Peer Momentum*

Efdal Ulas Misirli[†] Daniela Scida[‡] Mihail Velikov[§]

November 14, 2020

PRELIMINARY, CLICK HERE FOR LATEST VERSION.

Abstract

Using recent advances in network theory, we estimate the intra-industry connectedness for US publicly traded companies back to the 1920s. We develop a stock-level composite centrality measure that captures multiple dimensions of a stock's interdependence with its industry peers. Using our network and composite centrality estimates, we develop "peer momentum" trading strategies, which sort stocks on their industry peers' past month average returns weighted by the peers' influence in the industry. A "peripheral peer momentum" strategy that uses only peripheral stocks' influence as weights for the signal construction achieves an annualized Sharpe ratio of 0.70, survives a battery of robustness tests, and spans industry momentum.

JEL classification: G10, G11, G14, G40, Q41

Keywords: Asset Pricing, Anomalies, Network, Attention, Behavioral Finance, Connectedness

*The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of Richmond or the Federal Reserve System. We thank Jennifer Conrad and Francis Diebold for comments and suggestions. Cooper Killen provided excellent research assistance. All mistakes are ours alone.

[†]Federal Reserve Bank of Richmond, 502 S Sharp St, Baltimore, MD 21201. Email: ulas.misirli@rich.frb.org.

[‡]Federal Reserve Bank of Richmond, 530 E. Trade Street, Charlotte, NC 28202. Email: daniela.scida@rich.frb.org.

[§]Penn State University, 366 Business Building, University Park, PA 16802. Email: velikov@psu.edu.

1 Introduction

Understanding how information is incorporated into stock prices is a fundamental question in finance. An extensive behavioral finance literature documents a number of channels through which limited investor attention leads to slow diffusion of information into stock prices.[1] In some of these studies, the information that investors are slow to respond to comes from economically-linked firms, and in particular firms in the same industry. For example, Moskowitz and Grinblatt (1999) document a phenomenon called industry momentum, the tendency for stocks to follow their industry returns over subsequent months. Similarly, Hou (2007) argues that small firms in an industry tend to respond to industry-specific information. While it seems intuitive that the interdependence of firms within an industry has important asset pricing implications, measuring the connectedness between industry peers in a unified way has remained a challenge in the literature.

Motivated by this issue, we apply recent advances in the network literature to estimate the intra-industry connectedness for US publicly traded companies going back to the 1920s. Our estimation relies on volatility spillovers and is derived from co-movements in the stock market based on the methodology proposed by Diebold and Yilmaz (2014) and its extension for the estimation of larger networks in Demirer et al. (2017). This methodology requires high-frequency (e.g., daily) volatility estimates as input. Since realized daily volatility estimates require the use of intraday data, which are only available post 1983, instead we estimate conditional volatilities using the GARCH model of Bollerslev (1986). With the volatilities as inputs, we estimate generalized forecast error variance decomposition network matrices that pin down the manner in which firms within an industry are connected to each other.

With the network decomposition matrices in hand, we develop a new measure of centrality, termed 'composite centrality', that combines several dimensions of connectivity between firms

[1] See, e.g., Hirshleifer and Teoh (2003) for a theoretical model and Richardson, Tuna, and Wysocki (2010) for a survey of empirical studies that show investor underreaction to publicly available accounting information.

in an industry. Using principal component analysis (PCA), we extract the common variation between three of the most prominent measures of network centrality in the literature - degree, closeness, and eigenvector centrality. We define "composite centrality" as the first principal component of the three centrality measures for each cross-section in our sample.

Next, we study the determinants of composite centrality. Our large sample of all eligible publicly traded firms in the US allows us to see how network centrality is related to other proxies for the importance of a firm in a given industry. We find that the single most important determinant of composite centrality is the firm's market share of total revenues within the industry. This is a remarkable result, since our composite centrality measure does not rely on any accounting data for its estimation. We also find that firms, which are central in their industries, are larger, have high past performance, invest less, and have lower operating leverage.

In the second part of the paper we examine the implications of network centrality on asset pricing and the cross-sectional intra-industry trading strategies in particular. Our main hypothesis is that if our industry network estimates have any merit, and investors are slow to incorporate information about the inter-connectedness between stocks in a given industry, we should be able to exploit this slow diffusion of information and design profitable trading strategies. Moreover, if our hypothesis is true, we should also expect the influence of peripheral (i.e., not central) firms to be more important, since it is more difficult for investors to process those linkages.

Using our network and composite centrality estimates, we develop "peer momentum" trading strategies, which sort stocks on weighted average of their industry peers' past month returns. The weights come from our network decomposition matrices, which measure the connection between pairs of firms within an industry. We create a simple "peer momentum" strategy that uses all stocks within an industry to create our signal, and a "peripheral peer momentum" strategy that only uses the peripheral firms (i.e., those with composite centrality lower than the 90th percentile within the industry) to create our signal. The strategies are all value-weighted long/short portfolios from quintile sorts using NYSE breakpoints on our peer momentum signals.

Our peer momentum strategies are very profitable. Simple peer momentum yields average returns of 78 basis points per month (t-statistic of 5.82) between 1928 and 2015, with alphas ranging between 61 and 77 basis points per month depending on the factor model used. As expected, peripheral peer momentum (PPM) performs even better, achieving an average return of 86 basis points per month (t-stat of 6.53), and statistically significant alphas ranging between 71 and 85 basis points per month.

Given the stronger performance of the PPM strategy, we focus on it for the rest of the paper and study its robustness and link to industry momentum. In Fama-MacBeth regressions, PPM has strong predictive power for the cross-section of stock returns. In univariate Fama-MacBeth regressions, PPM has a t-statistic of 9.27. Controlling for classical individual stock momentum and industry momentum, as well as other common cross-sectional predictors such as size, value, profitability, investment, and short-term reversals, PPM maintains its strong predictive power with a t-statistic of 7.78.

Our PPM trading strategy also survives a battery of robustness tests. We show that most changes in strategy construction (e.g., decile sort, all-stock breaks, or equal-weighting) yield even better performance. The strategy works well in subsamples and within size quintiles. It also does not appear to be a different anomaly in disguise. Using double sorts, we show that it survives controlling for the 23 anomaly signals in Novy-Marx and Velikov (2016).

Finally, we study the relation between our peripheral peer momentum strategy and the industry momentum strategy of Moskowitz and Grinblatt (1999). The two strategies share plenty of the steps in the construction of the signals used for sorting stocks. Both PPM and industry momentum signals are weighted-averages of the past month's returns to stocks in the same industry. What differentiates PPM from industry momentum, however, is the weights used to calculate the signals. For a given stock ABC, industry momentum uses the value-weighted past month's return of ABC's industry. PPM, on the other hand, weights the past month's return for each of the ABC industry peers by the cell in the network decomposition matrix that indicates that peer's influence on ABC.

We show that PPM completely subsumes industry momentum. Using conditional double sorts and spanning tests controlling for one another, PPM continues to generate significant average returns, while industry momentum's returns become insignificant. We interpret this result to indicate that our network decomposition matrices, coupled with the composite centrality measure that indicates the peripheral firms, bring additional information for the impact of each industry peer on a given firm's stock returns, which is not incorporated by investors in a timely manner.

Our paper contributes to several different strands of the literature. While multiple papers capture the asset pricing implication of economically linked firms, those papers typically focus on a particular channel through which these links operate (e.g., Moskowitz and Grinblatt, 1999; Cohen and Frazzini, 2008). In contrast, our approach relies on the network estimation to extract only the relevant information about the links between firms in the same industry from their stock returns. Thus, our measure likely captures multiple channels and contributes to studies in this area.

The most closely related paper to ours is Hoberg and Phillips (2018), which introduces text-based industry momentum. Using their text-based network industry classification (TNIC) in Hoberg and Phillips (2016) to identify peer firms, Hoberg and Phillips (2018) propose an alternative peer momentum strategy that yields strong profits in a short sample period (July 1997-December 2012). Our peer momentum strategies are different from text-based industry momentum with respect to the industry classification of peers and the ranking period of signal construction. TNIC is based on 10-K (SEC EDGAR) product descriptions where each firm having its own distinct set of competitors. In Fama French industry classification (FFIC), however, peers are the same for each firm operating in the same industry. Hoberg and Phillips

(2018) show that TNIC overlaps on a limited domain with FFIC; hence, their text-based industry momentum becomes distinct from classic industry momentum with one month ranking period. Underlining the distinct and less visible nature of peer connections in TNIC, Hoberg and Phillips (2018) further show that their text-based industry momentum is profitable at different ranking periods varying from one month to one year. Our peer momentum strategies are built upon recent volatility spillovers among peers and are profitable at one month ranking period, however.

Another related paper is Lee, Sun, Wang, and Zhang (2019) technology momentum. Motivated by Bloom, Schankerman, and Van Reenen (2013) technological closeness measure, Lee et al. identify technologically linked firms and develop another one-month tech-peer momentum strategy. Unlike our peer momentum and classic industry momentum strategies, however, technologically linked firms often operate in different industries.

It is also worth emphasizing that our paper has broader scope than Hoberg and Phillips (2018) and Lee, Sun, Wang, and Zhang (2019). We also contribute to the network literature by being the first paper to estimate these networks for the entire eligible cross-section of US publicly traded firms going back to the 1920s. We also contribute to the network literature by introducing a new network centrality measure that captures multiple dimensions of connectedness. We show that our composite centrality measure is correlated to variables we would intuitively expect to be related to network centrality, and can be successfully applied to identify central and peripheral firms in an asset pricing application.

While we focus on an asset pricing application in this study, we believe our contribution opens new avenues for research. Peer interactions are proven to be quite important in different research areas such as fire-sales (Acharya, Bharath, and Srinivasan, 2007), relative performance evaluation (Albuquerque, 2009), corporate capital structures and financial policies (Leary and Roberts, 2014), among others. Future studies could make further advancements in these fields using our network decomposition estimates and composite centrality measure. [2]

2 Measuring Network Connectedness

2.1 Estimation methodology

Our analysis is based on network connectivity among firms in a given industry and across time. While there are several ways in which two firms can be connected with each other, there is no unique methodology to study networks nor one that encompasses all these possible connectivity dimensions. In this paper, we focus on connectivity derived from co-movements in the stock market based on the methodology proposed by Diebold and Yilmaz (2014) (DY2014 hereafter) and its extension to larger networks in Demirer, Diebold, Liu, and Yilmaz (2017) (DDLY2017 hereafter). We chose this methodology for several reasons. First, it is a well established methodology in the financial econometrics and financial networks literature. Second, we believe it is suitable for applications in asset pricing (e.g., peer momentum investment strategies), since it is based on stock market co-movements. Third, it captures several ways in which firms can be connected without the need to specify the actual source of the shock.

2.2 Volatility estimation

We estimate the network for each Fama-French industry every June starting 1928 until 2015. Each industry network is based on the volatility connectedness of the public firms that operate in the same industry. The volatility is, however, a latent variable and needs to be estimated. Diebold and Yilmaz (2014) study the network of financial institutions by employing daily realized stock volatilities as their volatility measure. They estimate realized volatilities using high frequency returns, which are available from 1999 onwards.

[2] Industry network data including measures for peer connectedness and network centrality are available from the authors upon request.

To extend the network estimation to the early periods of the 20th century, we construct conditional volatility measures using daily stock returns and the GARCH model of Bollerslev (1986). Daily returns are available at CRSP from 1926 onwards, and a GARCH(p,q) model expresses the daily conditional volatility σ^2_{t} as a function of its p lags and q lags of the squared unexpected stock return u^2_{t}:

$$\sigma^2_{\text{t}} = \alpha_0 + \sum_{i=1}^{q} \alpha_i u^2_{\text{t-i}} + \sum_{j=1}^{p} \beta_j \sigma^2_{\text{t-i}} \tag{1}$$

Volatility estimations of individual stock returns start in June 1928 and repeat every year using all available historical data prior to the estimation date. Before estimating the GARCH(p,q) models, we filter CRSP data in two different ways:

- We drop stocks that have missing daily return observations over a two year period that precedes the estimation date. A minimum of 504 return observations is necessary to produce reliable parameter estimates in the GARCH model.
- We drop stocks that fail the Engle (1982) test for ARCH effects at 1% level. An important reason for such failure is infrequent trading which leads to many zero return observations in consecutive days. We run this test to ensure that GARCH style models are appropriate to the data.

Figure 1 shows the coverage of our sample after the implementation of these filters. With regard to the number of observations, the coverage is less than 50% of CRSP universe for most years and ranges from 10% in late 1920s to close to 80% in 2010. The coverage drops mechanically at the opening of NASDAQ stock exchange and recovers slowly thereafter. With regard to the market capitalization, however, the coverage exceeds 50% for most years and reaches 90% towards the end of the sample. Such overrepresentation of large stocks is not surprising because investors trade large stocks more frequently than small stocks.

Next, for each stock that passes these filters, we run four different GARCH specifications, i.e. GARCH(1,1), GARCH(1,2), GARCH(2,1) and GARCH(2,2). Afterwards, we use the Akaike Information Criterion to select the best specification. The daily conditional volatility time series of the best specification starting from two years prior to the estimation date and ending at the estimation date become inputs to the industry network model explained in subsection 2.3.

2.3 Connectedness estimation

2.3.1 Variance decomposition network estimation

The methodology consists on estimating Vector Autoregressive (VAR) models of daily log realized volatility of stock returns on their own lags as well as those of other firms in the same industry. To deal with the curse of dimensionality that can arise in certain industry-years from having several firms, we use the shrinkage method extension of the DY2014 methodology proposed in DDLY2017 to estimate high dimensional VARs and ultimately networks. The final step of the network estimation consists of using a generalized forecast error variance decomposition based on the estimated VARs to pin down the weights of the edges connecting firms. This decomposition depends on the forecasting horizon whose length can be chosen based on the application at hand. According to DY2014, the connectedness horizon is important as it relates to issues of dynamics connectedness as opposed to purely contemporaneous connectedness.

The estimated network matrix is both weighted and directed. By construction, all nodes in the network are connected, yet the importance of the connections can vary significantly depending on the weight. Each entry of the network matrix corresponds to a weight representing the strength of the corresponding connection. While all nodes are connected to each other, the matrix is not restricted to be symmetric since it is directed (e.g., node i may affect node j more than the other way around). Consequently, the network matrix provides not only the strength of connections but also allows weights to vary depending on the direction of a connection (i.e., from i to j vs from j to i).

We estimate a network matrix for each of the 49 Fama-French industries and each of the 88 years of data using a forecasting horizon of 10 days.[3] The number of firms can significantly vary across years as shown in Panel B of 1. In particular, at the beginning of the time series some industries consists of only a few firms. As a result, we require a minimum of 5 firms in order to estimate the network. This leads us to a final count of 2836 estimated networks. Each firm corresponds to a node in the network, the names of which are listed in the rows and columns of the network matrix, and the edge weights between them correspond to the entries of the network.

2.3.2 Network centrality estimation

Centrality measures are widely used connectivity measures to study the topology of a network.[4] These measures aim at answering the question of who is the most important or "central" node in a given network. The answer to this question may vary depending on how we define what it means to be "central." For example, a node could be central because it has many connections within the network (only direct connections matter) or it could be central because it is connected to a few but well connected nodes (direct plus indirect connections matter). Several measures of centrality have been proposed in the network literature, we focus on three of the most popular centrality measures as they have proven to complement each other well.

Mathematically, a network of K individuals can be represented by a $K \times K$ adjacency matrix $\mathcal{A}$ of the network graph. By convention, elements in the i^{th} row of $\mathcal{A}$ indicate which nodes are affected by node i. The adjacency matrix is simply the transpose of the network

[3]We chose this horizon as it allows for dynamics of connections to take place while not being overly long. For example, the Basel accord requires a 10-day Value at Risk (VaR) horizon.

[4]For a detailed discussion on centrality measures see Jackson (2008).

matrix.[5] Consequently, the adjacency matrix considered in this paper is also weighted and directed. Let $[\mathcal{A}]_{ij}$ be the (i, j) element of the adjacency matrix $\mathcal{A}$ (i.e., i^{th} row and j^{th} column). We define the following centrality measures:

(i) *Degree centrality:* This measure represents how well a node is connected in terms of number of direct connections (in or out) and it is usually normalized by the number of total possible connections from a node (i.e., $K - 1$). As a result, the normalized measure ranges between 0 and 1. Formally, for node i we have:

$$\left.\begin{array}{lll} C_d^{i\leftarrow j} = \frac{\#\{j:\,[\mathcal{A}]_{ji}\neq 0\}}{K-1} & & [\text{In-degree}_i] \\ C_d^{i\rightarrow j} = \frac{\#\{j:\,[\mathcal{A}]_{ij}\neq 0\}}{K-1} & & [\text{Out-degree}_i] \end{array}\right\} \quad C_d^{i-j} = C_d^{i\leftarrow j} + C_d^{i\rightarrow j} \quad [\text{Total-degree}_i] \tag{2}$$

where C_d stands for degree centrality. The notation $i \leftarrow j$, $i \rightarrow j$, and $i - j$ indicates the direction of a connection or lack thereof. The numerator $\#\{j : [\mathcal{A}]_{ji} \neq 0\}$ counts the number of nodes j such that there is a connection from j to i (i.e., $[\mathcal{A}]_{ij} \neq 0\}$). Similar interpretation applies for $\#\{j : [\mathcal{A}]_{ij} \neq 0\}$. Therefore, in-degree (out-degree) counts the number of connections to (from) a node i from (to) other nodes "j" in the network; while total degree measures the total number of connections "in" and "out" for a given node i. Since there are at most $K - 1$ possible connections "in" or "out" for a given node, the denominator $K - 1$ is simply a normalizing factor that sets the measure between 0 and 1 (if desired). A high value of degree centrality (i.e., closer to 1) corresponds to a high number of in, out, or total direct connections; while a low value of degree centrality (i.e., closer to 0) corresponds to a low number of in, out, or total direct connections.

This is the simplest measure of centrality. However, degree centrality overlooks several important features of a network. For instance, it does not account for how well located a node

[5] If W denotes the network matrix, then $\mathcal{A} = W'$. Therefore, elements in the i^{th} row of W captures which nodes affect node i. Although the network matrix representation is often more intuitive, centrality measures are defined and coded in most software packages using the adjacency matrix notation. This is why we use the adjacency matrix representation

is in the network. That is, two nodes can have the same degree centrality but only one of them may be key for information to get transmitted from/to other nodes (i.e., lie in the middle of paths connecting many other nodes). The next measure takes some of these additional features into account.

(ii) *Closeness centrality:* It reflects how close a node is, on average, to any other node. Intuitively, it describes the extent of influence of a node on the network. Being close to every other node can be important when something is transmitted through the network. Formally, let $l(j,i)$ be the (shortest) distance between i and j given by their shortest path.[6,7] The closeness centrality of node i is defined by the inverse of the average length of the shortest paths to/from any other node j in the graph:

$$C_{cl}^{i} = \frac{K-1}{\sum_{j \neq i} l(j,i)}$$

Both these measures focus mostly on "quantity" rather than "quality" of connections. We are also interested in assessing not only how close a node is to many other nodes, but also whether it is connected to other "central" nodes or "key players" in the network. The last measure is based on the premise that a node's importance is determined by how important its neighbors are.

(iii) *Eigenvector centrality*: It generalizes degree centrality by taking into account the prestige of a node's neighbors. The centrality of each vertex i is proportional to the sum of centralities of its neighbors:

$$C_{eg}^{i} = \frac{1}{\lambda} \sum_{j:j \neq i} [\mathcal{A}]_{ij} C_{eg}^{j} \tag{3}$$

[6]If there is no path between j and i, the total number of vertices is used instead of the path length.

[7]Given that we have a weighted directed network, we compute the shortest path between node i and j by summing the inverse weights involved in each path connecting the two nodes and then choosing the path with the smallest value of total inverse weight.

where C^i_{eg} (C^j_{eg}) stands for eigenvector centrality of node i (j); $j \in \mathcal{M}(i)$, $\mathcal{M}(i)$ is the set of neighbors of i, and λ is a constant. The sum is over all nodes j different from i, and it is weighted by the weight of the connection from j to i (i.e., $[\mathcal{A}]_{ij}$). Note that if there is no connection from j to i, then the weight is 0. Equation (3) can be written in vector notation as the eigenvector equation:

$$\mathcal{A}C_{eg} = \lambda C_{eg} \tag{4}$$

A vector C_{eg} satisfying the above equation is an eigenvector of $\mathcal{A}$. In general, there are multiple eigenvalues λ for which a non-zero eigenvector solution exists. However, the additional requirement that all entries in the eigenvector must be non-negative implies (by the Perron-Frobenius theorem) that only the greatest eigenvalue of $\mathcal{A}$ results in the desired centrality measure.[8] The i^{th} component of the eigenvector C_{eg} then gives the relative centrality score of the vertex i. The eigenvector is only defined up to a common factor, so only the ratios of the centralities of the vertices are well defined. In order to define an absolute score we normalize the eigenvector such that the sum over all vertices equals 1.

(iv) *Composite centrality*: Thus far we have introduced three competing measures of centrality. While it is true that these measures are often times positively correlated, each is intended to capture possibly unique features of the network topology. Furthermore, there is as priori no compelling reason to use one of them over the others to pin down central nodes. Consequently, we propose using a composite centrality measure that combines the information embedded in all three of them.

Using principal component analysis (PCA), we extract the common variation between the three centrality measures discussed above - degree, closeness, and eigenvector centrality. At the end of each June of each year, we take an $N \times 3$ matrix of the centrality measures, where N is the

[8]See Newman (2016).

number of stocks with available centrality measures, and run a PCA on that matrix. Then, we define "composite centrality" as the first principal component (i.e., the eigenvector associated with the highest eigenvalue) of the matrix, rescaled so that the centrality measure is positive and adds up to one within each industry.

2.4 Understanding connectedness

In this subsection we discuss how to interpret the estimated connectedness measures among firms (i.e., network matrix and centrality measures). We use the beer and liquor industry as a leading example to illustrate the several aspects of the network estimation. We also discuss summary statistics across industries.

2.4.1 Network matrix

Table 1 shows an example of a variance decomposition matrix (i.e., network matrix) corresponding to the beer and liquor industry as of June 2006. The network matrix is 9 $\times$ 9 since there are nine firms in this year for this industry. The names in the rows and columns correspond to firms' tickers. The off-diagonal values of the matrix correspond to the forecast error variance decomposition and are interpreted as the importance of directional pairwise connections (i.e., network weights). The sum of the off-diagonal elements of the i^{th} row gives the share of the H-step forecast error variance of firm i coming from shocks arising in other firms. For example, we see that Anheuser-Busch (BUD) is the most influential connection for Molson-Coors (TAP) (i.e., 0.18 value in row TAP, column BUD or BUD $\rightarrow$ TAP connection). While the converse is also true, the weight of the TAP $\rightarrow$ BUD connection is much lower (0.03). This also serves to illustrate the directional feature of connections (i.e., $A \rightarrow B \neq B \rightarrow A$), which is captured by the asymmetry of the network matrix.

For completeness, we also report diagonal values although they are of no interest for pairwise connections analysis and are, therefore, disregarded throughout the discussion. Also, each row in the network matrix is standardized to sum to 1 to ease interpretation of weights. Lastly, it is important to note that by construction all elements of the matrix are non-zero, and any zeros displayed in the table are the result of rounding. This means that by assumption everyone is connected to everyone although with possibly different strengths of connections.

2.4.2 Network centrality

Table 2 illustrates the four centrality measures discussed in section 2.3.2 using the beer and liquor industry (June 2006) as an example. As explained in section 2.3.2 all centrality measures are standardized to sum to one. This is a useful normalization as it allow us to easily pin down central nodes in a given year. Indeed, we are interested in the relative ranking of nodes based on their centrality as opposed to the absolute value the measure takes.

As expected most centrality measures are positively correlated. However, closeness centrality is mostly negatively correlated with other measures. This negative correlation is an example of how the use of a set of centrality measures as opposed to only one can be beneficial. Interestingly, our composite centrality measure, which combines the three individual measures, is more aligned in terms of rank-order with eigenvector and total degree centrality for the beer and liquor industry. Overall, the most central firms across all measures (except closeness) are Anheuser-Busch and Molson-Coors. This makes sense since these are the largest firms in the industry. In third place of centrality we have Pyramid Breweries, which, contrary to the prior two firms, it is also quite central from a closeness point of view.

In addition to the beer and liquor industry example, we provide more detailed statistics for our composite measure. Table 3 shows the industries with the bottom and top composite centrality measure ranges across time. Given the high dimensionality of the data, we show time evolution across three different years: 1930, 1970, and 2010. We also present summary statistics for all industries across years in Table A.1 of the Appendix.

We can observe that the composition of industries with the smallest and largest ranges of values for composite centrality has changed over time. For example, in 1930, the Electrical Equipment industry had the highest range of composite centrality estimates. Out of the five firms in it, the industry was dominated by the Electric Storage Battery Company, which had a composite centrality of 0.81, while the Westinghouse Electric and Manufacturing Company (a predecessor of CBS Corp) had a zero composite centrality.

Fast forward to 2010, however, the industry with the highest composite centrality range, Textiles, only had a range of 0.25. This could partially be attributed to the fact that nowadays industries tend to have a lot more firms and that we rescale our composite centrality measure to add up to one within each industry. In the Textiles industry in 2010, however, Interface Inc, a global commercial flooring manufacturer, had the highest composite centrality measure of 0.27. UniFirst Corporation, a uniform rental company, and Mohawk Industries, a carpet and flooring manufacturer, had the second and third highest composite centrality measures of 0.17 and 0.12, respectively, while all other 10 companies had composite centrality measures less than 0.10.

Next, we show the relation of the composite centrality measure to standard centrality measures and firm characteristics using Fama-MacBeth regressions. The first three columns of Table 4 show that the composite measure is positively related to each of the three standard centrality measures that we discuss above, yet some of them have more footprint than others. Specifically, total degree becomes the main driver of the composite measure, while eigenvector and closeness have secondary and tertiary effects, respectively.

The last column of Table 4 shows the relation between the composite measure and important firm characteristics such as size, book-to-market, market share, capital spending, operating leverage, past one-year return. We find that central firms tend to be large firms with valuable growth opportunities, control a high market share in their industry, invest less in physical and R&D capital, have low operating leverage, and experienced significant gains in their stock price over the past year.

2.4.3 Network dynamics

The network topology for each industry vary over time. Firms enter and exit the industry thereby changing the number of peer firms and the distribution of market shares among peers. Similarly, some incumbent firms grow more than others or merge with their peers (or outsiders) thereby increasing their market power in the industry. To show this dynamic feature of the network, we analyze its evolution around an important event. Continuing with our beer and liquor industry example, we study the evolution of its network around the Molson-Coors merger.

Molson-Coors was formed in 2005 through the merger of Molson of Canada, and Adolph Coors of the United States. Adolph Coors became the parent of the merged company and changed its name to Molson Coors Brewing Company. The resulting brewing giant controlled 11% of US beer sales volume in 2005 and ranked third after Anheuser-Busch of the United States (controlling 49%) and SABMiller of Great Britain (controlling 18%). The merger was announced in July 2004 and legally completed in February 2005, yet the audited financial statements reflecting the revenue gains and the cost savings of the merger became available to shareholders at Molson-Coors annual stockholder meeting in May 2006.

Hence, we take the perspective of Coors investors and analyze the evolution of beer and liquor industry by comparing its network structure as of June 2004 with that as of June 2006. This two-year gap is also consistent with the two-year (504 trading day) rolling window of our network estimation procedure.

Figure 2 displays both networks. The network on the top pertains to June 2004, whereas the network on the bottom corresponds to June 2006. Each bubble reflects a node (incumbent firm) in the network and each arrow reflects the link among two industry peers. The size of each bubble is determined by the composite centrality measure of the incumbent firm attached to the node and it is standardized across all years for comparison purposes. The larger the centrality measure, the more influential (i.e., central) the incumbent firm is in the network. In this regard, the bottom graph provides a visual representation of the centrality measures in Table 2. Darker edge colors represent stronger connections based on network weights; cutoff points are based on the 0^{th}, 40^{th}, 60^{th}, 80^{th}, 90^{th},100^{th} average quantiles across all years. Since the network is directed, we use curved edges when directional links between same pair of nodes (i.e., $A \to B$ vs $B \to A$) fall in different buckets.

In June 2004, the beer and liquor industry had two central firms, namely Anheuser-Busch (ticker: BUD) and Adolph Coors (ticker: RYD). Anheuser-Busch was the largest brewery in the industry, but it has experienced moderate sales growth since the beginning of 2000 (21% increase from December 1999 to December 2003). Adolph Coors is the third largest brewery, but it has almost doubled its sales since the beginning of 2000 (94% increase). The second largest brewery (Miller) is not shown in the network as its parent (SABMiller) was not listed in NYSE. The industry also included other noncentral breweries, wineries, and hard liquor producers.

Over the two-year period from June 2004 to June 2006, the beer and liquor industry observed several structural changes. Anheuser-Busch continued to be a central firm and increased its influence in the industry by preserving its US market share and expanding ambitiously to international markets such as China and UK. Molson-Coors emerged as an important new entity after the merger. Compared to its pre-merger levels, the surviving parent company (Adolph Coors) experienced 28% increase in net sales. These enhancements in market power are visible through the increases in the centrality measures of both companies in the network (90% and 20% increase for BUD and TAP respectively).

Several other changes in the network are worth mentioning. Pyramid Breweries, a microbrewery specializing in craft beers acquired a private craft beer firm (Portland Brewing) and increased its power in the network. Similarly, Constellation Brands, an important wine producer, acquired another wine producer in the network (Robert Mondavi). Golden State Vintners was sold to a group of private investors and exited the industry. Brown-Forman, which owns valuable liquor brands such as Jack Daniels, changed its industry classification after altering its business focus from liquor production to wholesale trade.

3 Peer Momentum

3.1 Strategy construction

This subsection explains the construction of two alternative peer momentum strategies that we employ throughout the paper. The first strategy is simple peer momentum. For each firm in each industry, we track the past performance of its peers and form a peer impact portfolio. Simple peer momentum is the signal that captures the past one-month return of this portfolio. To illustrate the portfolio construction, we use the network matrix of beer & liquor industry in Table 1. Specifically, we set the own impacts reported as diagonal elements in the network matrix to zero and rescale the other elements such that each row adds up to one. This procedure gives us a new matrix that includes the portfolio weights of the peer impact portfolio. The elements in the first row, for example, represent the weights for Coca-Cola Bottling Consolidated Company (ticker: Coke). Multiplying these portfolio weights by the past one-month returns of Cokes peers yields the simple peer momentum signal for Coke. Similar calculations produce the simple peer momentum signals for other firms in the industry.

The second strategy is peripheral peer momentum. For each firm in each industry, we track the past performance of its peers *except* those whose composite centrality measure is at the 90th percentile of the industry or above and form an alternative peer impact portfolio. Specifically, we set the own impacts as well as the impacts of the most central firm(s) in the industry to zero and rescale the other elements such that each row adds up to one. Continuing with the beer & liquor industry example, we assign zero weights to Anheuser Busch's (ticker: Bud) past returns in peer impact portfolios (see Table A.2 for details). Multiplying the resulting portfolio weights by past one-month returns yields the peripheral peer momentum signals for the firms in beer & liquor industry.

We implement these signal construction techniques to each industry at the end of June each year t. The portfolio weights that we use in these techniques are the same from June of year t to May of year t+1. The past one-month returns vary each month, however, thereby generating different signals for each firm each month. Once we collect all signals, we sort stocks on them, form our simple and peripheral peer momentum portfolios, and rebalance them monthly.

3.2 Strategy performance

We form five value-weighted portfolios on simple and peripheral momentum over the sample period July 1928-December 2015. Quintile 1 is the loser portfolio holding stocks with low (L) peer momentum, whereas Quintile 5 is the winner portfolio holding stocks with high (H) peer momentum. In addition, we construct a zero-cost investment strategy that buys the winner portfolio and sells the loser portfolio (H-L). We calculate the average returns of five peer momentum portfolios and (H-L) strategy and report them in Table 5.

For each peer momentum (H-L) strategy, we also test whether it can add significantly to the investment opportunity set. To answer this question, we run spanning tests, and regress (H-L) return on market excess return, Fama and French (1993) three factors, Carhart (1997) four factors including classic momentum, and Fama and French (2015) five factors. A significant intercept term (alpha) would indicate that adding a peer momentum strategy to the investment opportunity set generates an attainable Sharpe ratio that significantly exceeds that which can be achieved with Fama-French strategies alone. Furthermore, if the latter factor portfolios capture risk dimensions that investors care about, the regression coefficients convey useful information about the riskiness of peer (H-L).

Table 5 reports the average returns and alphas (intercepts) of the regressions. Panel A covers simple peer momentum. Average returns increase monotonically from the loser portfolio (L) to the winner portfolio (H), and (H-L) strategy earns an average return of 0.78% per month with a t-statistic of 5.82. The abnormal returns relative to the CAPM and the three-, four-, and five-factor models are 0.77%, 0.74%, 0.61% and 0.65% per month, respectively. Adding the classic momentum factor (UMD or up minus down) reduces the intercept terms slightly, yet the abnormal returns are still statistically and economically significant. In short, one-month peer momentum is a distinct anomaly, which can contribute significantly to investment opportunity set, and the risks embedded in Fama and French factors do not affect its profits.[9]

Panel B of Table 5 reports the same statistics for peripheral peer momentum. Average returns increase monotonically from the loser portfolio (L) to the winner portfolio (H), and (H-L) strategy earns an average return of 0.86% per month with a t-statistic of 6.53. The zero-cost peripheral peer momentum strategy offers both higher rewards and lower volatilities than its simple momentum counterpart. In an unreported simple regression of the peripheral (H-L) on simple (H-L), we also estimate an alpha of 0.14% per month with a t-statistic of 3.22. This evidence shows that excluding the most central firms from strategy construction increases the Sharpe ratio of peer momentum strategies.

Figure 3 further illustrates the benefits of the latter exclusion using cumulative returns. A $1 investment in winner-minus-loser peripheral momentum portfolio in June 1928 yields $3,170 in December, 2015. Similar winner-minus-loser stratgies yield $1,299, $1,006 and $249 for simple peer momentum, one-month industry momentum and classic one-year price momentum, respectively. Hence, we focus mainly on peripheral peer momentum strategy hereafter.

Table 6 shows the results of the Fama-MacBeth regressions of firm returns on (peripheral) peer momentum, classic price momentum (measured over twelve to two months horizon), one-

[9] We refer the interested reader to Table A.3 in the appendix for the loadings of peer momentum portfolios on Fama and French (2015) factors.

month industry momentum and other important firm characteristics (controls) such as size (log(ME)), book-to-market (log(B/M)), gross profits-to-assets (GP/A), real investment, and short-run reversal $r_{1,0}$. The sample period covers July 1963 through December 2015.

The first specification of Table 6 shows that peer momentum is an important predictor for the cross-section of returns. Firms whose industry peers performed well over the previous month generate higher average returns than those whose peers performed poorly. The second and third specifications control for classic price momentum and one-month industry momentum, respectively. The peer momentum is almost orthogonal to classic price momentum, whereas controlling for industry momentum somewhat lowers the predictive ability of peer momentum.

The fourth specification reproduces the well-known predictive relations between important firm characteristics and future stock returns in our sample. Firms with low market equity, high book-to-market, high gross profitability, low investment and low recent return (over the previous month) earn higher average returns than those that have high market equity, low book-to-market, low gross profitability, high investment and high recent return. The fifth specification brings all predictive variables together to highlight our contribution. We find that peer momentum has strong incremental power in predicting future stock returns in the presence of important characteristics studied in asset pricing literature.

3.3 Relation to industry momentum

The Fama-MacBeth regressions of Table 6 suggest that peer momentum predicts average returns. One can criticize these regressions because they evaluate each observation equally and thereby putting a lot of weight on micro-cap stocks. In addition, these regressions are sensitive to outliers and can impose a potentially misspecified parametric relation between explanatory variables (in particular, between peer momentum and industry momentum), which can blur the economic significance of the results. This subsection attempts to address these issues by considering the performance of value-weighted portfolios sorted on peer momentum and industry momentum and non-parametrically testing the hypothesis that peer momentum is subsumed by industry momentum and vice versa.

We first check whether the spread in average returns on peer momentum portfolios is profitable controlling for industry momentum. Hence, we sort stocks initially into five portfolios based on industry momentum. Then, within each quintile, we sort stocks into five portfolios based on peer momentum.

Panel A of Table 7 shows the average returns of 5 x 5 industry momentum and peer momentum portfolios. Within the low industry momentum quintile, average returns increase monotonically from low peer momentum portfolio to high peer momentum portfolio. The long/short high-minus-low peer momentum strategy earns 0.71% per month. Controlling for Carhart (1997) four factors has minimal impact on the latter strategy as it yields an alpha of 0.64% per month.

Next, we average the returns of each peer momentum quintile over the five industry momentum portfolios. Thus these quintile peer momentum portfolios control for differences in industry momentum. The last row in Panel A of Table 7 reports their average returns as well as that of the corresponding winner-minus-loser peer momentum strategy.

The unconditional peer momentum strategy of buying winners and selling losers is highly profitable and yields an average return of 0.86% per month (see Table 5). Controlling for industry momentum reduces the magnitude of this profit to 0.31% per month, yet it is still economically significant and statistically different from zero (t-statistic: 3.37). Although the two industry-based momentum strategies are highly correlated with one another, peer momentum captures additional profits beyond industry momentum.

Finally, we change the order of the conditional sort to check whether the spread in average returns on industry momentum portfolios is profitable controlling for peer momentum. In this alternative double sort, we sort stocks initially into five portfolios based on peer momentum. Then, within each quintile, we sort stocks into five portfolios based on industry momentum.

Panel B of Table 7 shows the average returns of 5 x 5 peer momentum and industry momentum portfolios. Controlling for peer momentum distorts the cross-sectional relation between average returns and industry momentum characteristic. In fact, the long/short high-minus-low industry momentum strategy fails to earn significant profits at 5% significance level in any peer momentum quintile. Averaging the returns of each industry momentum quintile over the five peer momentum portfolios produces a similar outcome: Controlling for peer momentum reduces the magnitude of industry momentum profits to 0.19% per month, which is not significantly different from zero. In sum, peer momentum accounts for industry momentum.

Our results are robust to controlling for classic momentum or running spanning tests. For example, we repeat the same double-sort exercises using peer momentum and classic momentum signals. Table A.4 in the appendix shows that neither momentum strategy subsumes the other. Spanning tests using peer momentum, industry momentum and classic momentum long/short strategies produce similar results. Table A.5 in the appendix shows that peer momentum adds significantly to the investment opportunity set, spans industry momentum, and is unrelated to classic mometum over the sample period July 1928-December 2015.

3.4 Robustness

3.4.1 Placebo test

An important contribution of this paper is to measure peer interactions. For each firm in each Fama-French industry, for example, our network analysis identifies its peers and quantifies how much it is influenced by them using a variance decomposition matrix (see Table 1). Its diagonal elements show firms' own contribution, whereas off diagonal elements show peer contributions. We use off-diagonal elements to construct peer momentum signals because they help us quantify lead-lag relations among peers.

One can question the usefulness of our network approach in constructing peer momentum signals. For instance, one could argue that the variance decomposition matrix provides little value and that alternative portfolio weights could generate more profits than the weights of our peer momentum strategy. We alleviate such concerns by conducting a placebo test.

At the end of each month, we replace the variance decomposition matrix of an industry with a random matrix of the same size using a draw from the log-normal distribution. Afterwards, we set the diagonal elements of this random matrix to zero and rescale its off-diagonal elements such that each row adds up to one. We multiply the resulting weights by one-month past returns of the firms in the industry to calculate placebo momentum signals. We collect these signals in each industry each month and construct a time series for holding period returns of a long/short winner-minus-loser placebo momentum strategy.

To compare the performance of peer momentum strategy with that of a placebo momentum strategy, we compute three statistics. The first statistic is the average return over the full sample reflecting the rewards of the strategy. The second and third statistics are the intercept terms (alphas) from simple spanning regressions of one winner-minus-loser portfolio on the other. A significant intercept term would indicate that the strategy in the dependent variable would provide improvements in the Sharpe ratio.

We compute these three statistics for a given placebo momentum draw, repeat the simulation 1000 times, and report the distributions of the statistics in Figure 4. The upper panel shows the distribution of average returns and that of their t-statistics. The average returns of placebo momentum strategies vary between 0.53% and 0.71% and are significantly lower than the average return of the peripheral peer momentum strategy (0.86%). The t-statistics of the average returns vary between 4.93 and 6.57. The t-statistic of the peripheral peer momentum strategy equals 6.53, which is greater than 99.9% percentile of the distribution. Hence, peripheral peer momentum strategy stands out as one of the most profitable industry based one-month momentum strategies.

The middle panel of Figure 4 shows the distribution of placebo momentum alphas relative to peripheral peer momentum strategy and that of their corresponding t-statistics. The alphas vary between 0.01% and 0.20% and are mostly insignificant. The 90th percentile of the t-statistics equals 1.83, which is only marginally significant at 10% level. Hence, peripheral peer momentum strategy spans most of the industry based one-month momentum strategies.

The bottom panel of Figure 4 shows the distribution of peripheral peer momentum alphas relative to placebo momentum strategies and that of their corresponding t-statistics. The alphas vary between 0.21% and 0.38% and are all significant. The t-statistics vary between 2.29 and 4.17, and their median equals 3.28. Hence, our proposed peripheral peer momentum strategy is not subsumed by any placebo momentum strategy.

Taken together, all these analyses reveal that the peer interactions suggested by our network analysis provide valuable information to investors and that adding peripheral peer momentum strategy to the investment opportunity set leads to significant improvements in Sharpe ratios.

3.4.2 Other robustness tests

So far we have shown that peripheral peer momentum is a strong predictor of returns in the cross-section of equities. Its predictive power survives a battery of additional robustness tests, the results for which are reported in the appendix. Table A.6 shows that the peripheral peer momentum strategy works well across size quintiles. It reports a conditional double sort where stocks are first sorted into quintiles based on their market capitalization using NYSE breakpoints. Within each size quintile, stocks are further sorted based on their peripheral peer momentum in quintiles again. We report the average returns (Panel A), number of stocks (Panel B), and average market capitalization (Panel B) for each of the resulting 25 portfolios. The right-hand-side of Panel A also reports the returns, alphas, and loadings on the Carhart (1997) four-factor model for the peripheral peer momentum long/short strategies within each size quintile.

The peripheral peer momentum strategies yield significant returns across all five size quintiles. Their average returns and alphas vary between 60 and 88 basis points per month, with t-statistics exceeding three for all but the strategy executed within the smallest quintile. Even there, however, it yields 75 basis points per month and has an alpha of 93 basis points per month.

Our main strategy reported in Table 5, Panel B is constructed using value-weighted portfolios formed by a quintile sort using NYSE breakpoints. Table A.7 shows that the strategy is robust to alternative strategy construction choices. Panel A uses equal-weighting instead of value-weighting, Panel B uses all-stock breaks instead of NYSE breaks, and Panel C uses a decile sort instead of a quintile one. Among all of these variations, the lowest performance is shown in the five-factor alpha for the long/short strategy reported in Panel B, which equal 76 basis points per month, with a t-statistic of 4.52.

Table A.8 shows that our peripheral peer momentum strategy is not a different anomaly in disguise. It reports excess returns on conditional peripheral peer momentum strategies, constructed from conditional 5x5 double sorts on each of the 23 anomaly signals from Novy-Marx and Velikov (2016) first, and peripheral peer momentum second. All of the twenty-three conditional peripheral peer momentum strategies generate sizeable and significant average returns. The lowest performance comes from conditioning on industry momentum, where the conditional peripheral peer momentum strategy yields 31 basis points per month with a t-statistic of 3.37.

Table A.9 further splits our sample in half and reports the performance of the peripheral peer momentum strategy following the format of Table 5, Panel B. The strategy performes slightly better in the first part of the sample (1928-1963), with average returns and alpha ranging between 99 and 109 basis-points per month. However, even in the second half of the sample (1963-2015), the strategy continues to earn significant profits, with the average returns and alpha varying between 57 and 78 basis points per month. Across both samples, the t-statistics for the performance metrics we look at for the long/short portfolio exceed four everywhere.

Finally, we are interested in whether our peripheral peer momentum strategy survives trading costs, especially in light of the high average monthly turnover in its portfolio composition. We follow Novy-Marx and Velikov (2016) and measure trading costs by subtracting half of the Hasbrouck (2009) Gibbs effective bid-ask spread estimate every time a position is entered or exited. Both the long and short portfolio turn over, on average about 78% of their holdings each month, placing the strategy among the high-turnover anomalies in the taxonomy of Novy-Marx and Velikov (2016). However, its incredibly strong gross performance results in average returns *net* of trading costs of 35 basis points per month, with a t-statistic of 2.62.

4 Conclusion

In this paper, we use recent advances in the network literature to estimate the intra-industry connectedness for US publicly traded companies back to the 1920s. We develop a stock-level composite centrality measure that captures multiple dimensions of a stock's interdependence with its industry peers. Using our network and composite centrality estimates, we develop two "peer momentum" trading strategies, which sort stocks based on their industry peers' past month average returns weighted by the peers' influence in the industry. We show that a "peripheral peer momentum" strategy that uses only peripheral stocks' influence as weights for the signal construction achieves an annualized Sharpe ratio of 0.70, survives a battery of robustness tests, and spans industry momentum.

Our paper uses network estimation techniques in an asset pricing setting, but we are excited about potential future applications of network estimation more generally and the composite centrality measure in particular. For example, Acharya, Bharath, and Srinivasan (2007) show that industry-wide distress affects recovery rates. Perhaps adjusting for the impact of each industry peer using our approach can better inform our estimates of loss-given default. Similarly, Albuquerque (2009) shows that CEO compensation exhibits peer effects, so perhaps using the network structure we estimate here can better inform us about executive compensation practices. Finally, Leary and Roberts (2014) show that peer effects impact corporate financial policies, so another potential application of our measure could be to study whether our network estimates are related to commonalities in these various corporate policies.

References

Acharya, V., S. Bharath, and A. Srinivasan (2007). Does industry-wide distress affect defaulted firms? evidence from creditor recoveries. *Journal of Financial Economics 85*(3), 787–821.

Albuquerque, A. (2009). Peer firms in relative performance evaluation. *Journal of Accounting and Economics 48*(1), 69–89.

Bloom, N., M. Schankerman, and J. Van Reenen (2013). Identifying technology spillovers and product market rivalry. *Econometrica 81*(4), 1347–1393.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Carhart, M. M. (1997). On the persistence in mutual fund performance. *Journal of Finance 52*, 57–82.

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

Cooper, M. J., H. Gulen, and M. J. Schill (2008). Asset growth and the cross-section of stock returns. *Journal of Finance 63*, 1609–1651.

Demirer, M., F. X. Diebold, L. Liu, and K. Yilmaz (2017). Estimating global bank network connectedness. *Journal of Applied Econometrics*.

Diebold, F. and K. Yilmaz (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics 182*(1), 119–134.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica 50*(4), 987–1007.

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

Fama, E. F. and K. R. French (2015). A five-factor asset pricing model. *Journal of Financial Economics 116*, 1–22.

Hasbrouck, J. (2009). Trading costs and returns for U.S. equities: estimating effective costs from daily data. *Journal of Finance 64*, 1446–1477.

Hirshleifer, D. and S. H. Teoh (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics 36*, 337–386.

Hoberg, G. and G. M. Phillips (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy 124*(5).

Hoberg, G. and G. M. Phillips (2018). Text-based industry momentum. *Journal of Financial and Quantitative Analysis 53*(6), 2355–2388.

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies 20*(4), 1113–1138.

Jackson, M. O. (2008). *Social and economic networks*, Volume 3. Princeton University Press.

Leary, M. and M. Roberts (2014). Do peer firms affect corporate financial policy? *Journal of Finance 69*(1), 139–78.

Lee, C., S. Sun, R. Wang, and R. Zhang (2019). Technological links and predictable returns. *Journal of Financial Economics 132*(3), 76–96.

Moskowitz, T. and M. Grinblatt (1999). Do industries explain momenum? *Journal of Finance 54*, 1249–1290.

Newman, M. E. J. (2016). *Mathematics of Networks in "The New Palgrave Dictionary of Economics"*. Palgrave Macmillan, UK.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics 108*, 1–28.

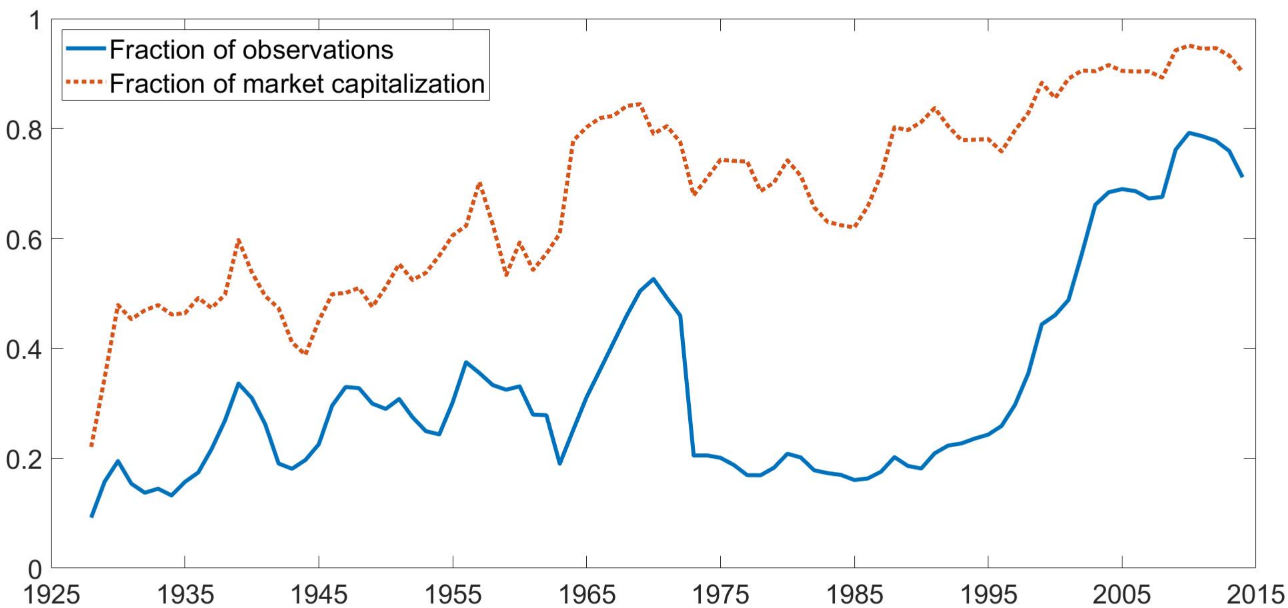
Novy-Marx, R. and M. Velikov (2016). A taxonomy of anomalies and their trading costs. *Review of Financial Studies 29*(1), 104–147.

Richardson, S., I. Tuna, and P. Wysocki (2010). Accounting anomalies and fundamental analysis: A review of recent research advances. *Journal of Accounting and Economics 50*, 410–454.

Figure 1: Sample coverage

This figure plots the coverage of our sample after the implementation of volatility estimation filters. Panel A shows the coverage with respect to all stocks in CRSP-Compustat-Merged database. Solid blue line displays the fraction of observations by count, whereas the dashed red line displays the fraction of observations by market value. Panel B shows the coverage of 49 Fama-French industries in our sample. The figure displays the nine most populated industries and shows how their firm count varies over time.

Panel A: All stock coverage



Panel B: Industry coverage

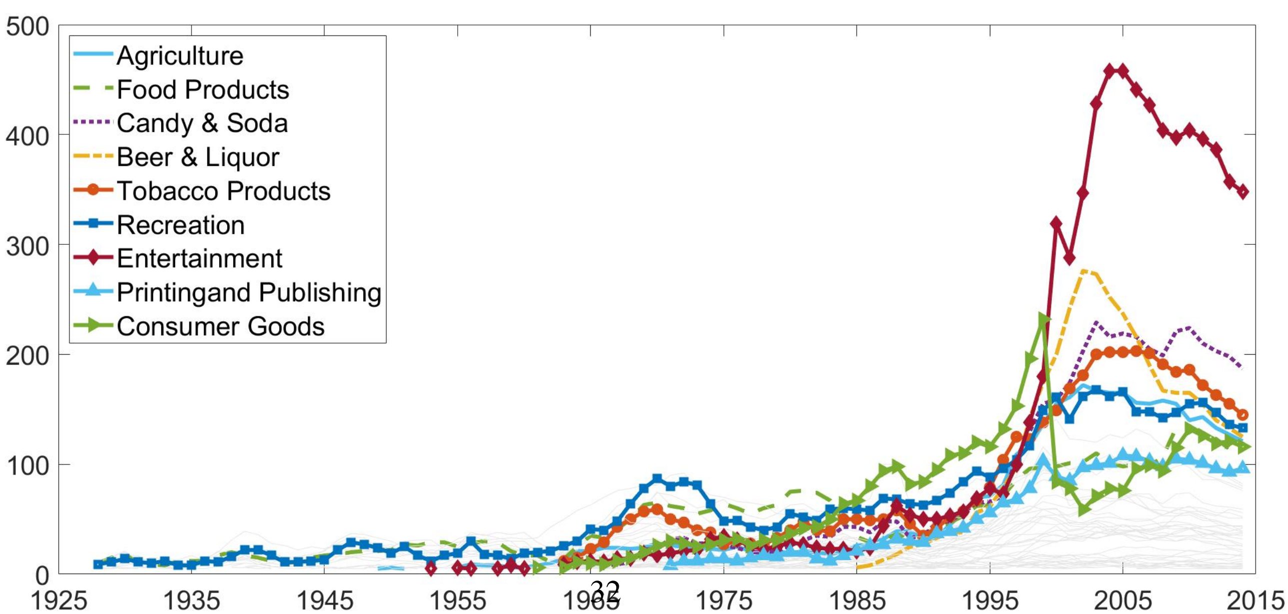
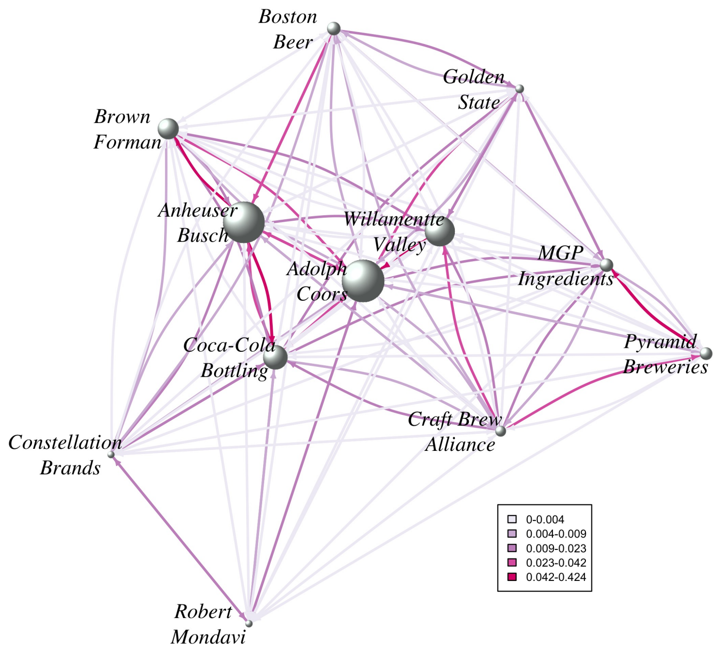


Figure 2: Beer and liquor industry network
This figure plots the network for the beer and liquor industry as of June 2004 (Panel A) and June 2006 (Panel B). Larger node size corresponds to higher composite centrality and is standardized across all years. Darker edge colors represent stronger connections based on network weights; cutoff points are based on the 0^{th}, 40^{th}, 60^{th}, 80^{th}, 90^{th},100^{th} average quantiles across all years.

(a) 2004



(b) 2006

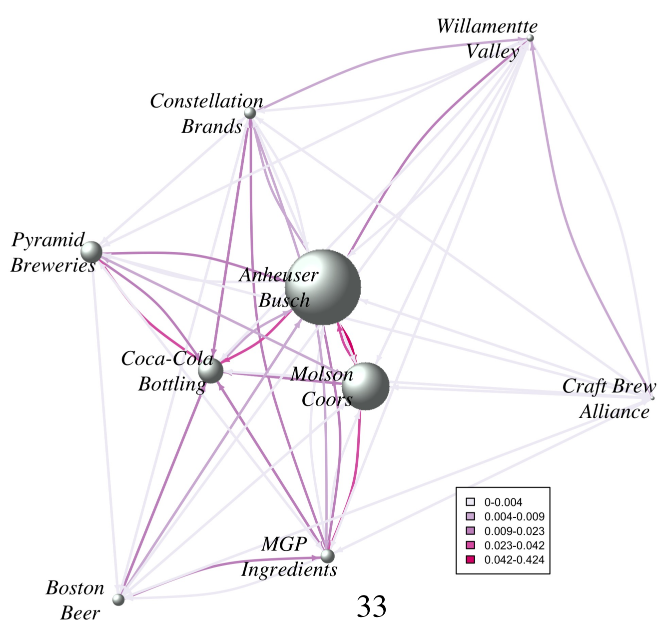


Figure 3: Strategy performance

This figure plots the performance of a $1 investment in alternative winner-minus-loser momentum strategies. The sample is monthly and spans the period July 1928-December 2015.

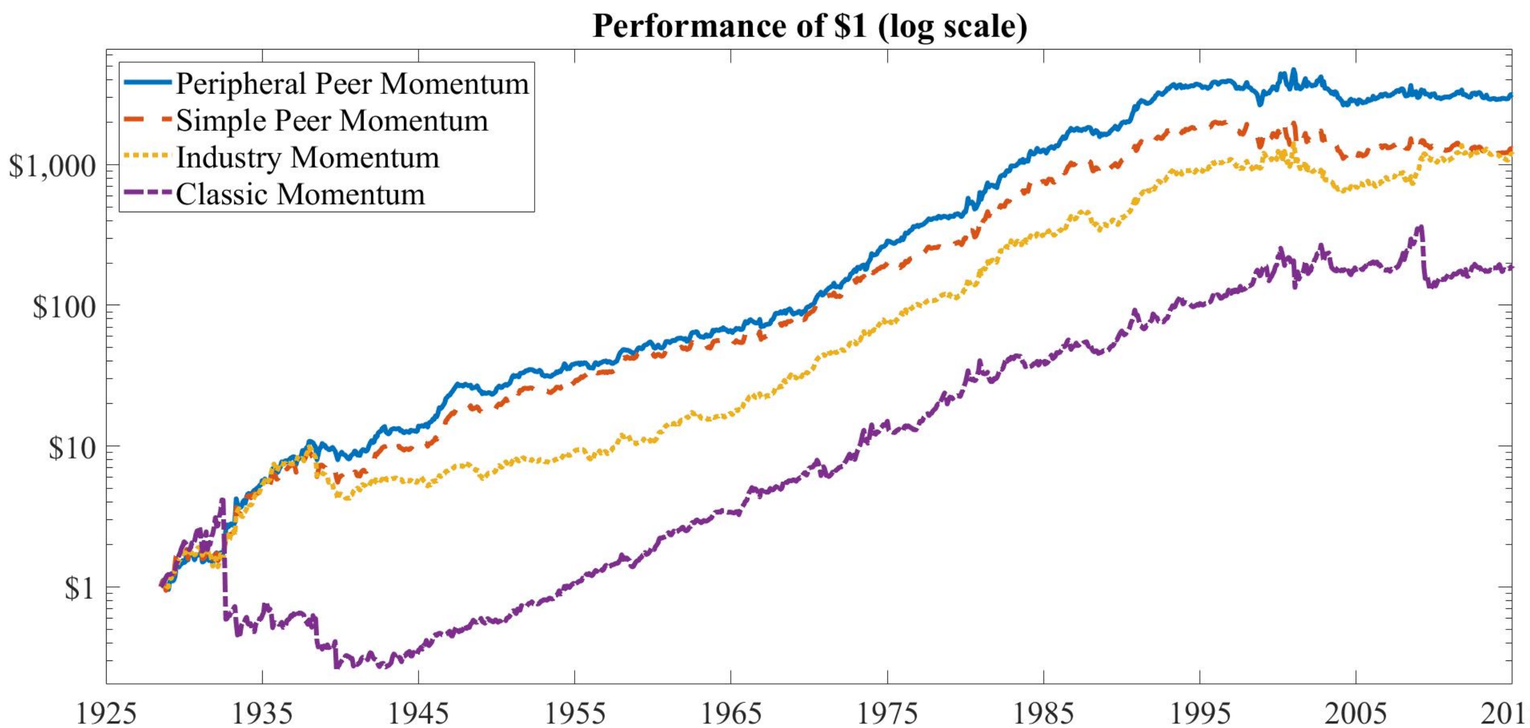


Figure 4: Placebo momentum performance

This figure compares the performance of peripheral peer momentum (PPM) strategy with that of a placebo momentum using the distributions of three statistics. The first statistic is the average return over the full sample July 1928-December 2015. The second and third statistics are the intercept terms (alphas) from simple spanning regressions of one winner-minus-loser portfolio on the other. We compute these three statistics (as well as their corresponding t-statistics) for a given placebo momentum draw, repeat the simulation 1000 times, and report the distributions of all statistics.

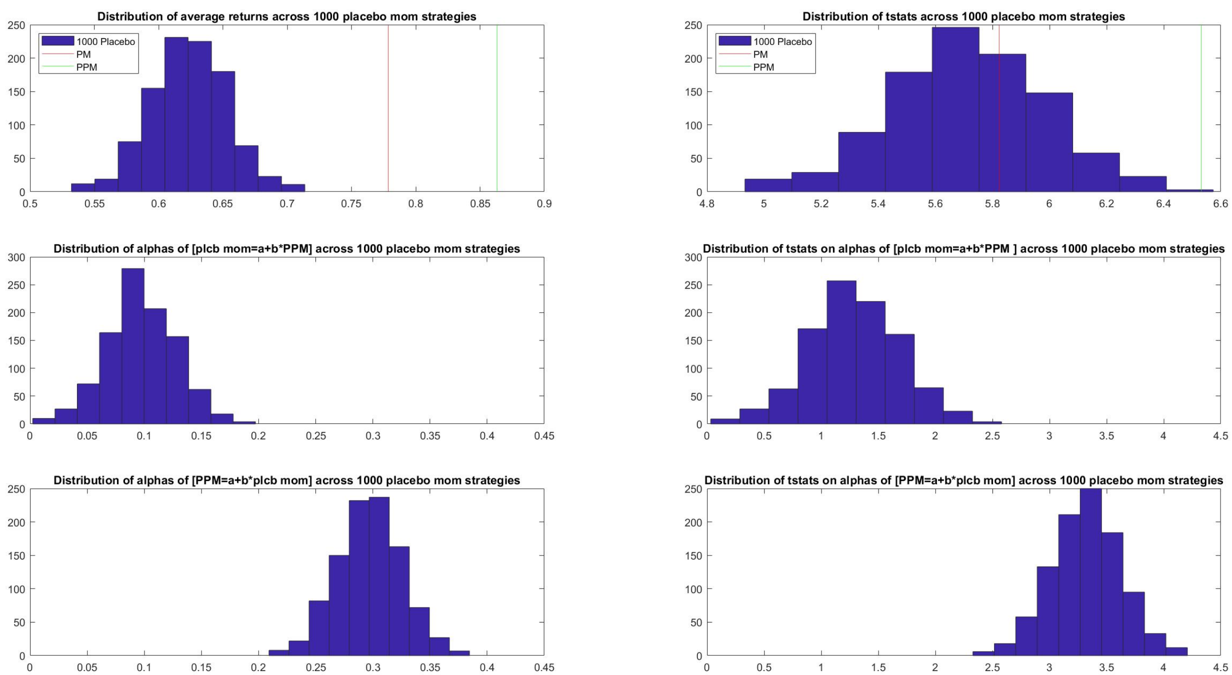


Table 1: Network Estimation: Example Industry Network

This table reports the variance decomposition matrix (i.e., network matrix) for the beer and liquor industry as of June 2006. The names in the rows and columns correspond to firms' tickers, and the off-diagonal values in the matrix are interpreted as the importance of pairwise directional connections (i.e., network weights). For completeness, we also report diagonal values although they are of no interest for pairwise connections analysis and can be ignored. Each row is standardized to sum to 1. By construction all elements are non-zero, and zeros displayed in the table are a result of rounding.

Variance decomposition for June, 2006									
	COKE	MGPI	BUD	TAP	STZ	WVVI	BREW	SAM	PMID
COKE	0.89	0.02	0.03	0.02	0.01	0.00	0.00	0.01	0.01
MGPI	0.01	0.92	0.01	0.04	0.01	0.00	0.00	0.01	0.00
BUD	0.00	0.00	0.93	0.03	0.00	0.00	0.00	0.01	0.02
TAP	0.01	0.00	0.18	0.80	0.00	0.00	0.00	0.00	0.00
STZ	0.02	0.01	0.00	0.00	0.95	0.00	0.00	0.00	0.00
WVVI	0.00	0.00	0.01	0.00	0.00	0.98	0.00	0.00	0.00
BREW	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00
SAM	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.97	0.00
PMID	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.95

Table 2: Network Estimation: Example Industry Centrality
This table reports closeness, eigenvector, and total degree centrality as well as the composite centrality measure for the beer and liquor industry corresponding to the network matrix estimated as of June 2006. Firms are sorted from higher to lower centrality based on the composite measure.

Centrality estimates for June, 2006				
Firm	Total degree	Closeness	Eigenvector	Composite
Anheuser-Busch Companies	0.25	0.08	0.45	0.34
Molson Coors Brewing Company	0.25	0.12	0.20	0.22
Coca-Cola Bottling Company	0.18	0.03	0.10	0.12
Pyramid Breweries	0.07	0.14	0.11	0.10
MGP Ingredients	0.09	0.12	0.04	0.06
Boston Beer	0.05	0.11	0.05	0.06
Constellation Brands	0.06	0.13	0.03	0.05
Willamette Valley Vineyards	0.03	0.16	0.01	0.03
Craft Brew Alliance	0.02	0.10	0.01	0.02

Table 3: Network Estimation: Bottom and Top Composite Centrality Range Industries

This table reports the bottom and top industries based on composite centrality measure ranges (maximum minus minimum) across three years (1930, 1970, and 2010).

1930		1970		2010	
Panel A: Lowest composite centrality ranges					
Automobiles and Trucks	0.11	Utilities	0.04	Banking	0.01
Steel Works Etc	0.12	Petroleum and Natural Gas	0.05	Business Services	0.02
Petroleum and Natural Gas	0.16	Electronic Equipment	0.05	Petroleum and Natural Gas	0.02
Transportation	0.17	Construction Materials	0.06	Electronic Equipment	0.02
Food Products	0.18	Retail	0.06	Computer Software	0.03
Panel B: Highest composite centrality ranges					
Utilities	0.23	Shipbuilding and Railroad Equipment	0.28	Shipping Containers	0.20
Entertainment	0.38	Restaurants, Hotels, Motels	0.28	Agriculture	0.22
Construction Materials	0.41	Fabricated Products	0.29	Defense	0.23
Chemicals	0.42	Candy and Soda	0.35	Beer and Liquor	0.25
Electrical Equipment	0.81	Beer and Liquor	0.41	Textiles	0.25

Table 4: Network Centrality: Composite Centrality Determinants

This table reports Fama-MacBeth regressions of composite centrality measure on standard centrality measures and important firm characteristics. Standard centrality measures are closeness, total degree and eigenvector centrality. Firm characteristics are firm size (log(ME)), the logarithm of book-to-market ratio (log(BM)), the component of market share that is orthogonal to firm size (MSHARE ORT), investment in physical and R&D capital (PRINV), operating leverage (OLEV), and past one-year return (R). Newey-West t-statistics are reported in brackets.

	(1)	(2)	(3)	(4)
const	-0.03	-0.05	-0.04	-1.40
	[-7.68]	[-7.34]	[-7.08]	[-3.38]
Closeness	1.11			
	[38.21]			
Total degree		1.54		
		[45.18]		
Eigenvector			1.27	
			[33.98]	
log(ME)				0.29
				[5.00]
log(BM)				-0.58
				[-2.45]
MSHARE ORT				56.44
				[17.79]
PRINV				-1.15
				[-4.82]
OLEV				-0.72
				[-8.35]
R				0.43
				[3.06]
adj. R^2	40.93	92.01	87.96	30.78
Avg. N	1,769	1,769	1,769	1,488

Table 5: Portfolio Performance

This table reports average excess returns and alphas of simple and peripheral peer momentum portfolios. At the end of each month, we sort stocks into five portfolios based on their industry peers' past month average returns weighted by the peers' influence in the industry. Simple peer momentum strategy uses all peers' influence, whereas "peripheral peer momentum" strategy uses only peripheral stocks' influence as weights for the signal construction. All portfolios are constructed using NYSE breaks. For each of the five portfolios, and for a portfolio long stocks with high peer momentum signal and short stocks with low peer momentum signal, the table reports average value-weighted returns in excess of the risk-free rate and alphas with respect to the CAPM, Fama and French (1993) three-factor model, Fama and French (1993) three-factor model augmented with the Carhart (1997) momentum factor, and the Fama and French (2015) five-factor model. Panel A (B) reports the statistics of simple (peripheral) peer momentum portfolios. T-statistics are in brackets. The sample period is 07/1928 to 12/2015 except for Fama and French (2015) five-factor model whose sample is available only after July 1963.

Panel A: Excess returns and alphas on simple peer momentum-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.20	0.49	0.72	0.88	0.97	0.78
	[0.99]	[2.68]	[3.96]	[4.75]	[4.88]	[5.82]
α^{CAPM}	-0.48	-0.15	0.09	0.23	0.30	0.77
	[-6.11]	[-2.58]	[1.35]	[3.65]	[3.61]	[5.75]
α^{FF3}	-0.47	-0.16	0.05	0.20	0.27	0.74
	[-5.97]	[-2.73]	[0.85]	[3.25]	[3.35]	[5.51]
$\alpha^{\text{FF3+UMD}}$	-0.37	-0.15	0.10	0.20	0.25	0.61
	[-4.59]	[-2.44]	[1.66]	[3.14]	[2.99]	[4.48]
α^{FF5}	-0.40	-0.24	0.12	0.11	0.25	0.65
	[-4.31]	[-3.75]	[1.95]	[1.78]	[2.84]	[4.01]
Panel B: Excess returns and alphas on peripheral peer momentum-sorted portfolios						
r^e	0.18	0.55	0.69	0.79	1.04	0.86
	[0.93]	[3.01]	[3.87]	[4.27]	[5.23]	[6.53]
α^{CAPM}	-0.49	-0.09	0.06	0.14	0.36	0.85
	[-6.25]	[-1.63]	[1.07]	[2.29]	[4.51]	[6.39]
α^{FF3}	-0.48	-0.10	0.04	0.11	0.34	0.81
	[-6.11]	[-1.81]	[0.67]	[1.88]	[4.23]	[6.12]
$\alpha^{\text{FF3+UMD}}$	-0.39	-0.07	0.07	0.12	0.32	0.71
	[-4.92]	[-1.24]	[1.13]	[2.09]	[3.86]	[5.20]
α^{FF5}	-0.48	-0.12	0.08	0.09	0.30	0.78
	[-5.11]	[-2.05]	[1.31]	[1.42]	[3.48]	[4.83]

Table 6: Fama-MacBeth Regressions

The table documents results from Fama-MacBeth regressions of the form $r_{tj} = \beta' \mathbf{x}_{t-1,j} + \varepsilon_{tj}$. The characteristics $\mathbf{x}_{t-1,j}$ include peripheral peer momentum ($r_{1,0}^{Peer}$), classic momentum ($r_{12,1}$), industry momentum $r_{1,0}^{Ind}$, the log of market capitalization (log(ME)), the log of the book-to-market ratio (log(BM)), gross profitability (GP/A), investment (I/A), and short-term reversals ($r_{1,0}$). Peripheral peer momentum strategy uses peripheral stocks' influence in the industry as weights for the signal construction, whereas industry momentum uses the market values of all stocks in the industry as weights. GP/A follows Novy-Marx (2013). I/A follows Cooper, Gulen, and Schill (2008). Independent variables are winsorized at the 1% level. T-statistics are in brackets. Sample period is 07/1963 to 12/2015.

	Regressions of the form $r_{tj} = \beta' \mathbf{x}_{t-1,j} + \varepsilon_{tj}$				
Coef.	(1)	(2)	(3)	(4)	(5)
$r_{1,0}^{Peer}$	10.13	9.56	6.73		5.91
	[9.27]	[9.79]	[6.94]		[7.78]
$r_{12,1}$		0.90			0.65
		[4.33]			[3.53]
$r_{1,0}^{Ind}$			6.99		7.49
			[5.81]		[7.23]
log(ME)				-0.08	-0.05
				[-2.02]	[-1.38]
log(B/M)				0.26	0.25
				[4.31]	[4.03]
GP/A				0.60	0.61
				[4.68]	[4.66]
Investment				-0.78	-0.63
				[-8.91]	[-5.81]
$r_{1,0}$				-5.81	-5.54
				[-14.50]	[-14.19]
R^2	0.95	2.63	1.35	3.44	6.46
n	630	630	630	630	630

Table 7: Conditional sort on industry momentum and peripheral peer momentum

This table presents results for conditional double sorts on industry momentum and peer momentum. In each month, firms are sorted first into quintiles based on industry momentum, then on peer momentum. Then, they are grouped into twenty-five portfolios based on the intersection of the two sorts. Panel A presents the average returns to the 25 portfolios, panel B documents the average number of firms and the average firm size for each portfolio. Time period is 07/1928 to 12/2015.

Panel A: Peer momentum conditional on industry momentum

Industry Momentum quintiles	Peer Momentum Quintiles (L)	(2)	(3)	(4)	(H)	Peer Momentum Strategies r^e	α_{FF4}	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}
(L)	-0.10	0.13	0.09	0.24	0.61	0.71	0.64	0.05	0.05	0.08	-0.01
						[4.04]	[3.52]	[1.44]	[0.84]	[1.43]	[-0.19]
(2)	0.34	0.48	0.48	0.81	0.55	0.22	0.13	0.04	-0.15	0.22	0.02
						[1.39]	[0.80]	[1.22]	[-3.02]	[4.71]	[0.43]
(3)	0.49	0.50	0.81	0.59	0.63	0.14	0.13	-0.01	-0.07	0.05	0.03
						[0.90]	[0.78]	[-0.34]	[-1.39]	[1.02]	[0.71]
(4)	0.83	0.78	0.88	1.02	1.12	0.30	0.28	-0.10	0.24	-0.13	0.10
						[1.79]	[1.69]	[-2.86]	[4.56]	[-2.54]	[2.64]
(H)	0.87	0.95	0.93	1.05	1.03	0.16	0.11	-0.05	-0.01	0.07	0.09
						[1.05]	[0.66]	[-1.46]	[-0.13]	[1.35]	[2.48]
All	0.50	0.51	0.62	0.77	0.82	0.31	0.29	-0.01	-0.01	0.08	0.00
						[3.37]	[3.00]	[-0.45]	[-0.39]	[2.65]	[0.18]

Panel B: Industry momentum conditional on peer momentum

Peer Momentum quintiles	Industry Momentum Quintiles (L)	(2)	(3)	(4)	(H)	Industry Momentum Strategies r^e	α_{FF4}	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}
(L)	0.28	0.41	0.21	0.29	0.46	0.19	0.26	0.00	-0.23	-0.12	0.03
						[1.01]	[1.38]	[0.05]	[-3.86]	[-2.05]	[0.65]
(2)	0.66	0.34	0.50	0.65	0.76	0.10	0.15	-0.12	-0.26	0.03	0.11
						[0.55]	[0.79]	[-3.21]	[-4.37]	[0.47]	[2.57]
(3)	0.65	0.59	0.66	0.64	0.98	0.32	0.40	-0.10	-0.14	-0.05	0.05
						[1.91]	[2.32]	[-2.98]	[-2.49]	[-0.97]	[1.32]
(4)	0.79	0.71	0.84	1.05	0.79	-0.00	0.07	-0.07	-0.23	-0.01	0.04
						[-0.01]	[0.36]	[-1.95]	[-3.90]	[-0.16]	[1.02]
(H)	0.69	1.01	0.94	0.96	0.98	0.30	0.20	0.02	-0.05	0.03	0.13
						[1.49]	[0.96]	[0.43]	[-0.76]	[0.44]	[2.69]
All	0.56	0.55	0.56	0.71	0.75	0.19	0.17	-0.06	-0.09	0.02	0.10
						[1.64]	[1.43]	[-2.39]	[-2.52]	[0.57]	[3.72]

Appendix

A.1 Network percentiles by industry

Table A.1: Network Estimation: Descriptive Statistics

This table reports summary statistics of composite centrality measure for each Fama-French industry. At the end of each network estimation period, we estimate the composite centrality measures of all firms in a given industry and calculate its cross-sectional mean, standard deviation, 25th and 75th percentiles. We repeat these calculations for each industry and each year from June 1928 to June 2015. The table reports the time series averages of these summary statistics.

Ind.	μ	σ	$p25$	$p75$	Ind.	μ	σ	$p25$	$p75$
Agric	0.16	0.09	0.08	0.23	Guns	0.16	0.1	0.08	0.25
Food	0.05	0.04	0.02	0.07	Gold	0.15	0.07	0.09	0.21
Soda	0.13	0.11	0.05	0.2	Mines	0.12	0.09	0.05	0.18
Beer	0.15	0.12	0.06	0.24	Coal	0.17	0.11	0.08	0.26
Smoke	0.17	0.12	0.07	0.27	Oil	0.04	0.03	0.01	0.05
Toys	0.09	0.07	0.03	0.13	Util	0.04	0.04	0.02	0.06
Fun	0.1	0.08	0.04	0.15	Telcm	0.06	0.05	0.02	0.09
Books	0.06	0.05	0.02	0.08	PerSv	0.08	0.06	0.03	0.11
Hshld	0.06	0.05	0.02	0.1	BusSv	0.03	0.03	0.01	0.04
Clths	0.07	0.07	0.02	0.12	Hardw	0.05	0.04	0.02	0.08
Hlth	0.06	0.04	0.03	0.08	Softw	0.02	0.02	0.01	0.03
MedEq	0.06	0.05	0.02	0.09	Chips	0.02	0.02	0.01	0.03
Drugs	0.05	0.04	0.02	0.08	LabEq	0.06	0.05	0.02	0.08
Chems	0.05	0.04	0.02	0.07	Paper	0.05	0.04	0.02	0.08
Rubbr	0.1	0.07	0.04	0.14	Boxes	0.1	0.08	0.04	0.15
Txtls	0.1	0.08	0.04	0.15	Trans	0.03	0.03	0.01	0.05
BldMt	0.04	0.04	0.01	0.06	Whlsl	0.05	0.04	0.02	0.07
Cnstr	0.07	0.06	0.03	0.1	Rtail	0.04	0.03	0.01	0.05
Steel	0.04	0.03	0.01	0.06	Meals	0.06	0.05	0.02	0.09
FabPr	0.15	0.11	0.07	0.23	Banks	0.05	0.04	0.01	0.07
Mach	0.04	0.03	0.01	0.06	Insur	0.04	0.03	0.01	0.05
ElcEq	0.07	0.06	0.02	0.1	RlEst	0.12	0.09	0.06	0.17
Autos	0.05	0.04	0.02	0.07	Fin	0.03	0.02	0.01	0.04
Aero	0.09	0.07	0.04	0.13	Other	0.1	0.08	0.03	0.14
Ships	0.16	0.11	0.07	0.24					

A.2 Portfolio construction

Table A.2: Portfolio weights used to construct peripheral peer momentum signals: An example

This table explains the construction of peripheral peer momentum strategy using the beer & liquor industry example from Table 1. For each firm in each industry, we track the past performance of its peers *except* those whose composite centrality measure is at the 90th percentile of the industry or above and form a peer impact portfolio. Specifically, we set the own impacts reported as diagonal elements in the network matrix as well as the impacts of the most central firm(s) in the industry (e.g. Anheuser Busch [ticker: Bud] in beer & liquor industry) to zero and rescale the other elements such that each row adds up to one. Multiplying the resulting portfolio weights by past one-month returns yields peripheral peer momentum signals. We implement these signal construction techniques to each industry at the end of June each year. Once we collect all signals, we sort stocks on them, form our peripheral peer momentum portfolios, and rebalance them monthly.

Portfolio weights used to construct peripheral peer momentum signals in beer & liquor industry from June 2006 to May 2007									
	COKE	MGPI	BUD	TAP	STZ	WVVI	BREW	SAM	PMID
COKE	0.00	0.23	0.00	0.28	0.17	0.03	0.02	0.15	0.12
MGPI	0.17	0.00	0.00	0.52	0.08	0.01	0.03	0.17	0.02
BUD	0.06	0.03	0.00	0.49	0.04	0.00	0.02	0.09	0.27
TAP	0.54	0.07	0.00	0.00	0.12	0.01	0.01	0.06	0.20
STZ	0.50	0.20	0.00	0.07	0.00	0.08	0.02	0.06	0.07
WVVI	0.25	0.06	0.00	0.12	0.28	0.00	0.26	0.03	0.01
BREW	0.20	0.23	0.00	0.01	0.02	0.36	0.00	0.05	0.14
SAM	0.88	0.02	0.00	0.01	0.02	0.01	0.02	0.00	0.04
PMID	0.75	0.01	0.00	0.19	0.01	0.00	0.03	0.03	0.00

A.3 Portfolio loadings

Table A.3: Loadings on Fama and French (2015) five-factor model
At the end of each month, we sort stocks into five portfolios based on their industry peers' past month average returns weighted by the peers' influence in the industry. Simple peer momentum strategy uses all peers' influence as weights for signal construction, whereas "peripheral peer momentum" strategy uses only peripheral stocks' influence as weights. All portfolios are constructed using NYSE breaks. For each of the five portfolios, and for a portfolio long stocks with high peer momentum signal and short stocks with low peer momentum signal, the table reports the loadings on Fama and French (2015) five-factors. Panel A (B) reports the statistics of simple (peripheral) peer momentum portfolios. T-statistics are in brackets. The sample period is July1963 to December 2015.

Panel A: Simple peer momentum-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
β_{MKT}	1.07	1.04	0.96	0.99	0.99	-0.08
	[46.89]	[67.79]	[66.01]	[65.39]	[46.34]	[-2.12]
β_{SMB}	-0.05	-0.03	-0.04	-0.03	-0.09	-0.04
	[-1.55]	[-1.59]	[-2.12]	[-1.57]	[-3.08]	[-0.76]
β_{HML}	0.04	0.09	0.11	0.03	-0.00	-0.04
	[0.84]	[3.01]	[3.86]	[1.10]	[-0.07]	[-0.52]
β_{RMW}	0.07	0.19	0.20	0.15	0.01	-0.06
	[1.45]	[6.11]	[6.84]	[4.93]	[0.22]	[-0.72]
β_{CMA}	-0.05	0.01	0.02	0.21	0.05	0.09
	[-0.72]	[0.25]	[0.39]	[4.71]	[0.72]	[0.81]
Panel B: Peripheral peer momentum-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
β_{MKT}	1.07	1.05	0.96	1.01	1.00	-0.07
	[46.78]	[71.24]	[65.94]	[68.70]	[46.93]	[-1.85]
β_{SMB}	-0.04	-0.09	0.00	-0.11	-0.07	-0.03
	[-1.19]	[-4.50]	[0.10]	[-5.24]	[-2.38]	[-0.59]
β_{HML}	0.03	0.10	0.11	0.08	0.00	-0.03
	[0.68]	[3.32]	[3.86]	[2.68]	[0.02]	[-0.38]
β_{RMW}	0.15	0.15	0.23	0.12	0.01	-0.14
	[3.16]	[4.91]	[7.76]	[3.79]	[0.27]	[-1.68]
β_{CMA}	-0.01	-0.02	0.04	0.09	0.06	0.08
	[-0.22]	[-0.56]	[0.83]	[2.12]	[0.97]	[0.65]

A.4 Additional conditional sorts

Table A.4: Conditional sort on classic momentum and peer momentum

This table presents results for conditional double sorts on classic momentum and peer momentum. In each month, firms are sorted first into quintiles based on classic momentum, then on peer momentum. Then, they are grouped into twenty-five portfolios based on the intersection of the two sorts. Panel A presents the average returns to the 25 portfolios, panel B documents the average number of firms and the average firm size for each portfolio. Time period is 07/1928 to 12/2015.

Panel A: portfolio average returns and time-series regression results

	Peer Momentum Quintiles					Peer Momentum Strategies					
Momentum quintiles	(L)	(2)	(3)	(4)	(H)	r^e	α_{FF4}	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}
(L)	-0.00	0.43	0.23	0.31	0.63	0.63	0.43	-0.12	0.35	0.19	0.19
						[2.78]	[1.85]	[-2.67]	[4.77]	[2.71]	[3.54]
(2)	0.13	0.48	0.65	0.73	0.87	0.76	0.62	-0.02	0.02	0.17	0.12
						[4.63]	[3.68]	[-0.72]	[0.28]	[3.47]	[2.98]
(3)	0.07	0.42	0.70	0.84	0.86	0.79	0.76	-0.00	-0.09	0.13	0.01
						[5.29]	[4.89]	[-0.09]	[-1.93]	[2.85]	[0.22]
(4)	0.36	0.66	0.83	0.84	1.25	0.89	0.74	0.05	-0.07	0.17	0.08
						[5.86]	[4.76]	[1.75]	[-1.34]	[3.67]	[2.32]
(H)	0.80	0.93	1.06	1.10	1.29	0.49	0.40	-0.03	-0.05	0.08	0.14
						[2.69]	[2.10]	[-0.75]	[-0.84]	[1.39]	[3.13]

Panel B: Portfolio average number of firms and market capitalization

	Peer Momentum Quintiles: Average n					Peer Momentum Quintiles: Average market capitalization ($\$10^6$)				
Momentum quintiles	(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)
(L)	52	51	50	50	53	750	687	658	689	674
(2)	43	42	42	41	42	1403	1287	1237	1330	1348
(3)	41	40	40	39	40	1557	1549	1596	1568	1630
(4)	41	39	39	40	40	1653	1552	1594	1570	1833
(H)	47	45	44	45	46	1418	1274	1335	1449	1613

A.5 Spanning tests

Table A.5: Spanning tests

The table shows the results of time series regressions of alternative winner-minus-loser momentum portfolios on each other and Fama and French (2015) five-factors. The alternative momentum strategies are peripheral peer momentum, industry momentum and classic momentum. The sample period is 07/1928 to 12/2015 except for regressions including Fama and French (2015) five-factors whose sample is available only after July 1963.

Coef.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Peer Momentum as a dependent variable			Industry Momentum as a dependent variable			Momentum as a dependent variable		
Intercept	0.86	0.31	0.16	0.79	0.11	0.18	0.79	0.71	0.66
	[6.53]	[3.36]	[1.49]	[5.77]	[1.13]	[1.45]	[4.16]	[3.62]	[3.03]
Per. Peer Mom.					0.78	0.85		0.03	0.22
					[35.58]	[28.33]		[0.43]	[2.72]
Ind. Mom.		0.70	0.66					0.03	0.09
		[35.58]	[28.33]					[0.51]	[1.29]
Mom.		0.01	0.05		0.01	0.03			
		[0.43]	[2.72]		[0.51]	[1.29]			
MKT			-0.01			-0.02			-0.18
			[-0.25]			[-0.55]			[-3.44]
SMB			0.03			-0.07			0.04
			[0.86]			[-1.74]			[0.53]
HML			0.05			-0.03			-0.56
			[0.95]			[-0.53]			[-5.49]
CMA			-0.12			0.05			0.40
			[-2.26]			[0.85]			[3.71]
RMW			-0.06			0.09			0.46
			[-0.85]			[1.09]			[3.01]
n	1050	1050	630	1079	1050	630	1069	1050	630
$\bar{R}^2(\%)$	0	55	59	0	55	59	0	-0	12

A.6 Robustness to strategy construction

Table A.6: Conditional sort on market equity and peer momentum

This table presents results for conditional double sorts on market equity and peer momentum. In each month, firms are sorted first into quintiles based on size, then on peer momentum. Then, they are grouped into twenty-five portfolios based on the intersection of the two sorts. Panel A presents the average returns to the 25 portfolios, panel B documents the average number of firms and the average firm size for each portfolio. Time period is 07/1928 to 12/2015.

Panel A: portfolio average returns and time-series regression results												
	Peer Momentum Quintiles					Peer Momentum Strategies						
	(L)	(2)	(3)	(4)	(H)	r^e	α_{FF4}	β_{MKT}	β_{SMB}	β_{HML}	β_{UMD}	
Size quintiles (L)	0.84	1.12	1.00	1.22	1.59	0.75	0.83	-0.27	0.06	-0.04	0.13	
						[2.59]	[2.80]	[-4.56]	[0.70]	[-0.46]	[1.92]	
(2)	0.51	0.56	0.90	1.21	1.23	0.73	0.88	-0.11	-0.23	-0.16	0.04	
						[3.43]	[4.06]	[-2.49]	[-3.42]	[-2.39]	[0.87]	
(3)	0.47	0.80	1.04	1.21	1.07	0.60	0.61	-0.10	-0.02	-0.06	0.12	
						[3.30]	[3.26]	[-2.74]	[-0.33]	[-1.08]	[2.87]	
(4)	0.39	0.67	0.91	1.11	1.12	0.73	0.63	-0.01	0.02	0.22	0.01	
						[4.82]	[4.10]	[-0.40]	[0.42]	[4.68]	[0.34]	
(H)	0.18	0.54	0.66	0.74	1.00	0.82	0.68	0.02	-0.02	0.18	0.09	
						[6.07]	[4.87]	[0.82]	[-0.39]	[4.24]	[2.89]	
Panel B: Portfolio average number of firms and market capitalization												
	Peer Momentum Quintiles					Peer Momentum Quintiles						
	Average n					Average market capitalization ($\$10^6$)						
	(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)		
Size quintiles (L)	77	75	72	72	78	92	95	91	92	92		
(2)	37	36	36	36	38	167	167	167	167	171		
(3)	34	33	33	33	35	304	303	301	303	314		
(4)	35	34	34	35	35	709	691	690	700	718		
(H)	39	38	38	38	39	5483	5234	5190	5226	5788		

Table A.7: Robustness to strategy construction

At the end of each month, we construct one-way sorted peripheral peer momentum portfolios and a zero-cost portfolio long stocks with high peer momentum signal and short stocks with low peer momentum signal. Panel A reports results using equal-weighted portfolios, constructed from a quintile sort with NYSE stock breakpoints. Panel B reports results using value-weighted portfolios, constructed from a quintile sort with all stock breakpoints. Panel C reports results using value-weighted portfolios, constructed from a decile sort with NYSE breakpoints. For all of these portfolios, the table reports average returns in excess of the risk-free rate and alphas with respect to the CAPM, Fama and French (1993) three-factor model, Fama and French (1993) three-factor model augmented with the Carhart (1997) momentum factor, and the Fama and French (2015) five-factor model. T-statistics are in brackets. The sample period is 07/1928 to 12/2015 except for Fama and French (2015) five-factor model whose sample is available only after July 1963.

Panel A: Quintile sort, equal-weighted, NYSE breaks						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.28	0.63	0.89	1.04	1.27	0.99
	[1.21]	[2.79]	[4.08]	[4.71]	[5.47]	[7.94]
α^{CAPM}	-0.51	-0.15	0.13	0.27	0.48	0.99
	[-5.24]	[-1.91]	[1.69]	[3.48]	[5.06]	[7.91]
α^{FF3}	-0.59	-0.27	0.03	0.15	0.36	0.95
	[-6.88]	[-4.09]	[0.40]	[2.37]	[4.48]	[7.61]
$\alpha^{\text{FF3+UMD}}$	-0.39	-0.12	0.15	0.28	0.46	0.85
	[-4.59]	[-1.87]	[2.31]	[4.48]	[5.60]	[6.61]
α^{FF5}	-0.67	-0.25	0.09	0.20	0.51	1.18
	[-6.44]	[-3.56]	[1.46]	[3.22]	[6.01]	[7.88]
Panel B: Quintile sort, value-weighted, all stock breaks						
r^e	0.17	0.51	0.72	0.78	1.03	0.86
	[0.86]	[2.80]	[3.99]	[4.21]	[5.11]	[6.36]
α^{CAPM}	-0.50	-0.13	0.08	0.13	0.34	0.85
	[-6.33]	[-2.29]	[1.43]	[2.10]	[4.21]	[6.22]
α^{FF3}	-0.49	-0.14	0.06	0.10	0.32	0.81
	[-6.15]	[-2.40]	[1.14]	[1.69]	[3.95]	[5.95]
$\alpha^{\text{FF3+UMD}}$	-0.39	-0.11	0.09	0.11	0.31	0.70
	[-4.83]	[-1.85]	[1.63]	[1.87]	[3.72]	[5.04]
α^{FF5}	-0.47	-0.16	0.16	0.10	0.29	0.76
	[-4.80]	[-2.63]	[2.86]	[1.58]	[3.22]	[4.52]

Table 3 (Continued): Robustness to strategy construction

Panel C: Decile sort, value-weighted, NYSE breaks											
	(L)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(H)	(H-L)
r^e	0.12	0.25	0.46	0.67	0.65	0.73	0.88	0.77	1.04	1.06	0.95
	[0.57]	[1.26]	[2.28]	[3.68]	[3.50]	[3.91]	[4.69]	[3.80]	[4.91]	[5.15]	[5.77]
α^{CAPM}	-0.55	-0.42	-0.22	0.04	0.02	0.09	0.24	0.09	0.34	0.38	0.94
	[-5.43]	[-5.13]	[-2.76]	[0.61]	[0.28]	[1.22]	[3.17]	[1.00]	[3.51]	[3.94]	[5.68]
α^{FF3}	-0.53	-0.44	-0.25	0.02	-0.01	0.06	0.21	0.03	0.29	0.36	0.89
	[-5.19]	[-5.35]	[-3.05]	[0.37]	[-0.17]	[0.86]	[2.89]	[0.41]	[3.12]	[3.69]	[5.38]
$\alpha^{\text{FF3+UMD}}$	-0.44	-0.37	-0.18	0.03	0.03	0.09	0.22	0.08	0.27	0.33	0.77
	[-4.20]	[-4.42]	[-2.14]	[0.42]	[0.43]	[1.24]	[2.89]	[0.89]	[2.84]	[3.35]	[4.57]
α^{FF5}	-0.55	-0.42	-0.27	-0.02	0.04	0.14	0.14	0.04	0.21	0.34	0.89
	[-4.55]	[-4.57]	[-3.27]	[-0.27]	[0.55]	[1.82]	[1.97]	[0.50]	[2.48]	[3.04]	[4.44]

A.7 Robustness to other anomalies

Table A.8: Robustness to other anomalies: Double sorts

The table reports average excess returns for conditional peripheral peer momentum (PPM) strategies, constructed from double-sorts on each of the twenty-three anomaly signals from Novy-Marx and Velikov (2016) and PPM. In each month, we sort all firms in the CRSP/COMPUSTAT merged database into quintiles using the signal for one of the twenty-three anomalies. Then, within each quintile, we sort stocks into quintiles depending on their PPM. Firms are grouped into five PPM portfolios by combining the firms across the characteristic quintiles. The table reports value-weighted average excess returns for the five PPM portfolios and for a portfolio that is long stocks in the high PPM portfolio and short stocks in the low PPM portfolio. T-statistics are in brackets.

Panel A: Excess returns on conditional ISS_PAYNORM-sorted portfolios

Anomaly	Size	Gross Profitability	Value	ValProf	Accruals	Asset Growth	Investment	Piotroski F-score	Net Issuance (M)	ROE	Failure Probability
(L)	0.18	0.33	0.26	0.30	0.07	0.28	0.29	0.24	0.21	0.18	0.13
	[0.93]	[1.79]	[1.32]	[1.64]	[0.35]	[1.54]	[1.59]	[1.00]	[1.07]	[0.78]	[0.54]
(2)	0.56	0.57	0.54	0.54	0.51	0.54	0.53	0.50	0.57	0.50	0.44
	[3.01]	[3.39]	[2.98]	[3.26]	[2.58]	[3.26]	[3.14]	[2.32]	[3.10]	[2.38]	[2.05]
(3)	0.69	0.72	0.68	0.74	0.60	0.71	0.73	0.77	0.66	0.65	0.69
	[3.99]	[4.51]	[3.73]	[4.62]	[3.21]	[4.44]	[4.51]	[3.78]	[3.73]	[3.33]	[3.44]
(4)	0.76	0.74	0.80	0.66	0.56	0.69	0.70	0.65	0.85	0.69	0.65
	[4.16]	[4.70]	[4.30]	[4.16]	[3.08]	[4.32]	[4.37]	[3.16]	[4.39]	[3.42]	[3.20]
(H)	1.01	0.85	0.98	0.88	0.78	0.86	0.88	0.82	0.98	0.79	0.82
	[5.21]	[5.05]	[5.12]	[5.19]	[3.93]	[5.15]	[5.12]	[3.74]	[5.06]	[3.62]	[3.74]
(L-H)	0.83	0.53	0.72	0.58	0.70	0.59	0.59	0.58	0.77	0.62	0.69
	[6.36]	[3.95]	[5.87]	[4.30]	[4.19]	[4.42]	[4.33]	[3.13]	[6.23]	[3.55]	[3.94]

Anomaly	ValMomProf	ValMom	Idiosyncratic Volatility	Momentum	PEAD (SUE)	PEAD (CAR3)	Industry Momentum	Industry Relative Reversals	High-frequency Combo	Short-run Reversals	Seasonality	IRR (LowVol)
(L)	0.29	0.34	0.26	0.25	0.25	0.14	0.50	0.22	0.42	0.19	0.28	0.33
	[1.62]	[1.88]	[1.54]	[1.32]	[1.03]	[0.60]	[2.68]	[1.12]	[2.11]	[1.00]	[1.44]	[1.75]
(2)	0.53	0.56	0.55	0.50	0.55	0.45	0.51	0.54	0.60	0.46	0.60	0.57
	[3.17]	[3.45]	[3.42]	[2.77]	[2.49]	[2.09]	[2.84]	[2.96]	[3.27]	[2.55]	[3.33]	[3.25]
(3)	0.78	0.77	0.73	0.73	0.74	0.76	0.62	0.69	0.70	0.78	0.75	0.68
	[4.79]	[4.72]	[4.81]	[4.11]	[3.62]	[3.83]	[3.42]	[3.89]	[3.85]	[4.39]	[4.20]	[3.88]
(4)	0.72	0.71	0.79	0.78	0.70	0.65	0.77	0.78	0.72	0.77	0.83	0.84
	[4.50]	[4.51]	[4.94]	[4.25]	[3.43]	[3.18]	[4.25]	[4.29]	[3.98]	[4.14]	[4.39]	[4.66]
(H)	0.85	0.88	0.85	1.02	0.80	0.74	0.82	1.01	0.85	1.06	1.00	1.00
	[5.08]	[5.18]	[5.27]	[5.20]	[3.59]	[3.37]	[4.40]	[5.10]	[4.46]	[5.47]	[5.19]	[5.23]
(L-H)	0.56	0.54	0.59	0.76	0.55	0.60	0.31	0.79	0.44	0.87	0.71	0.67
	[4.21]	[4.09]	[5.10]	[6.12]	[3.04]	[3.42]	[3.37]	[5.96]	[3.58]	[7.61]	[5.88]	[5.37]

A.8 Subsample Analysis

Table A.9: Subsample Analysis
At the end of each month, we sort stocks into five peripheral peer momentum (PPM) portfolios using NYSE breaks. For each of the five portfolios, and for a portfolio long stocks with high PPM and short stocks with low PPM, we calculate average value-weighted returns in excess of the risk-free rate and alphas with respect to the CAPM, Fama and French (1993) three-factor model, Fama and French (1993) three-factor model augmented with the Carhart (1997) momentum factor. Panel A reports the results for the early subsample (1928-1963), and Panel B reports the results for the late subsample (1963-2015). T-statistics are in brackets.

Panel A: Excess returns and alphas on peripheral peer momentum-sorted portfolios: 1928-1963						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.32	0.71	0.73	1.01	1.41	1.09
	[0.84]	[2.00]	[2.01]	[2.68]	[3.44]	[4.60]
α^{CAPM}	-0.56	-0.14	-0.13	0.11	0.46	1.02
	[-4.03]	[-1.24]	[-1.19]	[0.96]	[3.01]	[4.31]
α^{FF3}	-0.55	-0.14	-0.14	0.10	0.45	0.99
	[-3.93]	[-1.25]	[-1.23]	[0.86]	[2.95]	[4.22]
$\alpha^{\text{FF3+UMD}}$	-0.49	-0.12	-0.09	0.15	0.51	1.00
	[-3.43]	[-1.08]	[-0.77]	[1.32]	[3.26]	[4.11]
Panel B: Excess returns and alphas on peripheral peer momentum-sorted portfolios: 1963-2015						
r^e	0.09	0.44	0.67	0.65	0.80	0.71
	[0.43]	[2.33]	[3.85]	[3.61]	[4.21]	[4.64]
α^{CAPM}	-0.43	-0.06	0.21	0.17	0.32	0.75
	[-4.78]	[-1.01]	[3.39]	[2.77]	[3.81]	[4.86]
α^{FF3}	-0.43	-0.08	0.17	0.15	0.32	0.75
	[-4.71]	[-1.27]	[2.75]	[2.46]	[3.78]	[4.78]
$\alpha^{\text{FF3+UMD}}$	-0.33	-0.04	0.18	0.13	0.24	0.57
	[-3.57]	[-0.61]	[2.93]	[2.17]	[2.86]	[3.63]
α^{FF5}	-0.48	-0.12	0.08	0.09	0.30	0.78
	[-5.11]	[-2.05]	[1.31]	[1.42]	[3.48]	[4.83]