the link to be more apparent, and thus, priced in the market. Following this line of thought, one would expect investor limited attention towards customer-supplier links to decline when shocks towards related companies explain more of future variation in returns. Thus, we argue that these results are in line with the other aspects of our thesis, which exhibit that investors have increased their attention towards customer-supplier links, namely that abnormal returns have disappeared due to greater attention along with a lower predictability from past returns of related firms.

6 Conclusion

In this thesis, we investigate the effectiveness of Cohen and Frazzini (2008)'s customer-momentum strategy, for which they originally found an annualized abnormal return of 18.6%. To investigate if there exist abnormal returns from this strategy today, we construct an extended dataset by the same means as Cohen and Frazzini. In addition, we investigate the topic of predictability in customer-supplier links. We find no evidence of abnormal returns in the market today, which answers hypothesis (i). We then utilize the methodology of Mclean and Pontiff (2016) to investigate possible reasons to why the abnormal returns disappeared, for which we find a drastic rise in trading activity post-publication. Through our panel-VAR setup, we investigate predictability and find strong evidence for predictability between the customer and supplier, which answers hypothesis (ii). Lastly, we run Granger causality tests and find evidence of bi-directional predictive causality, providing additional evidence that there is predictability between customers and suppliers. This concludes our thesis by answering hypothesis (iii).

As the customer-momentum strategy does not yield any abnormal returns in the market today, but we nevertheless find evidence for predictability, we argue this is partly evidence for market efficiency (Fama, 1991). In the case of the customer-momentum strategy, perhaps investors incorporated a strategy shown to yield abnormal returns, and arbitraged these away, however, without fully incorporating the information in the market, thus leaving some predictability on the table. We are not able to draw clear conclusions by using Mclean and Pontiff (2016)'s methodology, as the results could be somewhat weakened by coinciding factors. However, we believe the presented results are strong enough to argue for reduced returns being partially influenced by publication. Together with the gradual decline of the customer-momentum strategy, slight decline in predictability, increased connectivity and increased trading we conclude that investors are more attentive towards customer-supplier links, which we believe is a combination of publication and general decline of investor limited attention towards customer-supplier links.

Future interesting research could be to extend our use of panel-VAR to investigate if some groups of customers and suppliers hold greater predictive power, as this could vary between industries, degree of dependence and competitiveness among suppliers/customers. As we find evidence for predictability, we argue there could be other trading strategies that could earn abnormal returns, perhaps a strategy motivated by the aforementioned topic of greater predictability among some groups. As a concluding remark, Compustat is currently working on a linking suite product which will provide the missing link between customer names and GVkey. Hence, future datasets will be much more easily attained, as the need for hand-matching is removed.

Bibliography

Andrews, D. W. and Lu, B. (2001). Consistent model and moment selection procedures for gmm estimation with application to dynamic panel data models. *Journal of Econometrics*, 101(1):123–164.

Aobdia, D., Caskey, J., and Ozel, N. B. (2014). Inter-industry network structure, information transfers, and the cross-predictability of earnings and stock returns. *Review of Accounting Studies*, 19(3):1191–1224.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1):29–51.

Billio, M., Getmansky, M., Lo, A., and Pelizzon, L. (2012). Econometric measures of systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104(3):535–559.

Brooks, C. (2015). *Introductory Econometrics for Finance*. Cambridge University Press.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57.

Cen, L., Hertzel, M. G., and Schiller, C. M. (2017). Speed matters: Limited attention and supply-chain information diffusion. *Working Paper*.

Chan-Lau, J. (2017). Variance decomposition networks: Potential pitfalls and a simple solution. *IMF Working Papers*, 17(107):1.

Cohen, L. and Frazzini, A. (2008). Economic links and predictable returns. *The Journal of Finance*, 63(4):1977–2011.

Diebold, F. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1).

Fama, E. F. (1991). Efficient capital markets: Ii. *The Journal of Finance*, 46(5):1575.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49(3).

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (1997). Industry costs of equity. *Journal of Financial Economics*, 43(2):153–193.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):31–47.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029.

Herskovic, B. (2017). Networks in production: Asset pricing implications. *Journal of Finance*, Forthcoming.

Holtz-Eakin, D., Newey, W., and Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica*, 56(6):1371–1395.

Hong, H. and Stein, J. C. (2007). Disagreement and the stock market. *Journal of Economic Perspectives*, 21(2):109–128.

Hong, H., Torous, W., and Valkanov, R. (2007). Do industries lead stock markets? *Journal of Financial Economics*, 83(2):367–396.

Huberman, G. and Regev, T. (2001). Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *The Journal of Finance*, 56(1):387–396.

Johnston, J. and Pashler, H. (1998). Attentional limitations in dual-task performance. *Psychology Press*.

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Love, I. and Zicchino, L. (2006). Financial development and dynamic investment behavior: Evidence from panel var. *The Quarterly Review of Economics and Finance*, 46(2):190–210.

Mclean, R. D. and Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1):5–32.

Menzly, L. and Ozbas, O. (2006). Cross-industry momentum. *Working Paper*.

Menzly, L. and Ozbas, O. (2010). Market segmentation and cross-predictability of returns. *The Journal of Finance*, 65(4):1555–1580.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, 42(3):483.

Müller, S. (2017). Economic links and cross-predictability of stock returns: Evidence from characteristic-based "styles". *Review of Finance*.

Pandit, S., Wasley, C. E., and Zach, T. (2011). Information externalities along the supply chain: The economic determinants of suppliers' stock price reaction to their customers' earnings announcements. *Contemporary Accounting Research*, 28(4).

Peng, L. (2005). Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis*, 40(02):307.

Peng, L. and Xiong, W. (2005). Investor attention: Overconfidence and category learning. *Journal of Financial Economics*, 80(3):563–602.

Rapach, D., Strauss, J., Tu, J., and Zhou, G. (2015). Industry interdependencies and cross-industry return predictability. *Research Collection Lee Kong Chian School Of Business*.

Schweitzer, F., Fagiolo, G., Sornette, D., White, D. R., Vespignani, A., and Vega-Redondo, F. (2009). Economic networks: What do we know and what do we need to know? *Science*, 325(5939):422–425.

TheWorldBank (2018a). *Stocks traded, total value (current US$)*. `https://data.worldbank.org/indicator/CM.MKT.TRAD.CD?end=2017&locations=XU-US&start=1980` [Accessed: 230518].

TheWorldBank (2018b). *Stocks traded, turnover ratio*. `https://data.worldbank.org/indicator/CM.MKT.TRNR?end=2017&locations=XU-US&start=1980` [Accessed: 230518].

Appendix A

Data Construction

This appendix is meant to provide a complete and detailed walkthrough of how we collect and proceed to create our dataset. The CRSP/Compustat name of variables are in parentheses after a variable name.

First, we extract data on customer-supplier relations from WRDS- Compustat historical customer segment in the period 1998.01-2018.04. We choose company code ”GVKEY”, such that we immediately have a linkage between Compustat and CRSP in terms of suppliers. Further, we extract the variables Company Name (conm) and Standard Industry Classification Code (SIC), which are for the suppliers. Additionally, the variables Customername (CNMS) and Customer Sales (SALECS) are also extracted, which are for the customers. We then remove the characters ” . , ’ from all Customernames and replace - / and \” with space and set all the customer names in uppercase letters. At a later stage we must match these company names with another dataset containing both company names and PERMNO number, therefore, these edits are made to best facilitate this matching process, as suppliers report their customers in altering ways.[I]The matching of customer names is a tedious and time-consuming process, hence, we first want to remove any observations that cannot be used due to missing information or due to non-matches in regard to suppliers.

We initially drop 2,767 of 375,345 observations due to either missing customer name, supplier GVKEY or Date of customer-supplier relation. We then move forward with matching the suppliers to their PERMNO numbers. As we initially extracted GVKEY, we use this to match the companies with their PERMNO numbers, instead of company names. We link supplier GVKEY and Supplier PERMNO number from CRSP - CRSP/COMPUSTAT merged database- Linking Table. We match the suppliers by exporting the GVKEY from our set, inserting these in a text file and utilizing WRDS to do the rest by applying these as our company code, instead of searching the entire database. By searching the entire database one would receive a different GVKEY format, 001004 instead of 1004, and therefore, providing WRDS with a text file containing the 1004-format avoids this issue. We choose the linking options LU, LC and LS, as Compustat recommends selecting LINKTYPES LC, LU, and LS for the same results. As these represent the vast majority of the links between CRSP securities and Compustat companies, without introducing duplicate data. Further, we gather the variables PERMNO (Historical CRSP PERMNO Link to COMPUSTAT Record) and ”Last and First Effective Date of Link”. Some companies have the same GVKEY, but differing PERMNO through time. Therefore, we utilize the first effective date of link variable and generate unique ’PERMNO sets’ rising by date. This dataset is then merged with the original Customer-Supplier dataset. From thereon, we are able to identify which PERMNO set correctly matches with the supplier at a given point in time by matching the reporting date of the customer-supplier relation with the date interval between ’PERMNO

[I] E.g. Mc Donald’s vs mc-donalds

sets' together with the GVKEY. We drop 31,692 observations for which we can not find a PERMNO match for the supplier.

The next step is to match the customer company names with their corresponding PERMNO-number. First, we gather data from CRSP/Compustat Merged database - Fundamentals annual, from the period 1997.01-2017.12 (last available date). Further, we select company code GVKEY so we are able to gather accounting data at a later stage. Again, we choose the linking options LC, LU and LS. Finally, we extract the variables SIC, Company Name and PERMNO (Historical CRSP PERMNO Link to COMPUSTAT Record). The Company Name variable is stripped of the same unwanted characteristics described above, towards the names of customers. Then we match the variable company name with the customers in the customer-supplier set. Even though we have cleaned the company names to a certain degree by removing some characters, the customer names are reported in different ways (COCA COLA CORPORATION vs COCACOLA CRP). As such, a procedure relying on exact name matches will result in a substantial loss of observations. Therefore, we utilize the user-written "matchit" function in Stata. This function compares text strings and outputs a similarity score between 0 and 1 with 1 being an exact match. This function allows for a broad range of vectoral decompositions of texts, or in other words, ways of comparing texts. Through testing different settings, we found tokens, which compares whole words, to be best in the setting of company names. Cohen and Frazzini (2008) used a similar method for pre-1998 observations, using a Soundex algorithm which compares names by sound, however, according to Cohen and Frazzini most company names where in this period written as abbreviations. Therefore, when we use a similar Soundex algorithm we obtain a lot of false positives as we impose the algorithm on complete names. Therefore, we move forward with comparing whole words. We specify the function to emphasize unique words/names trough "w(simple)", which effectively reduces the scores weighting of frequent words such as corp, inc, enterprise etc. and enhance the weighting of rare words. The effect and reasoning for doing so is to be able to review potential matches where frequent words such as the difference between "crp" and "corporation" would otherwise exclude or produce a low score. Further, it also gives a higher score to rare company name matches which are by intuition more probable.

The final specification we make is to only review potential matches with a score above 0.9245. We set this cut-off as we do not see a lot of matches below this score. This leads to removal of some correct matches, however, we correct for this at a later stage. The reasoning for this cut-off is to reduce manual labour, as the possible combinations with the cut-off are 10,000+ and when one merges the matches with the complete set and reviews this set, it is easily visible which matches are missed. This process also corrects for the matches for which the token setting is not ideal for, such as the example of COCA COLA CORPORATION vs COCACOLA CRP, where the first name consists of three tokens and the latter consists of two, which results in a low score. The set of possible matches are then manually reviewed. We are conservative with matching names in this process so we do not match a customer with a wrong firms stock return and accounting data. Matches which are initially ambiguous towards a match/non-match are reviewed through investigating the suppliers SIC code and public information. Through public information we can review whether the potential, but not exact match is caused by a company name change, such as switching between "inc" and "corp", or whether it is a subsidiary, sister or holding company. We characterize these related companies as "affiliates", but we do not include them in our final dataset. The reason for this is that affiliate companies are likely to have a varying effect on the matched company return due to ownership structure, payout policies and general cooperation between the entities. As we have identified the customers with their respective PERMNO number, we are able to extract SIC code and GVkey. We collect this data from merged CRSP/Compustat linking table. The process above results in a total of 52,031 observations. The drop of around 250,000 might seem like a lot, however, many observations have customers reported as ”Not Reported”, countries, regions and so on.

The final restriction we have towards our sample, following Cohen and Frazzini (2008), we only include observation for companies with non-missing values of book and market equity at fiscal year-end. We utilize the SAS-code written by Denys Glushkov at WRDS research applications[II] to collect market and book values. We only utilize the first stages of the code, as these are the only stages relevant to us. Further, we match the book and market values towards our sample of customer and supplier firms. Fiscal year-end for suppliers is simply the same reported date as when first extracting the data from the customer segment file. Naturally, fiscal year-end for the customer is not always the same as the supplier. As such, we match book- and market values towards the customer on the following fiscal year-end if it is not the same. Around 25,000 observations are dropped due to missing market and/or book values, as such the final sample consists of 25,867 customer-supplier relations from which we can gather returns.

[II] https://wrds-web.wharton.upenn.edu/wrds/research/applications/sas_files/market_to_book.sas

Appendix B

Robustness and Controls

For cross-referencing and robustness between datasets and reuslts

In this Appendix, we show tables of summary statistics and the customer-momentum strategy portfolio results for the extended dataset we have constructed, Cohen and Frazzini's original (cleaned) and Cohen and Frazzini's re-cleaned, and slightly extended dataset. We do this for robustness between datasets and to provide evidence that our data collection process (Appendix A) is adequate. Further, we show additional robustness and controls for our other results.

We start by showing the customer-momentum portfolio returns. For Table A.1, A.2 and A.3, alpha is the intercept on a regression of monthly excess return from the rolling customer-momentum strategy.

Table A.1: Customer-Momentum Strategy, Abnormal Returns: C&F's cleaned data 1981-2004

In this table we show the abnormal returns from running the Cohen and Frazzini strategy for 1981-2004. The results comes from Equation 1 and adjustments of it. Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *. The explanatory variables are the Fama and French 3-factors, as well as Carharts Momentum-variable. t-values in parentheses.

Panel A: Value-Weights	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	L/S
Excess returns	-0.832***	-0.366	-0.126	0.309	0.662**	1.493***
	[-2.95]	[-1.53]	[-0.51]	[1.12]	[1.98]	[3.59]
3-factor alpha	-0.610**	-0.272	-0.128	0.413	0.869***	1.479***
	[-2.21]	[-1.17]	[-0.50]	[1.43]	[2.92]	[3.46]
4-factor alpha	-0.383	-0.240	-0.055	0.458	0.915***	1.298***
	[-1.39]	[-1.01]	[-0.21]	[1.57]	[3.00]	[2.98]
Panel B: Equal-Weights	**Q1 (Low)**	**Q2**	**Q3**	**Q4**	**Q5 (High)**	**L/S**
Excess returns	-0.744**	-0.134	0.165	0.379*	0.622**	1.366***
	[-2.58]	[-0.61]	[0.71]	[1.73]	[2.44]	[5.14]
3-factor alpha	-0.663***	-0.215	0.027	0.371**	0.583***	1.246***
	[-3.12]	[-1.28]	[0.15]	[2.23]	[3.03]	[4.51]
4-factor alpha	-0.457**	-0.067	0.192	0.438**	0.739***	1.196***
	[-2.17]	[-0.40]	[1.08]	[2.58]	[3.85]	[4.23]

Table A.2: Customer-Momentum Strategy: C&F's re-cleaned data 1981-2004

In this table we show the abnormal returns
from running the Cohen and Frazzini strategy for 1981-2004.
The results comes from Equation 1 and adjustments of it.
Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *.
The explanatory variables are the Fama and French 3-factors,
as well as Carharts Momentum-variable.
t-values in parentheses.

Panel A: Value-Weights	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	L/S
Excess returns	-0.381	0.303*	0.167	0.399*	0.658***	1.039***
	[-1.56]	[1.75]	[0.96]	[1.71]	[2.88]	[3.40]
3-factor alpha	-0.129	0.338*	0.297*	0.530**	0.826***	0.955***
	[-0.55]	[1.96]	[1.70]	[2.19]	[3.64]	[2.99]
4-factor alpha	0.088	0.409**	0.317*	0.510**	0.928***	0.836**
	[0.38]	[2.32]	[1.77]	[2.05]	[4.01]	[2.57]
Panel B: Equal-Weights	**Q1 (Low)**	**Q2**	**Q3**	**Q4**	**Q5 (High)**	**L/S**
Excess returns	-0.261	0.157	0.374**	0.568***	0.739***	1.001***
	[-1.10]	[0.79]	[2.00]	[3.14]	[3.22]	[4.89]
3-factor alpha	-0.223	0.103	0.307***	0.448***	0.769***	0.992***
	[-1.37]	[0.85]	[2.63]	[3.86]	[5.36]	[4.62]
4-factor alpha	0.016	0.185	0.431***	0.492***	0.864***	0.840**
	[0.11]	[1.51]	[3.77]	[4.16]	[5.97]	[2.58]

In Table A.1, the results from Cohen and Frazzini's cleaned dataset[III] is shown. These results (and data used) are very similar to the ones used in their original paper, which we find evidence for when comparing it to the tables in their paper. Table A.2 shows the re-cleaned dataset from Cohen and Frazzini, for the same overlapping years as Table A.1. The main results, L/S, have kept the same significance levels in the recleaned dataset, however, the coefficients have dropped, implying an economically significant decrease. The most notifiable change is for quintile 1, the short-position quintile, which in the old set, Table A.1, is greatly significant, while for the re-cleaned set there is no significance.

For Table A.3, we use the re-cleaned dataset of Cohen and Frazzini. As seen in Table A.3, both the significance and the coefficients have decreased from Table A.2, which can be a symptom of decreasing return predictability from 1981-2004 to 2000-2009, or that there are some specifics that affect the returns like the IT-bubble and the financial crisis that reduces the returns significantly in 2000-2009, but not enough to affect the significance in 1981-2004.

In Table A.4 and A.5 we split the post-publication dummy into two new dummies, where one is for the first half and the other is for the second half of the original post-publication dummy. The argument is based on the logic of investor limited attention, where attention is a scarce cognitive resource and the attention to one task requires a substitution of another. It then follows, that there exists the possibility of a switch between strategies

[III] This is the dataset from Andrea Frazzini's homepage

Table A.3: Customer-Momentum Strategy: Overlapping Years - Cohen and Frazzini's 2.0 re-cleaned data 2000-2009

In this table we show the abnormal returns
from running the Cohen and Frazzini strategy for 1981-2004.
The results comes from Equation 1 and adjustments of it.
Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *.
The explanatory variables are the Fama and French 3-factors,
as well as Carharts Momentum-variable.
t-values in parentheses.

Panel A: Value-Weights	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	L/S
Excess returns	-0.269	0.301	0.333	0.082	0.513	0.783
	[-0.55]	[1.04]	[1.02]	[0.19]	[1.21]	[1.27]
3-factor alpha	-0.353	0.184	0.353	0.114	0.521	0.874*
	[-0.77]	[0.65]	[1.11]	[0.26]	[1.28]	[1.39]
4-factor alpha	-0.322	0.203	0.361	0.084	0.539	0.862*
	[-0.70]	[0.71]	[1.14]	[0.19]	[1.32]	[1.36]
Panel B: Equal-Weights	**Q1 (Low)**	**Q2**	**Q3**	**Q4**	**Q5 (High)**	**L/S**
Excess returns	0.339	0.503	0.748**	0.950***	0.907**	0.568
	[0.82]	[1.56]	[2.42]	[3.01]	[2.33]	[1.42]
3-factor alpha	-0.027	0.061	0.305	0.474**	0.469*	0.496
	[-0.09]	[0.32]	[1.58]	[2.25]	[1.94]	[1.21]
4-factor alpha	0.016	0.073	0.327*	0.483**	0.485**	0.468
	[0.06]	[0.39]	[1.73]	[2.29]	[2.01]	[1.15]

of abnormal returns, where resources are drawn to the highest return at all times. This could lead to time-varying returns of the portfolio, which the dummies could then capture. The switch of attention is arbitrarily set midway, but we argue it is interesting to review nonetheless.

In Table A.4 and A.5, we show the extended results from the Mclean and Pontiff regressions shown in Table 3. For regression 1 and 2 in Table A.4, there is significant evidence for post-sample decay in the returns, which we discuss under Results and Analysis. When splitting the post-publication dummy into two separate dummies, we find no evidence for a significant reduction in the returns for the ”Late dummy” or the ”Early dummy”. Further, in Table A.5, which shows Mclean and Pontiff's regressions for the equal-weighted returns, we find no evidence for a reduction.

In Figure A.1 we show the customer-momentum portfolio return compared to the CRSP all-stocks value- and equal-weighted indices. This is shown as a robustness to the Mclean and Pontiff (2016) regressions. We provide this graph as a visual inspection of the cumulative returns of our strategy, compared to value- and equal-weighted indices of all CRSP stocks, to review the performance of the strategy relative to the performance of CRSP stocks. One can clearly see that the returns from our portfolios had the highest slopes until approximately year 2000, and then they are overtaken by the CRSP indices indicating a general decline in performance.

Further, in Table A.6 we show trading activity dynamics for the stocks that are not

Table A.4: McLean and Pontiff regressions on Value-weighted returns: 1981-2017

In this table, we test for a reduction in the value-weighted abnormal returns from Cohen and Frazzini's strategy. We run Equation 2 to obtain the results. Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *. t-values in parentheses.

Value-weighted returns	1	2
Post-sample	**-0.014***	**-0.014**
dummy	**[-1.89]**	**[-1.90]**
Post-publication	**-0.009**	
dummy	**[-1.72]**	
Early dummy		-0.007
		[-1.00]
Late dummy		-0.009
		[-1.47]
Constant	0.01	0.01
	[3.61]	[3.57]
Observations	444	444

included in quintile 1 or 5 (the extremes) of the customer-momentum portfolio. Similar to what we find in the main part of the analysis, we find significantly higher trading and dollar volume for these stocks. Together with the trend in the general market, this might imply that the stocks in all the quintiles have had increased trading activity. This might further be evidence for why the returns have decreased.

Next, we investigate summary statistics of the overlapping years of our dataset and Cohen and Frazzini's dataset for robustness and cross-referencing. In Table A.7 we show summary statistics of our dataset for the subsample 2000-2009, while in Table A.8 we show summary statistics for Cohen and Frazzini's dataset for the same subsample, 2000-2009. These overlapping years are the most important for our summary statistics, as this is essentially the only period we are able to cross-reference our data collection process with theirs. We argue that the results are very similar, and when taking the other results discussed above into account, we believe that this is evidence for a very adequate data collection process.

In Table A.9, we show summary statistics for 2009-2017, which is the relevant subsample used for the customer-momentum strategy shown in chapter 5. And lastly, for our summary statistics, in Table A.10, we show the summary statistics that are relevant for the Mclean and Pontiff regressions.

For our two next tables, Table A.11 and A.12, we utilize the panel-VAR and Granger-causality test to investigate if there are any differences between above and below median percentage sales to customers, which we believe should matter, based on the strength of the customer-supplier relationships. Most of these results are discussed in chapter 5. It is worth noting that there is no predictivity or causality towards the customer for our extended dataset of 2009-2017 when reviewing above and below median percentage sales. This is in contrast to the results when we review the whole sample which is not affected by the split

Table A.5: McLean and Pontiff regressions on Equal-weighted returns: 1981-2017

In this table, we test for a reduction in the equal-weighted abnormal returns from Cohen and Frazzini's strategy. We run Equation 2 to obtain the results. Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *. t-values in parentheses.

Equal-weighted returns	1	2
Post-sample	-0.005	-0.005
dummy	[-1.02]	[-1.01]
Post-publication	-0.003	
dummy	[-0.76]	
Early dummy		-0.002
		[-0.39]
Late dummy		-0.003
		[-0.62]
Constant	0.009	0.009
	[4.47]	[4.42]
Observations	444	444

of sample and need for reported sales and is thus much larger, which we argue is the reason for changing results. However, there is strong evidence for predictivity and causality from the customer towards the supplier, for both above and below median percentage sales.

For the Panel-VAR of above and below median percentage sale to customer for Cohen and Frazzini (2008)'s dataset of 1981-2004, the panels are not exactly balanced around the median. One reason is that there might be an overweight of panels at the median cut-off, which gives an unbalanced median.

For our two last tables, Table A.13 and A.14, we show variance decomposition for both datasets used in the main analysis. We also show the difference of the variance in returns based on the shocks when changing the Cholesky ordering of the variables, i.e. supplier first or customer first.

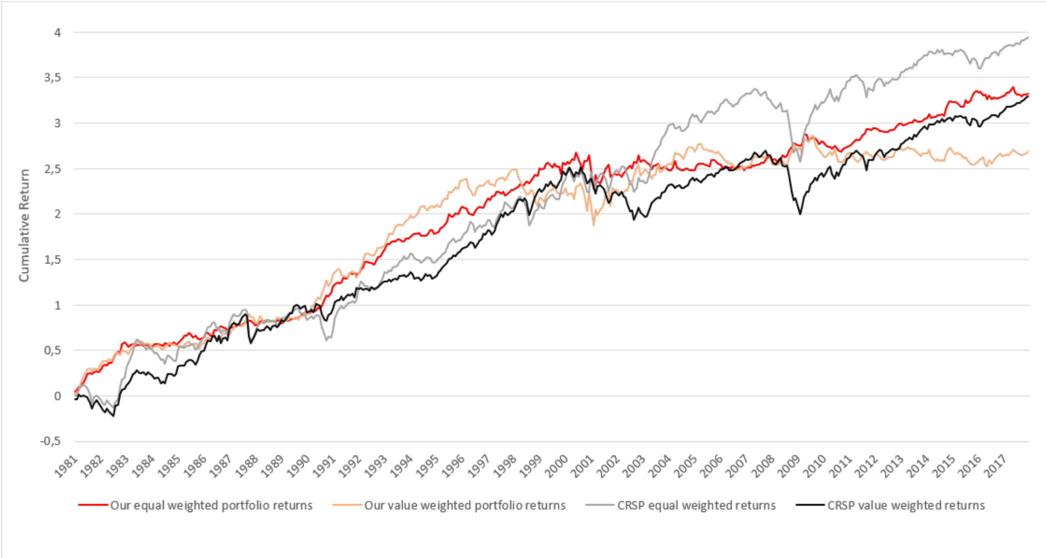

Figure A.1: Portfolio returns compared to CRSP index returns

Table A.6: McLean and Pontiff regressions on Trading Activity Dynamics for robustness: 1981-2017

In this table, we check for Trading Activity Dynamics for stocks not in the extreme quintiles of the Cohen and Frazzini strategy. The regressors are Trading Volume, which is measured as shares traded, and Dollar Volume, which is measured as shares traded multiplied by price. The Equation is similar to Equation 2, with dynamics instead of returns as regressors. Statistical significance at the 1, 5 and 10% level is indicated by a ***, ** and *. t-values in parentheses.

Variables	Trading Volume	Dollar Volume
Post-sample	**2.40*** **	**2.40*** **
dummy	**[14.81]**	**[13.75]**
Post-publication	**2.46*** **	**2.66*** **
dummy	**[22.23]**	**[22.32]**
Observations	444	444

Table A.7: Summary Statistics our data: 2000-2009

This table shows the summary statistics for
our dataset for the years 2000-2009.
"Stock universe" is all stocks in CRSP.
Link duration is number of years the firms are connected without breaks.
Same industry is based on the Fama & French industry definitions
Size percentiles is based on the size
of a customer or supplier in regards to the CRSP stock universe.
Percentage of sales to customer is the average sales
from a supplier the customers in the dataset count for.

	Min	Max	Mean	SD	Median
Panel A: Time Series (Annual Observations)					
Number of firms	900	1077	996	61	988
Number of customers	401	485	440	29	441
Full sample % coverage of stock universe (EW)	17.1	23.6	21.4	2.0	21.7
Full sample % coverage of stock universe (VW)	56.8	67.2	61.8	2.8	61.8
Supplier % coverage of stock universe (EW)	14.6	19.8	18.0	1.6	18.5
Supplier % coverage of stock universe (VW)	9.8	19.2	15.2	2.9	16.2
Customer % coverage of stock universe (EW)	5.4	10.7	8.0	1.2	8.1
Customer % coverage of stock universe (VW)	52.0	60.9	56.4	2.3	56.3
% of customer-supplier in the same industry	18.9	21.7	20.4	1.1	20.8
Link duration (Years)	1.0	10.0	2.6	2.1	2.0
Panel B: Firms (Pooled Firm-Year Observations)					
Supplier size percentile	0.00	0.99	0.50	0.29	0.51
Customer size percentile	0.01	0.99	0.81	0.21	0.90
Number of customers per firm	1.00	17.00	1.64	1.23	1.00
Percentage of sales to customer	0.00	100	17.77	15.98	13.75

Table A.8: Summary Statistics C&F data: 2000-2009

This table shows the summary statistics for
Cohen and Frazzini's dataset for the years 2000-2009.
"Stock universe" is all stocks in CRSP.
Link duration is number of years the firms are connected without breaks.
Same industry is based on the Fama & French industry definitions
Size percentiles is based on the size
of a customer or supplier in regards to the CRSP stock universe.
Percentage of sales to customer is the average sales
from a supplier the customers in the dataset count for.

	Min	Max	Mean	SD	Median
Panel A: Time Series (Annual Observations)					
Number of firms	947	1306	1129	115	1120
Number of customers	505	598	547	36	540
Full sample % coverage of stock universe (EW)	21.8	26.8	25.3	1.5	25.8
Full sample % coverage of stock universe (VW)	68.1	71.9	69.6	1.2	69.6
Supplier % coverage of stock universe (EW)	17.7	21.2	20.3	1.1	20.7
Supplier % coverage of stock universe (VW)	11.6	21.5	17.1	3.2	17.6
Customer % coverage of stock universe (EW)	6.9	13.2	10.0	1.5	10.0
Customer % coverage of stock universe (VW)	62.6	66.6	64.4	1.4	64.3
% of customer-supplier in the same industry	18.7	23.1	21.4	1.8	21.7
Link duration (Years)	1.0	10.0	2.6	2.1	2.0
Panel B: Firms (Pooled Firm-Year Observations)					
Supplier size percentile	0.00	0.99	0.49	0.29	0.50
Customer size percentile	0.01	0.99	0.81	0.20	0.89
Number of customers per firm	1.00	21.00	1.78	1.29	1.00
Percentage of sales to customer	0.00	100	18.34	15.72	14.00

Table A.9: Summary Statistics: 2009-2017

This table shows the summary statistics for
our dataset for the years 2009-2017.
”Stock universe” is all stocks in CRSP.
Link duration is number of years the firms are connected without breaks.
Same industry is based on the Fama & French industry definitions
Size percentiles is based on the size
of a customer or supplier in regards to the CRSP stock universe.
Percentage of sales to customer is the average sales
from a supplier the customers in the dataset count for.

	Min	Max	Mean	SD	Median
Panel A: Time Series (Annual Observations)					
Number of firms	714	900	803	53	792
Number of customers	417	497	442	28	433
Full sample % coverage of stock universe (EW)	23.6	25.4	24.2	0.6	24.0
Full sample % coveage of stock universe (VW)	63.4	67.8	65.5	1.7	65.9
Supplier % coverage of stock universe (EW)	18.5	20.0	19.6	0.4	19.8
Supplier % coverage of stock universe (VW)	15.2	17.5	16.3	0.7	16.5
Customer % coverage of stock universe (EW)	9.2	12.9	10.8	0.8	10.8
Customer % coverage of stock universe (VW)	58.2	62.3	59.9	1.4	59.8
% of customer-supplier in the same industry	18.0	21.2	19.8	1.3	19.5
Link duration (Years)	1.0	9.0	3.2	2.4	2.0
Panel B: Firms (Pooled Firm-Year Observations)					
Supplier size percentile	0.00	0.99	0.53	0.28	0.55
Customer size percentile	0.02	0.99	0.81	0.20	0.89
Number of customers per firm	1.00	23.00	2.06	2.14	1.0
Percentage of sales to customer	0.00	100	14.56	15.68	11.50

Table A.10: Summary Statistics: 1981-2017

Summary Statistics: 1981-2017					
This table shows the summary statistics for the dataset consisting of CF re-cleaned data in period 1980-1999 and our dataset from 2000 and onwards. ”Stock universe” is all stocks in CRSP. Link duration is number of years the firms are connected without breaks. Same industry is based on the Fama & French industry definitions Size percentiles is based on the size of a customer or supplier in regards to the CRSP stock universe. Percentage of sales to customer is the average sales from a supplier the customers in the dataset count for.					
	Min	Max	Mean	SD	Median
Panel A: Time Series (Annual Observations)					
Number of firms	640	1515	1017	225	1006
Number of customers	300	641	465	72	462
Full sample % coverage of stock universe (EW)	16.0	25.4	21.7	2.3	21.8
Full sample % coverage of stock universe (VW)	51.7	67.9	62.6	3.6	63.1
Supplier % coverage of stock universe (EW)	11.6	20.0	17.4	2.1	17.9
Supplier % coverage of stock universe (VW)	6.7	19.2	13.7	3.1	13.6
Customer % coverage of stock universe (EW)	3.6	16.6	8.3	2.4	7.7
Customer % coverage of stock universe (VW)	47.0	64.1	57.7	3.5	58.0
% of customer-supplier in the same industry	18.0	25.6	21.3	2.0	21.1
Link duration (Years)	1.0	34	2.9	3.0	2.0
Panel B: Firms (Pooled Firm-Year Observations)					
Supplier size percentile	0.00	0.99	0.50	0.28	0.50
Customer size percentile	0.00	0.99	0.83	0.20	0.91
Number of customers per firm	1.00	23.00	1.67	1.30	1.00
Percentage of sales to customer	0.00	100	18.46	16.86	14.00

Table A.11: Panel-VAR Above and Below median Percentage sales: C&F data 1981-2004

We utilize C&F's re-cleaned data in this table.
All customer-supplier links are treated as Panel-data,
where we run an adjustment of Equation 3 to obtain these results.
We utilize C&F's re-cleaned data in this table.
All customer-supplier links are treated as Panel-data,
where we run an adjustment of Equation 3 to obtain these results.
Coefficients are multiplied by 100. Statistical significance at the
1, 5 and 10% level is indicated by a ***, ** and *. z-values in parentheses.

	Below median		Above median	
Variables	Customer Returns	Supplier Returns	Customer Returns	Supplier Returns
Customer Returns t-1	**-1.4***** **[-2.56]**	**5.3***** **[7.35]**	-0.6 [-1.32]	**7.9***** **[11.38]**
Supplier Returns t-1	**2.9***** **[8.23]**	**2.0***** **[3.36]**	**1.7***** **[6.22]**	0.8 [1.41]
Average # of Months of the links (panels)	18	18	22	22
# of Observations	78575	78575	94726	94726
# of Panels	4353	4353	4237	4237
Granger Causality	***	***	***	***

Table A.12: Panel-VAR Above and Below median Percentage sales: Our data 2009-2017

We utilize our data in this table.
All customer-supplier links are treated as Panel-data,
where we run an adjustment of Equation 3 to obtain these results.
Coefficients are reported in percentages. Statistical
significance at the 1, 5 and 10% level is indicated by a ***, ** and *. z-values in parentheses.

	Below median		Above median	
Variables	Customer Returns	Supplier Returns	Customer Returns	Supplier Returns
Customer Returns t-1	1.4 [1.26]	**7.4***** **[6.19]**	-0.2 [-0.26]	**5.4***** **[4.43]**
Supplier Returns t-1	-0.9 [-1.62]	**-4.4***** **[-4.41]**	-0.3 [-0.64]	**-3.1***** **[-3.83]**
Average # of Months of the links (panels)	24	24	30	30
# of Observations	20815	20815	26229	26229
# of Panels	875	875	876	876
Granger Causality	-	***	-	***

Table A.13: Variance Decomposition for customer and suppliers: C&F data 1981-2004

We run a unit-root test, to review if the panel-VAR satisfies the stability conditions, for which we reject the null hypothesis of unit-root. Then we run Variance decomposition on the VAR-results for C&F's data from 1981-2004.
Coefficients are multiplied by 100.

Ordering	Supplier Customer		Customer Supplier	
Impulse variable Response variable	Supplier	Customer	Supplier	Customer
Supplier **1**	1	0	93.99	6.01
2	99.82	0.18	93.78	6.22
3	99.82	0.18	93.78	6.22
4	99.82	0.18	93.78	6.22
5	99.82	0.18	93.78	6.22
Customer **1**	6.01	93.99	0	1
2	6.10	93.90	0.09	99.91
3	6.10	93.90	0.09	99.91
4	6.10	93.90	0.09	99.91
5	6.10	93.90	0.09	99.91

Table A.14: Variance Decomposition for customer and suppliers: Our data 2009-2017

We run a unit-root test, to review if the panel-VAR satisfies the stability conditions, for which we reject the null hypothesis of unit-root. Then we run Variance decomposition on the VAR-results for our data from 2009-2017.
Coefficients are multiplied by 100.

Ordering	Supplier Customer		Customer Supplier	
Impulse variable Response variable	Supplier	Customer	Supplier	Customer
Supplier **1**	1	0	91.28	8.72
2	99.87	0.13	91.23	8.77
3	99.87	0.13	91.23	8.77
4	99.87	0.13	91.23	8.77
5	99.87	0.13	91.23	8.77
Customer **1**	8.72	91.28	0	1
2	8.73	91.27	0.01	99.99
3	8.73	91.27	0.01	99.99
4	8.73	91.27	0.01	99.99
5	8.73	91.27	0.01	99.99