Competition links and stock returns

Assaf Eisdorfer, Kenneth Froot, Gideon Ozik, Ronnie Sadka*

April 2019

ABSTRACT

We implement a PageRank-type algorithm on the competitor-links across 10-Ks to produce a dynamic measure of firm competitiveness. We call this measure 'C-Rank'. We find that firm C-Rank has a positive effect on subsequent stock returns, controlling for firm size and other common risk factors. A hedge high-minus-low C-Rank portfolio yields an annual 6-factor alpha of about 16 percent. This effect can potentially be explained by the high risk that a firm faces when it is targeted as a competitor by many strong companies. We further find that managers often choose to list strong firms as competitors even if they are less related to their company.

Keywords: Text analysis; Competition; Asset pricing

JEL Classifications: G12, G14

---

* Eisdorfer: University of Connecticut, email: assaf.eisdorfer@uconn.edu. Froot: Harvard Business School, email: kfroot@hbs.edu. Ozik: EDHEC, email: gideon.ozik@edhec.edu. Sadka: Boston College, email: ronnie.sadka@bc.edu.

1. Introduction

The information content of corporate financial statements has been explored for many years. Starting with the early studies on the effect of unexpected net income on stock returns (e.g., Ball and Brown (1968) and Beaver, Clarke, and Wright (1979)), through a deeper look at the accounting numbers such as discretionary accruals (e.g., Sloan (1996)), to the resent more sophisticated learning procedures such as text analysis (e.g., Hoberg and Phillips (2010, 2016) and Cohen, Malloy, and Nguyen (2018)). The common objective of these studies is to assess what investors can learn about a certain company from the information embedded in its financial statements.

In this study we move a step forward and ask what we can learn about a company from the point of view of *its competitors*. Financial statements typically include a competition section in which the company can list and discuss the firms it views as competitors. A company that is mentioned as a competitor by many other companies is more likely to be a strong competitor. Yet counting the number of such mentions is not sufficient, as not all mentions created equal: a mention by a major competitor should count more than a mention by a minor competitor. Weighting these mentions by competition strength means therefore that the competition status of all firms should be evaluated simultaneously. We use an advanced text analysis technology to identify the competitor lists in the cross-section of financial statements. We implement a PageRank-type algorithm in the manner of Page et al. (1999) on the cross-mentions in the most recent annual reports of all firms in each month to produce a firm-level competition rank (referred to as 'C-Rank').

A unique feature of our competition C-Rank measure is that it is not an independent assessment based on observed firm characteristics, such as firm size or product market share, but rather it reflects the view of all companies on who are the strong competitors in the market. This feature therefore has the potential to impose an element of risk on a competitive firm. If many strong companies identify a certain firm as a competitor, they likely to enhance their strategies to compete with that firm. A high C-Rank of a firm may thus suggest that the firm is "targeted" by strong competitors, which can increase the uncertainty about the firm's future performance and value. If this form of risk is recognized by the market, it should be compensated by high stock returns.

Our first test therefore addresses the predictive ability of the C-Rank competition status over future stock returns. Due to the positive correlation between C-Rank and firm size, we first eliminate the size effect on stock returns. We run monthly cross-sectional regressions of C-Rank

on current firm size, and use the regression residuals as our sorting variable. This residual thus captures the firm's level of competitiveness relative to where it is "expected" to be according to its size. We find that the C-Rank residual positively predicts stock returns; an investment strategy that buys high C-Rank stocks and shorts low C-Rank stocks generates a monthly 6-factor alpha of 1.35 percent (about 16 percent in annual terms). This effect is robust to various subsamples and investment horizons, and is confirmed using Fama-MacBeth (1973) regressions.

This positive association between firm competition status and stock returns supports the "targeting competitors" risk story outlined above. Yet we acknowledge that the relatively high stock returns earned by high C-Rank firms may also represent a temporary mispricing, particularly because obtaining the C-Rank is not an easy task for the retail investor. We do not have however a reasonable explanation for why high C-Rank firms will exhibit a different mispricing pattern than low C-Rank firms. In addition, observing stock returns in the few months before and after the financial statements' filings does not point to stock mispricing.

It is important to note that while larger firms are likely to be stronger competitors, our competition measure contains relevant information that is not captured by firm size. First, in all our tests of the information content of C-Rank we control for the effect of firm size. Second, we show that the largest companies in the market are not necessarily the stronger competitors. And moreover, we find that high competition status is associated with high returns, which goes against the negative effect of firm size on stock returns.

We produce also two industry-based forms of C-Rank; the first one is 'cross-industry', which considers only mentioned firms that operate in different industries than the filing firm; and the second is 'within-industry', which considers only mentioned firms from the same industry as the filing firm. The results based on cross-industry C-Rank are similar to those based on the full market C-Rank. However, the effect of within-industry competitiveness on stock returns are insignificant and even negative under some models. This may suggest that the firm's real competition status is generated mostly by competing with companies outside the industry.

We next turn our focus to firms' strategies in listing competitors in their 10-Ks. Assigning a competition rank to every firm in the market allows us also to assess what any specific firm "thinks" about its competition strength. Specifically, the firm's own view on where it stands in the market competition can be given by the average C-Ranks of the companies it mentions as competitors. Comparing the firm's own view of its competition strength with the market view (the

firm's C-Rank) allows us to explore whether firms tend to over-position themselves in the market competition.

We find a clear negative relation between the strength of the competitors mentioned in the reports and the similarity between them and the filing firms. Particularly, when a strong company is mentioned as a competitor in the financial report, it is less likely to produce similar products than the filing firm, it is less likely to operate in the same industry as the filing firm, and it is less likely to mention back the filing firm as a competitor. This suggests that firms often inflate their competition status by listing strong companies as competitors even though they are not their real competitors.

We further assess whether highly self competition view leads indeed to a better competition status. As expected, the competition status of the filing firms is more likely to improve from one filing year to the next when the strong companies mentioned as competitors are fairly similar to the filing firm. This further indicates that firms often associate themselves with strong and unrelated companies just to enjoy the benefits coming from an enhanced competition status.

Our paper fits in several streams of literature. We first contribute to the literature that uses an advanced text analysis technology to identify relevant information. Hoberg and Phillips (2010, 2016) apply text analysis on the product description section in 10-Ks to measure the similarity between the products of each pair of companies. Garcia and Norli (2012) extract U.S. state name counts from annual reports and find that less state-diversified firms earn higher stock return. Cohen, Malloy, and Nguyen (2018) show that when companies make an active change in their reporting practices, this conveys an important signal about future firm operations.

A closely related study is Li, Lundholm, and Minnis (2013); using a text analysis they measure the competitiveness of a firm by the ratio of the number of competition-related words to the total number of words in the report. The innovation of our measure over that of Li, Lundholm, and Minnis (2013) is twofold. First, as discussed above, a key aspect of our measure is that the information about a firm's competition status is not limited to the financial report of the firm itself, but rather it is gathered from the reports of its competitors. Second, we implement a ranking algorithm to simultaneously capture the market's view on the competition strength of each specific company at any given time.

Our paper also contributes to the asset pricing literature. We identify a firm-specific characteristic, C-Rank, that has a predictive ability over future returns, controlling for well-

documented return predictors such as size, value, momentum, and others. We provide a potential rational risk-based explanation for this effect, where additional evidence also points out to this direction and less to a mispricing explanation.

Finally, our results can be informative with respect to managerial incentives. There is a long line of literature that explores managerial manipulation of accounting numbers driven by incentives related to compensation considerations, debt contracts, political benefits, and others (see, for example, Healy (1985), Liberty and Zimmerman (1986), Healy, Kang, and Palepu (1987), Jones (1991), Holthausen, Larcker, and Sloan (1995), Perry and Williams (1994), and Sweeney (1994)). Our results indicate that managers can also manipulate the firm's competition status to enhance the firm's reputation and value. This behavior also fits managerial empire-building interests (see Jensen (1986)), as managers enjoy the benefits from running a company that is perceived as a major competitor.

2. Generating competition rank

Our procedure to measure the relative competition strength of a firm is based on the entire cross-sectional pool of financial statements. Each month we observe the most recent annual report of each firm (if available). That is, all companies are represented in the competition measurement in each month. Using an advanced technology of text analysis, we record the companies that are listed as competitors by each firm (the appendix describes the text analysis procedure). This provides a system of links between all firms in each month, where each firm can both mention other firms and be mentioned by other firms.

We observe a total of 119,785 10-Ks filed by 11,304 firms over the period 1995-2017, out of which 68,952 reports (58%) include a competition section. The number of firms that are mentioned as competitors in a single report's competition section ranges between zero (61 percent of the reports) to 35. Also, most firms, about 69 percent, are not mentioned at all in other reports. The company with the most mentions at a given time is IBM Corp. ('IBM') which was recognized as a competitor by 136 companies in the annual reports filed during 1997, followed by Microsoft Corp. ('MSFT') with 113 mentions in the reports filed during 1999. Figure 1 presents the distributions of the number of firms mentioned as competitors in a report's competition section, and the number of reports in which a firm is mentioned as a competitor.

Our measure is based on the notion that the competition status of a firm is determined not only by the number of times that the firm is mentioned by other firms, but also by the competition status of the mentioning firms. That is, the competition status of all firms should be evaluated simultaneously. For that matter we employ the PageRank algorithm developed by the founders of Google, Larry Page and Sergey Brin (Page et al., (1999)).

The basic aim of the PageRank algorithm is to assess the importance of a website page by counting the number and quality of links to the page, assuming that more important websites are likely to receive more links from other websites. In the same way we assume that more competitive companies are likely to be mentioned as competitors by other companies. Each month we run a PageRank-type algorithm that solves a system of n (=number of reports) simultaneous equations to produce a firm-level competition status. Referred to this measure as 'C-Rank'.[1] As stated above, to make sure that all firms participate in calculating monthly C-Ranks, for each month we use the most recent annual report of every firm over the past twelve months. For example, the C-Ranks for April 2005 are based on all annual reports from May 2004 to April 2005.

The primary C-Rank measure is based on the cross-section of all firms in the market. We also produce two alternative industry-related measures, based on the eleven GIC industry classifications. The first measure considers only competitors outside the industry of the filing firm (cross-industry measure), and the second measure considers only competitors from the same industry of the filing firm (within-industry measure).

3. C-Rank distribution and correlations

Our procedure produces the three C-Rank measures for each firm in each month over the period 1995-2017. Our sample includes 1,664,271 firm-month C-Rank values. Table 1 presents key statistics of the C-Rank distribution. The C-Rank values of the full market and cross-industry are substantially lower than that of the within-industry one. This is because the latter uses a fewer firms in the C-Rank algorithm thus assigns a relatively higher value to each firm. All C-Rank means are higher than the medians, indicating a positively skewed distribution. This makes sense in a competitive environment as a small number of firms are perceived as the lead competitors while most firms are not considered as major competitors. On a monthly average, about 60 percent

[1] We use a damping factor of 0.7 in applying the algorithm. The results remain similar when using different damping factors between 0.5 to 0.9. See Page et al. (1999) for details on the PageRank algorithm.

of the firms are not mentioned by any other firm, and thus received the lowest C-Rank in a given month. Not surprisingly, the three C-Rank measures are positively correlated; the full market and cross-industry C-Ranks exhibit a higher correlation coefficient, 0.59, whereas the correlations between these measures and the within-industry C-Rank are 0.20 and 0.27. These gaps in correlations are expected as the sample of firms used to generate the full market and cross-industry C-Ranks are more overlapped than with the sample used for the within-industry C-rank. These correlation values indicate that each form of C-Rank captures also different aspects of competition.

A reasonable question regarding our C-Rank measure is whether it simply captures the market capitalization of the firm, as larger firms are typically more competitive. We confirm that C-Rank captures information that is incremental to firm size in several ways. We show that the largest companies in the market are not necessarily the stronger competitors. We control for the effect of firm size in all of our tests. And moreover, we find that high competition status is associated with high returns, which goes against the negative effect of firm size on stock returns.

Table 2 presents the top five firms by the full market C-Rank every year over the period 1995-2017, as well as the largest companies over the same period. The C-Rank leaders show a remarkable pattern where from 1995 to 2012 IBM and Microsoft were the only companies with top C-Ranks, and since 2013 Google LLC ('GOOGL', now Alphabet Inc.) gets the top C-Rank every year. Comparing the top competitors to the list of the largest companies indicate indeed that very large firms are likely to be also very strong competitors, as several firms appear in both lists. Yet this association does not seem extremely strong as some of the largest firms in the market do not have the equivalent competition status. For example, General Electric Co. ('GE') had the largest market capitalization in eight years between 1995 and 2005, yet it was in the top five competitors only once during these years, in 1995. Similarly, Exxon Mobil Corp. ('XOM') had been constantly the largest company between 2006 and 2011, but was never in the group of the top five competitors. This provides a first indication that our competition measure contains information that is not captured by firm size.

We further calculate the correlation between C-Rank and firm size and other common risk factors: market-to-book ratio, past stock return, profitability, investment intensity, market beta, and idiosyncratic volatility. All market and accounting data are obtained from CRSP and COMPUSTAT. Because most firms are not mentioned as competitors in any report, we show the correlations both for the full sample and for the sample of competitive firms only (those with at

least one mention in other reports). To eliminate time effects, we compute the cross-sectional correlations for each month over the sample period and report the time-series averages in Table 3. As expected, there is a positive correlation between C-Rank and firm size: 0.24 to 0.58. This is consistent with the results in Table 2, indicating that competition represents a firm characteristic that is not entirely captured by the size of the firm. High C-Rank firms are also more profitable and with lower idiosyncratic volatility, yet all average correlations for these and the other characteristics are fairly low. This suggests that C-Rank is not likely representing any of these risk factors.

4. C-Rank and stock returns

An important feature of our measure of the competitiveness status of a firm, the C-Rank, is that it is not an independent assessment based on observed firm-specific characteristics, such as firm size or product market share, or even the competitive nature of the text in the firm's 10-K (Li, Lundholm, and Minnis (2013)). Rather, C-Rank reflects the view of the entire cross-section of companies on who are the strong competitors in the market. This feature therefore has the potential to impose an element of risk on a competitive firm. If a certain firm is identified as a competitor by many strong companies, then it puts the identified firm in a risky position as the other companies likely to enhance their strategies to compete with that firm. That is, a high C-Rank of a firm may suggest that the firm is "targeted" by strong competitors, which likely increase the uncertainty about the firm's future performance and value. If this form of risk is recognized by the market, we should expect a stock return premium on high C-Rank firms.

4.1 Portfolio sorts

We first examine the association between C-Rank and future stock returns using portfolio sort analysis. Due to the positive correlation between C-Rank and firm size, we first eliminate the size effect on stock returns. We run monthly cross-sectional regressions of C-Rank as of three months earlier (assuming it takes three months to release financial reports) on current firm size, and use the regression residuals as our sorting variable. Each month over the period 1995-2017 we divide all stocks into five equal-sized portfolios according to their C-Rank-residual. The portfolios are equal-weighted and held for one month.

Table 4 reports the monthly returns on each portfolio as well as the returns to the hedge portfolio that is long the highest C-Rank-residual firm portfolio and short the lowest C-Rank-residual firm portfolio. In addition to reporting the average return in excess of the risk-free rate, we also report the alphas from factor models. All factor returns are downloaded from Ken French's website. All returns and alphas are in percent per month and numbers in parentheses denote the corresponding *t*-statistics. Panels A, B, and C report the results for the full market, cross-industry, and within-industry C-Ranks, respectively.

The full market C-Rank's results in Panel A show that returns and factor-model alphas are generally monotonically increasing as one moves from quintile 1 (least competitive stocks) to quintile 5 (most competitive stocks). The long/short hedge portfolio has an excess return of 0.93% per month (*t*-statistic=3.76). Factor-model alphas are consistent with the excess return, ranging between 0.77% (CAPM) to 1.35% (6-factor model), all are statistically significant (*t*-statistics between 3.17 and 7.00). The results in Panel A therefore uncover a clear strong relation between the firm's competitiveness and subsequent returns.

Panels B and C reveals interesting results with regard the industry role of firm competitiveness. The results in Panel B based on cross-industry C-Rank are similar to those based on the full market C-Rank; the monthly returns/alphas are between 0.78% to 1.30% (*t*-statistics between 2.72 and 5.22). However, the effect of within-industry competitiveness on stock returns (Panel C) are insignificant and even negative under some models. This may suggest that the firm's real competition status is generated mostly by competing with companies outside the industry.

The upper graph in Figure 2 shows the performance of the C-Rank-residual hedge portfolio over the period 1995-2017. While the effect seems somewhat stronger in the early years, it is consistently upward sloping over the sample period, yielding a cumulative excess return of 253% and 6-factor alpha of 369%.

We further examine the robustness of the results to different subsamples and return horizons in Table 5. To reduce the clutter in the table, we report only the 6-factor alphas for each portfolio. To facilitate comparison with the main results, we also report the full-sample results in the first row of the table. We consider three different kinds of subsamples. The first simply tabulates results when excluding the month of January. The second subsample excludes recession periods. We use NBER recession dummy as an indicator of the health of the economy for this exercise. Third, we tabulate the results separately for the early years (1995-2006) and the late years (2007-2017).

Panel A of the table shows that the hedge portfolio alpha is somewhat lower when excluding January, but is still significant; the 6-factor alphas is 0.80% with a *t*-statistic of 5.01. The results seem insensitive to the state of the economy, as excluding recessions shows a significant 6-factor alphas of 1.39%. Consistent with Figure 2, the effect of C-Rank is stronger in the early years, yielding an alpha of 1.96% per month (*t*-statistic 5.76), yet is still significant in the late years with an alpha of 0.77% and *t*-statistic of 4.48.

We look at the horizon effects in Panel B of Table 5. Specifically, we consider holding periods of 3, 6, 12, and 18 months. This means that we have overlapping portfolios. We take the equal-weighted average of these overlapping portfolios similar to the approach of Jegadeesh and Titman (1993). The 6-factor alphas of the hedge portfolio are positive and statistically significant for horizons up to 18 months, although they decline monotonically as we increase the horizon, from 1.35% for one-month horizon to 0.91% for 18-month horizon. All portfolio sort results are therefore robust to different subsamples and horizons.

4.2 Fama-MacBeth regressions

We further examine the association between C-Rank and subsequent returns using Fama and MacBeth (1973) regressions. Beyond serving as an additional diagnostic check, these regressions offer the advantage of controlling for other well known determinants of the cross-sectional patterns in returns and thus check for the marginal influence of C-Rank on our results. Accordingly, we run these cross-sectional regressions and report the results in Table 6. The dependent variable is the excess stock return and the main independent variables is C-Rank. We use the C-Rank value and not the C-Rank/size residual as in the portfolio sort analysis because we can control for size directly in the regressions. Yet because the cross-sectional distributions of C-Rank are not identical (i.e., they depend on the number of reports available at each month), we standardize the monthly C-Rank values. The control variables are log market capitalization, log market-to-book, past six-month return, profitability, investment intensity, market beta, and idiosyncratic volatility. We winsorize all independent variables at the 1% and 99% levels to reduce the impact of outliers. All reported coefficients are multiplied by 100 and we report Newey-West (1987) corrected (with twelve lags) *t*-statistics in parentheses.

Because most firms are essentially not competitive (firms that are not mentioned in other reports and thus get the same lowest C-Rank value), we examine the effect of C-Rank among

competitive firms only (those with at least one mention in other reports). The result show that the full market C-Rank has a positive and significant effect on stock return for the full sample (*t*-statistic=2.99). When removing the non-competitive firms the results are weaker, but still significant (*t*-statistics=2.37). This suggests that the effect of C-Rank on stock returns is not coming only from the difference between non-competitive and competitive firms, but is also important within the competitive firms. These results therefore corroborate the portfolio sort analysis, indicating that a higher competition status is associated with higher stock returns. The lower graph of Figure 2 further shows that the effect of C-Rank on stock returns is consistently positive over time.

The industry-based C-Ranks however exhibit weaker to no effect. The cross-industry C-Rank has a marginally positive effect (*t*-statistics of 1.86 and 1.77 for the full and competitive firms samples), where the within-industry C-Rank does not show any effect in both samples. This is consistent with the portfolio sort results and may suggest that a firm's competition status is reflected more by the pool of firms that operate in different industries than by those operating in the same industry. Altogether, the results in Table 6 suggest that to capture the real competitiveness power of a firm, the view of all potential competitors should be considered, those from both inside and outside industry. This conjecture is consistent with the strong precative ability of the full market C-Rank over future returns across all models and methodologies.

The high stock returns gained by companies with high competition status is consistent with our risk story. That is, being recognized by strong companies as a competitor imposes uncertainty as to the firm's future performance and value, which is compensated by high returns. Yet, finding a relation between firm characteristic and future returns, which is not captured by well documented risk factors (size, value, etc.) can also raise the question if this relation represents a temporary mispricing. For the case of our study, such a mispricing story implies that the high monthly returns of the high C-Rank firms represent either a correction to prior underpricing of these firms or current overpricing of these firms that should be corrected in the following months.

We acknowledge that it may be difficult for investors to obtain the C-Rank competition status of the firm as it involves high computer processing power and implementation of advanced technologies. While this can produce deviations of market stock prices from fundamental values, we do not have a reasonable explanation for why high C-Rank firms will exhibit different mispricing pattern than low C-Rank firms. Nevertheless, to assess whether the relatively high

monthly stock returns of high C-Rank firms represent a temporary mispricing, we look at the monthly returns around the release of the reports (assumed to be three months after the filing month).

We replicate the Fama-MacBeth regression from the left column in Table 6 (full market C-Rank for the full sample), where the C-Rank is measured for each month from three months before the month in which the most recent 10-K filings were used to calculate the C-Rank to nine months after. Figure 3 shows the regression coefficients of C-Rank. The coefficients are highly positive and fairly persistent over this period with some deterioration after the release month, which is expected as the C-Rank dynamically changes. That is, the high return gained by high C-Rank firms in three months after the filing date does not seem to correct for prior negative returns and does not revert in the subsequent months. This evidence does not point therefore to a potential misvaluation of firms' competitive edge.

5. Do firms "over-position" themselves in market competition?

Assigning a C-Rank competition strength value to each firm in every month can also provide us with an idea on how each firm see itself as a competitor. Specifically, the firm's own view on where it stands in the market competition can be assessed by the average C-Ranks of the firms it mentions as competitors. If the firm's own view on its competition status (i.e., the average C-Ranks of the firms it mentions) is higher than the market's view on its competition status (i.e., the firm's C-Rank), it can indicate one of two scenarios. The first is that the firm indeed has developed a strong competition status, which has not been recognized yet by other firms. The second is that the firm is "over-positioning" itself in the market by listing strong companies as competitors even though they are not their real competitors.[2] It is easy to understand why managers would have incentives to over-position their companies in the market competition. Such strategy can enhance the managers' and firms' reputation, can "put" them in the major league of competitors, can attract more attention in the financial market, the financial media, and others.

[2] We note that as long as there is some cross-sectional variation in C-Rank, then by construction a firm's C-Rank value is expected to be lower than the average C-Ranks of the firms it mentions. This is the essence of the PageRank algorithm; the leading firms build their high rankings from being mentioned by many weaker firms. However, because our cross-sectional samples are relatively large (typically thousands of annual reports), this expected gap is negligible. That is, significant actual gaps likely to represent high self-positioning.

An anecdotal example of high self-positioning is given by Bsquare Corp. ('BSQR'). Bsquare is a relatively small company that provides business-to-business (B2B) technology products and services, headquartered in Bellevue, WA. The market capitalization of the company as of December 2017 was $61.3M. In its annual report for 2017 Bsquare listed industry giants as competitors, including GE, IBM, and MSFT. For example, the company states: "*Although we are closely partnered with AWS and Microsoft, there are elements of their solutions with which we compete directly*". These mentioned companies are thousands of times larger than Bsquare with much higher C-Rank values. None of these companies recognize Bsqaure back as a competitor.

To assess whether firms tend to over-position themselves by listing strong companies that are not their real competitors we consider three proxies for the relation between the filing firm and each of the companies it lists. The first is the product similarity between the filing firm and the mentioned firm. We use the product similarity score provided by the Hoberg and Phillips data library (see Hoberg and Phillips (2010, 2016) and Hoberg, Phillips, and Prabhala (2014)). The second proxy is a dummy variable indicating if the mentioned firm operates in the same industry as the filing firm. And the third proxy is a dummy variable indicating if the mentioned firm mentions back the filing firm as a competitor, i.e., a mutual mention. To sum, real competitors are more likely to produce similar products, to operate in the same industry, and to recognize each other as competitors.

We conduct two tests. In the first we consider only financial reports that recognize at least five firms as competitors (a total of 5,691 reports). For each report we compute the correlation between the C-Rank of the mentioned firm and its similarity to the filing firm. We then average this correlation across all reports. A negative average correlation will indicate that the stronger the competition status of the mentioned firm is, it is less related to the filing firm; i.e., the filing firm shows a better competition status by listing strong firms as competitors even though these firms are not really related to the filing firm. The correlation averages reported in Table 7 are all negative, -0.23 for the product similarity score and the mutual-mention indicator, and -0.11 for the same-industry indicator, and are all statistically significant at any common level. These results provide a first indication that firms tend to enhance their competition status in 10-K filings by associating themselves with strong companies that are less related to them.

In the second test we consider all reports that specify at least one competitor (a total of 21,970 reports) and examine the relation between the firm's relative self-positioning and the average

similarity between the filing firm and each of its listed competitors. The firm's relative self-positioning is measured by the gap between the firm's own view of its competition status (the average C-Ranks of the firms it mentions) and the market's view of its competition status (the firm's C-Rank). The average similarity between the filing firm and the firms it mentions is based on the three proxies used above: the Hoberg-Phillips' product similarity score, the same-industry indicator, and the mutual-mention indicator.

We run a pool regression of the average similarity on the relative self-positioning. We control for the firm's C-Rank as it would be more difficult for strong competitors to enhance their competition status. We control for the firm's total product similarity score (also obtained from Hoberg and Phillips' data library) as it affects the extent to which a competitor can be viewed as similar to the filing firm. And we also control for firm size and market-to-book ratio. If firms tend to over-position themselves by listing strong companies that are not necessarily related to them, then we should expect a negative relation between the firm's relative self-positioning and the average similarity between the listed firms and the filing firms.

Table 8 presents the regression results. The coefficient of the relative self-positioning is negative and statistically significant in all specifications and similarity proxies (*t*-statistics between -4.67 to -14.89). This provides further evidence that the choice of the filing firms of which companies to identify as competitors takes into account the competition strength of the firms; weak firms that are listed as competitors are more similar to the filing firm, whereas strong firms that are listed as competitors are less similar to the filing firms. The results in Tables 7 and 8 suggest therefore that firms tend to over-position themselves in the market competition.

6. Future competitiveness

Firms that value their competition status higher than the market does can reflect two cases, as discusses above. The first case (real positioning) is that the firm indeed imposes strong competition to major companies, but it has not yet been recognized by them. In this case it is more likely that the real competitive power of the firm will be revealed to its competitors at some point, which will improve its C-Rank. The second case (fake positioning) is that the firm presents an artificial view of a strong competitor in order to enhance its reputation. In this case we should not expect any specific change in the firm's C-Rank.

To identify if firms fall in the first or the second case, we use the three similarity measures. We divide all filing firms into two groups by the average of the Hoberg-Phillips' product similarity score between the filing firm and each of the competitors it mentions. And in the same way we divide all filing firms into two groups by the average of the same-industry indicator, and by the average of the mutual-mention indicator. Higher values of these similarity averages indicate that the firm's relative high view of its competition status represents real competitive power, and thus likely to improve its C-Rank. Low averages of the similarity proxies however are more likely to represent an artificial over-positioning of the filing firm, thus less likely to improve its C-Rank.

This effect is examined in the regressions presented in Table 9. The dependent variable is the change in the firm's C-Rank from current filing to next year filing. The independent variable is the firm's relative self-positioning, which is measured by the gap between the firm's current own view of its competition status (the average C-Ranks of the firms it mentions) and the current market's view of its competition status (the firm's C-Rank). The control variables include the firm's current C-Rank, size, and market-to-book ratio.

All three similarity proxies exhibit a similar pattern; a current high self-positioning of a firm better predicts an improvement in next-year C-Rank when the high positioning relies on competitors that are similar to the filing firms. Specifically, for firms that list strong competitors that produce similar products, high self-positioning seems to improve the C-Rank (*t*-statistic=1.90). If the mentioned firms produce different products, over-positioning shows a weaker predictive ability (*t*-statistic=1.45). For firms that list strong competitors from the same industry, high self-positioning has a significant effect on a change in C-Rank (*t*-statistic=2.07), whereas for firms that list competitors from other industries, the effect is insignificant (*t*-statistic=0.46). For filing firms that mention strong competitors that mention them back, the effect of high positioning is also positive and significant (*t*-statistic=2.97), whereas for firms that enhance their competition status by mentioning firms that do not mention them back, the effect is weaker (*t*-statistic=1.53).

The evidence that high self-positioning does not lead to the same market view positioning when the mentioned firms are not similar to the filing firm provides further indication that firms often mention strong and unrelated companies as competitors just to artificially inflate their competition status.

7. Conclusions

We produce a dynamic measure of firm competitiveness by analyzing the cross-references of firms to their competitors in annual financial statements. Our procedure is based on an advanced text analysis technology that allows identifying competitors in financial reports, and on a PageRank-type algorithm that simultaneously assesses the value of each firm's reference in its competitors' reports.

Our results yield two major findings. The first one is that firms with higher competition ranking (C-Rank) gain higher subsequent stock returns. This effect is significant controlling for firm size and other common risk factors. The long-short investment strategy that buys high C-Rank stocks and shorts low C-Rank stocks generates an annualized 6-factor alpha of about 16 percent. Fama-MacBeth regressions corroborate this effect. A potential explanation for this result is that being recognized and thus targeted by many strong companies as a competitor induces a risk to the firm's future performance and value, which is compensated by high returns. The second finding is that firms tend to "inflate" their competition position by listing strong companies as competitors even if they operate in different industries and produce different products.

References

Ball Ray, and Philip Brown, 1968, An empirical evaluation of accounting income numbers, *Journal of Accounting Research* 6, 159-178.

Beaver, William H., Roger Clarke, and William F. Wright, 1979, The association between unsystematic security returns and the magnitude of earnings forecast errors, *Journal of Accounting Research* 17, 316-340.

Cohen, Lauren, Christopher Malloy, and Quoc Nguyen, 2018, Lazy prices, Harvard Business School working paper.

Garcia, Diego, and Øyvind Norli, 2012, Geographic dispersion and stock returns, *Journal of Financial Economics* 106, 547-565.

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

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1-22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Frazzini, Andrea, and Lasse H. Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1-25.

Healy, Paul M., 1985, The effect of bonus schemes on accounting decisions, *Journal of Accounting and Economics* 7, 85-107.

Healy, Paul M., Sok-Hyon Kang, and Krishna G. Palepu, 1987, The effect of accounting procedure changes on CEO's cash salary and bonus compensation, *Journal of Accounting and Economics* 9, 7-34.

Hoberg, Gerard, and Gordon Phillips, 2010, Product market synergies and competition in mergers and acquisitions: A text-based analysis, *Review of Financial Studies* 23, 3773-3811.

Hoberg, Gerard, and Gordon Phillips, 2016, Text-based network industries and endogenous product differentiation, *Journal of Political Economy* 124, 1423-1465.

Hoberg, Gerard, Gordon Phillips, and Nagpurnanand Prabhala, 2014, Product market threats, payouts, and financial flexibility, *Journal of Finance* 69, 293-324.

Holthausen, Robert W., David F. Larcker, and Richard G. Sloan, 1995, Annual bonus schemes and the manipulation of earnings, *Journal of Accounting and Economics* 19, 29-74.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jensen, Michael C., 1986, Agency costs of free cash flow, corporate finance, and takeovers, American Economic Review 76, 323–329.

Jones, Jennifer J., 1991, Earnings management during import relief investigations, *Journal of Accounting Research* 29, 193-228.

Li, Feng, Russell Lundholm, and Michael Minnis, 2013, A measure of competition based on 10-K filings, *Journal of Accounting Research* 51, 399-436.

Liberty, Susan E., and Jerold L. Zimmerman, 1986, Labor union contract negotiations and accounting choice, *Accounting Review* 61, 692-712.

Newey, Whitney K., and Kenneth D. West, 1987, A simple positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703-708.

Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd, 1999, The PageRank citation ranking: Bringing order to the web, *Technical Report*. Stanford InfoLab. URL http://ilpubs.stanford.edu:8090/422/.

Perry, Susan E., and Thomas H. Williams, 1994, Earnings management preceding management buyout offers, *Journal of Accounting and Economics* 18, 157-179.

Sloan, Richard G., 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review* 71, 289-315.

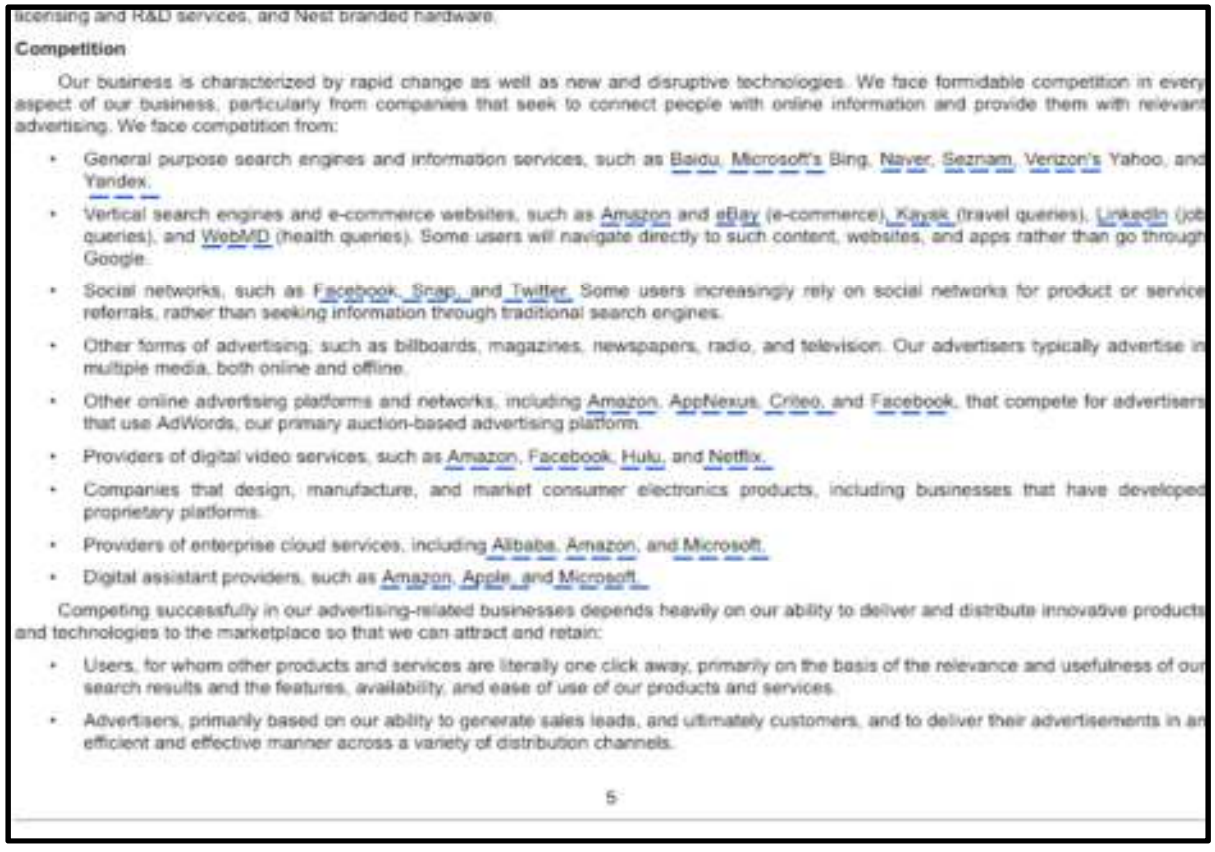
Sweeney, Amy P., 1994, Debt covenant violations and managers' accounting responses, *Journal of Accounting and Economics* 17, 281-308.

Appendix. Text analysis of competition sections in 10-Ks

Dataset

We match company tickers to CIKs, identifiers used by SEC-Edgar, and download from SEC-Edgar the 10-K filings. We observe a total of 119,785 10-Ks filed by 11,304 firms over the period 1995-2017. The focus of this paper is Part I / Item 1 – Business of the 10-K form. Although reporting firms are not required to designate a competition section in Item 1, we find that 68,952 of the forms used in this study (58%) include a designated section for competition. And about 39% of these competition sections include names of the company's competitors.

The example below is an extract from the 2017 10-K form filed by Alphabet Inc., parent company of Google. In Part I / Item 1 – Business, Alphabet designates a section to discuss its competitive environment. In this section it lists both the areas in which it faces competition (e.g., general search engines, vertical search engines, social networks, etc.) and the companies it considers as competitors in each of the areas.



licensing and R&D services, and Nest branded hardware.

Competition

Our business is characterized by rapid change as well as new and disruptive technologies. We face formidable competition in every aspect of our business, particularly from companies that seek to connect people with online information and provide them with relevant advertising. We face competition from:

- General purpose search engines and information services, such as Baidu, Microsoft's Bing, Naver, Seznam, Verizon's Yahoo, and Yandex.
- Vertical search engines and e-commerce websites, such as Amazon and eBay (e-commerce), Kayak (travel queries), LinkedIn (job queries), and WebMD (health queries). Some users will navigate directly to such content, websites, and apps rather than go through Google.
- Social networks, such as Facebook, Snap, and Twitter. Some users increasingly rely on social networks for product or service referrals, rather than seeking information through traditional search engines.
- Other forms of advertising, such as billboards, magazines, newspapers, radio, and television. Our advertisers typically advertise in multiple media, both online and offline.
- Other online advertising platforms and networks, including Amazon, AppNexus, Criteo, and Facebook, that compete for advertisers that use AdWords, our primary auction-based advertising platform.
- Providers of digital video services, such as Amazon, Facebook, Hulu, and Netflix.
- Companies that design, manufacture, and market consumer electronics products, including businesses that have developed proprietary platforms.
- Providers of enterprise cloud services, including Alibaba, Amazon, and Microsoft.
- Digital assistant providers, such as Amazon, Apple, and Microsoft.

Competing successfully in our advertising-related businesses depends heavily on our ability to deliver and distribute innovative products and technologies to the marketplace so that we can attract and retain:

- Users, for whom other products and services are literally one click away, primarily on the basis of the relevance and usefulness of our search results and the features, availability, and ease of use of our products and services.
- Advertisers, primarily based on our ability to generate sales leads, and ultimately customers, and to deliver their advertisements in an efficient and effective manner across a variety of distribution channels.

5

In total Alphabet lists twenty individual companies as competitors. These include domestic US firms such as Verizon and Microsoft, foreign firms (e.g., Baidu), and also private companies and private subsidiaries of public companies such as Hulu and Yahoo respectively. Interestingly, some of the listed competitors appear multiple times as Alphabet considered them as competitors in multiple areas. Amazon which is listed five times in considered by Alphabet as a competitor in e-commerce search, online advertising, digital video, enterprise cloud, and digital assistance services.

Identifying firms in competition section

Once a designated competition section is found on a 10-K filing, our process attempts to identify which specific companies it lists. Since competitors are referred to by names using natural language, matching listed firms to security identifiers requires some additional text and language processing. We use an open-source natural language processing (NLP) tool, StanfordNER,[3] which is designed to label names of "things" in sequences of words. Each of the 68,952 designated competition sections is passed to the StanfordNER tool which is required to provide a list of text parts that are likely names of organizations. We consider each name of organization as a potential public company by matching against databases of public companies.

We apply a matching process that first searches for organization name on Edgar-SEC database, then on company name column of the CRSP master file, and finally we search Wikipedia using suspected organization names and in the cases of public companies parse the ticker following a "traded as" tag.[4] On average, we find 1,940 unique firms mentioned on 10-K filing of other companies each year.

[3] Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling. Proceedings of the 43nd Annual Meeting of the Association for Computational Linguistics (ACL 2005), pp. 363-370.http://nlp.stanford.edu/~manning/papers/gibbscrf3.pdf https://nlp.stanford.edu/software/CRF-NER.shtml

[4] To increase the probability of matching suspected names of organizations to public companies we remove generic strings and suffixes such as Corp., LTD, LLC, etc which are oftren used prior to processing the matching algorithm. We then use the standard text matching algorithms Sequence Matcher and Levenshtein Distance.

Table 1. C-Rank distribution

The table shows descriptive statistics of the C-Rank measures as described in Section 2, where all statistics are multiplied by 100. The sample includes 1,664,271 firm-month observations over the period 1995-2017.

	Mean	Stdev	min	p25	p50	p75	max
C-Rank full market	0.0172	0.0093	0.0123	0.0136	0.0149	0.0167	0.3422
C-Rank cross-industry	0.0193	0.0097	0.0143	0.0153	0.0174	0.0191	0.3673
C-Rank within-industry	0.2075	0.2317	0.0519	0.0942	0.1154	0.2324	10.0000

Table 2. Top competitors against largest companies over the sample period

The left panel shows the five companies (by ticker symbol) with the highest full market C-Rank competition status each year over the sample period. The right panel shows the largest companies for the same period.

	Top competitors					Largest firms				
Year	1st	2nd	3rd	4th	5th	1st	2nd	3rd	4th	5th
1995	IBM	HPQ	GE	NIPNY	ITC	GE	T	XOM	KO	MRK
1996	IBM	MSFT	HPQ	WMT	MSI	GE	KO	XOM	INTC	MSFT
1997	IBM	MSFT	HPQ	LU	JNJ	GE	KO	MSFT	XOM	MRK
1998	IBM	MSFT	HPQ	LU	MSI	MSFT	GE	INTC	WMT	XOM
1999	MSFT	IBM	LU	HPQ	MSI	MSFT	GE	CSCO	WMT	XOM
2000	MSFT	IBM	LU	HPQ	A	GE	XOM	PFE	CSCO	C
2001	IBM	MSFT	MSI	SIEGY	HPQ	GE	MSFT	XOM	C	WMT
2002	IBM	MSFT	HPQ	CSCO	GOOGL	MSFT	GE	XOM	WMT	PFE
2003	IBM	MSFT	CSCO	WMT	JNJ	GE	MSFT	XOM	PFE	C
2004	IBM	MSFT	WMT	CSCO	NVS	GE	XOM	MSFT	C	WMT
2005	IBM	WMT	MSFT	A	PFE	GE	XOM	MSFT	C	PG
2006	MSFT	IBM	WMT	ELMG	ABT	XOM	GE	MSFT	C	BAC
2007	IBM	MSFT	WMT	GE	GSK	XOM	GE	MSFT	T	PG
2008	MSFT	WMT	IBM	GE	A	XOM	WMT	PG	MSFT	GE
2009	IBM	MSFT	GE	ELMG	WMT	XOM	MSFT	WMT	AAPL	JNJ
2010	MSFT	WMT	GE	IBM	CSCO	XOM	AAPL	MSFT	GE	WMT
2011	MSFT	IBM	GE	BAC	ELMG	XOM	AAPL	MSFT	IBM	CVX
2012	MSFT	GOOGL	GE	WMT	IBM	AAPL	XOM	WMT	MSFT	GE
2013	GOOGL	MSFT	AAPL	WMT	IBM	AAPL	XOM	GOOGL	MSFT	GE
2014	GOOGL	MSFT	IBM	FB	WMT	AAPL	XOM	MSFT	JNJ	WFC
2015	GOOGL	FB	IBM	MSFT	MDT	AAPL	MSFT	XOM	AMZN	GE
2016	GOOGL	FB	PFE	NVS	MDT	AAPL	MSFT	XOM	AMZN	JNJ
2017	GOOGL	NVS	MDT	FB	PFE	AAPL	MSFT	AMZN	FB	JNJ

Table 3. Correlation between C-Rank and firm characteristics

The table shows the time-series averages of monthly cross-sectional correlations between the three C-Rank measures and firm characteristics. Firm size is computed as stock price multiplied by the number of shares outstanding (in logs). Market-to-book ratio is the market value of equity divided by the book value of equity (in logs). Past return is based on monthly stock returns over the last six months skipping the most recent month (see Jegadeesh and Titman (1993)). We estimate profitability by return on equity (ROE), computed by the annual income before extraordinary items divided by the previous year's book equity value. We estimate investment by the annual change in gross property, plant, and equipment, plus the change in inventories, scaled by lagged book value of assets. Market beta is estimated using a regression of a firm overlapping 3-day log return on the equivalent market return over the past year (see Frazzini and Pedersen (2014) for a similar procedure). We calculate idiosyncratic volatility for each month by the standard deviation of the residuals of regression of daily stock returns on the daily Fama and French (1993) three factors. Panel A shows the correlations for the full sample and Panel B for a subsample of competitive firms, which includes only firms that are recognized as competitors by other firms at least once. The sample period is 1995-2017.

Panel A. Full sample			
	C_Rank full market	C_Rank cross-ind	C_Rank within-ind
Log(size)	0.568	0.410	0.239
Log(market-to-book)	0.019	0.008	-0.003
Past return	0.008	0.003	0.012
Profitability	0.080	0.051	0.072
Investment	-0.018	-0.001	0.083
Beta	0.018	0.001	-0.109
Idiosyncratic volatility	-0.107	-0.069	-0.137
Panel B. Competitive firms			
	C_Rank full market	C_Rank cross-ind	C_Rank within-ind
Log(size)	0.583	0.398	0.361
Log(market-to-book)	0.025	0.006	0.000
Past return	0.003	0.002	0.017
Profitability	0.121	0.063	0.127
Investment	-0.039	-0.001	0.053
Beta	-0.074	-0.081	-0.147
Idiosyncratic volatility	-0.170	-0.112	-0.182

Table 4. Returns of portfolios sorted on C-Rank

We run monthly cross-sectional regressions of C-Rank as of three months earlier on current firm size, and use the regression residuals as our sorting variable. Each month we divide all stocks into five equal-sized portfolios according to their C-Rank-residual. The portfolios are equal-weighted and held for one month. The table shows the portfolios' mean excess monthly stock returns (in excess of the risk-free rate) and alphas from factor models. The CAPM uses the market factor. The factors in the 3-factor model are the Fama and French (1993) factors. The factors in the 4-factor model are the Fama-French factors augmented with a momentum factor. The factors in the 5-factor model are the Fama and French (2015) factors. The factors in the 6-factor model are the Fama-French factors augmented with a momentum factor. Panels A, B, and C show the results for the full market, cross-industry, and within-industry C-Rank measures. All returns and alphas are in percent per month and the corresponding *t*-statistics are in parentheses. The sample period is 1995-2017.

Panel A. Full market C-Rank	1-low	2	3	4	5-high	high-low
Mean excess return	0.77	0.87	0.80	0.97	1.70	0.93
	(2.60)	(2.40)	(2.25)	(2.55)	(4.03)	(3.76)
CAPM alpha	0.03	0.04	0.02	0.23	0.80	0.77
	(0.28)	(0.20)	(0.09)	(0.86)	(3.13)	(3.17)
3-factor alpha	-0.06	-0.10	-0.10	0.17	0.76	0.83
	(-0.98)	(-1.33)	(-0.84)	(0.77)	(3.68)	(3.73)
4-factor alpha	-0.03	0.00	0.06	0.46	1.07	1.11
	(-0.53)	(0.02)	(0.56)	(2.47)	(6.25)	(5.65)
5-factor alpha	-0.13	-0.04	0.00	0.41	1.03	1.16
	(-1.96)	(-0.48)	(0.03)	(1.85)	(4.97)	(5.32)
6-factor alpha	-0.10	0.03	0.12	0.61	1.24	1.35
	(-1.62)	(0.47)	(1.04)	(3.18)	(7.23)	(7.00)

Panel B. Cross-industry C-Rank	1-low	2	3	4	5-high	high-low
Mean excess return	0.80	0.87	0.93	0.90	1.61	0.81
	(2.71)	(2.39)	(2.39)	(2.42)	(3.87)	(2.87)
CAPM alpha	0.05	0.02	0.07	0.16	0.82	0.78
	(0.55)	(0.09)	(0.33)	(0.63)	(2.76)	(2.72)
3-factor alpha	-0.01	-0.10	-0.08	0.06	0.78	0.79
	(-0.16)	(-1.46)	(-0.65)	(0.33)	(3.00)	(2.95)
4-factor alpha	0.02	0.00	0.09	0.34	1.12	1.09
	(0.38)	(-0.07)	(0.86)	(2.02)	(4.88)	(4.46)
5-factor alpha	-0.03	-0.06	0.05	0.25	1.08	1.10
	(-0.41)	(-0.94)	(0.38)	(1.27)	(4.06)	(4.04)
6-factor alpha	0.00	0.00	0.15	0.44	1.30	1.30
	(-0.03)	(-0.01)	(1.50)	(2.56)	(5.58)	(5.22)

Panel C. Within-industry C-Rank	1-low	2	3	4	5-high	high-low
Mean excess return	0.96	0.93	0.97	1.33	0.94	-0.01
	(1.98)	(2.87)	(3.06)	(3.36)	(2.98)	(-0.04)
CAPM alpha	-0.10	0.19	0.27	0.52	0.24	0.34
	(-0.35)	(1.13)	(1.51)	(2.03)	(1.36)	(1.15)
3-factor alpha	-0.03	0.06	0.14	0.40	0.10	0.13
	(-0.16)	(0.57)	(1.23)	(2.21)	(0.75)	(0.58)
4-factor alpha	0.19	0.22	0.30	0.62	0.25	0.06
	(1.04)	(2.26)	(2.91)	(3.77)	(1.94)	(0.26)
5-factor alpha	0.46	0.06	0.14	0.55	0.07	-0.39
	(2.64)	(0.55)	(1.12)	(2.94)	(0.52)	(-1.86)
6-factor alpha	0.59	0.17	0.25	0.70	0.18	-0.41
	(3.71)	(1.75)	(2.34)	(4.15)	(1.36)	(-1.98)

Table 5. Robustness Checks on portfolios sorted on C-Rank-residual

We run monthly cross-sectional regressions of the full market C-Rank as of three months earlier on current firm size, and use the regression residuals as our sorting variable. Each month we divide all stocks into five equal-sized portfolios according to their C-Rank-residual. The portfolios are equal-weighted and held for one month. The table reports 6-factor alphas where the factors are the Fama-French (2015) factors augmented with a momentum factor. All alphas are in percent per month and the corresponding *t*-statistics are in parentheses. The full sample period is 1995 to 2017. The full sample period is broken up into subsamples in Panel A. Recession periods are based on NBER recession dummy. The holding period is increased to 3, 6, 12, and 18 months in Panel B.

	1-low	2	3	4	5-high	high-low
Full sample	-0.10	0.03	0.12	0.61	1.24	1.35
	(-1.62)	(0.47)	(1.04)	(3.18)	(7.23)	(7.00)
Panel A. Subsamples						
Excluding January	-0.08	0.04	-0.08	0.03	0.72	0.80
	(-1.25)	(0.59)	(-0.77)	(0.21)	(5.03)	(5.01)
Excluding Recessions	-0.11	0.02	0.15	0.69	1.28	1.39
	(-1.67)	(0.29)	(1.38)	(3.58)	(7.03)	(6.74)
1995-2006	-0.18	0.12	0.35	0.98	1.78	1.96
	(-1.75)	(1.03)	(1.86)	(3.24)	(6.04)	(5.76)
2007-2017	-0.04	-0.06	-0.05	0.22	0.73	0.77
	(-0.72)	(-0.83)	(-0.44)	(0.96)	(4.33)	(4.48)
Panel B. Longer investment horizons						
3 months	-0.10	0.05	0.21	0.64	1.13	1.23
	(-1.57)	(0.70)	(1.88)	(3.35)	(6.95)	(6.80)
6 months	-0.09	0.08	0.25	0.68	1.06	1.15
	(-1.35)	(1.12)	(2.19)	(3.49)	(6.83)	(6.69)
12 months	-0.05	0.06	0.29	0.71	0.96	1.01
	(-0.80)	(0.88)	(2.40)	(3.65)	(6.66)	(6.49)
18 months	-0.03	0.08	0.32	0.71	0.88	0.91
	(-0.38)	(1.14)	(2.60)	(3.67)	(6.28)	(6.07)

Table 6. Fama-MacBeth regressions on C-Rank

We run cross-sectional Fama and MacBeth (1973) regressions each month of excess stock returns. The independent variables are C-Rank (full market, cross-industry, and within-industry, all are standardized monthly), log market capitalization, log market-to-book ratio, past six-month return, profitability, investment intensity, market beta, and idiosyncratic volatility, as described in Table 3. We run the regressions on the full sample and on a subsample of competitive firms, which includes only firms that are recognized as competitors by other firms at least once. All coefficients are multiplied by 100 and Newey-West corrected *t*-statistics (with twelve lags) are in parentheses. The sample period is 1995-2017.

	C-Rank full market		C-Rank cross-industry		C-Rank within-industry	
	All firms	Competitive firms	All firms	Competitive firms	All firms	Competitive firms
Intercept	2.86	3.43	2.71	2.51	2.71	3.29
	(4.86)	(4.62)	(4.59)	(2.50)	(4.77)	(4.79)
C-Rank	5.35	3.02	2.51	1.90	-0.04	2.46
	(2.99)	(2.37)	(1.86)	(1.77)	(-0.01)	(0.68)
Log(size)	-0.11	-0.15	-0.10	-0.10	-0.10	-0.14
	(-3.19)	(-3.49)	(-2.92)	(-1.74)	(-3.04)	(-3.70)
Log(market-to-book)	-0.03	0.03	-0.04	0.00	-0.03	0.06
	(-0.74)	(0.54)	(-0.81)	(-0.01)	(-0.80)	(0.95)
Past return	0.56	0.47	0.53	0.31	0.54	0.50
	(2.49)	(1.40)	(2.42)	(0.83)	(2.45)	(1.49)
Profitability	0.78	0.37	0.77	0.53	0.77	0.44
	(3.74)	(1.42)	(3.62)	(1.50)	(3.74)	(1.76)
Investment	-1.74	-1.25	-1.76	-1.42	-1.76	-1.31
	(-4.87)	(-2.55)	(-4.87)	(-1.80)	(-4.95)	(-2.75)
Beta	0.00	0.02	0.00	0.08	-0.01	0.02
	(-0.02)	(0.09)	(-0.04)	(0.42)	(-0.04)	(0.11)
Idiosyncratic volatility	-30.96	-22.32	-30.45	-15.14	-30.48	-20.15
	(-6.24)	(-3.83)	(-6.09)	(-1.54)	(-6.05)	(-3.55)

Table 7. Correlation between self-position C-Rank and measures of firm similarity

We consider only filing firms that mention at least five competitors in their annual reports (5,691 reports over 1995-2017). For each mentioned competitor we calculate three measures: (i) the similarity between the products of the competitor and the filing firm, based on the product similarity scores of Hoberg and Phillips (2010, 2016), (ii) a dummy variable indicating if the competitor operates in the same industry as the filing firm, and (iii) a dummy variable indicating if the competitor mentions back the filing firm, i.e., a mutual mentioning. For each filing firm, we calculate the correlation between the three similarity measures and the competitor's C-Rank.

	Measures of similarity between the filing firm and its listed competitors		
	Hoberg-Phillps' product similarity score	Same-industry indicator	Mutual-mention indicator
Sample mean	0.0465	0.7964	0.0407
Sample Stdev	0.0662	0.4027	0.1977
Mean correlation with competitor C-Rank	-0.2268	-0.1072	-0.2323
P-value	<.0001	<.0001	<.0001

Table 8. Regression of the similarity between the filing firm and the competitors it mentions on the filing firm's relative self-positioning

For each filing firm we measure the extent to which the firm "over-positions" its competition status as follows. We first calculate the average C-Rank of each of the firms that are mentioned as competitors in its report, referred to as 'own view C-Rank'. The relative self-positioning is then given by the difference between the firm's own view C-Rank and its actual C-Rank. We also measure the similarity between each firm and the competitors it lists in its report, based the variables used in Table 7. The first measure is the average Hoberg-Phillips' product similarity score between the filing firm and its competitors. The second measure is the average of the dummy variable indicating if the competitor operates in the same industry as the filing firm. And the third measure is the average of the dummy variable indicating if the competitor also recognizes the filing firm in its report. We regress each of these average similarity measures on the filing firm's relative self-positioning. The control variables include the filing firm's C-Rank, the total product similarity score of the filing firm, and the filing firm's (log) size and (log) market-to-book ratio. All regressions include industry-year fixed effects (by GIC classification), and the standard errors are clustered by firm. The coefficients' *t*-statistics are in parentheses. The sample contains 21,970 annual reports over 1995-2017.

	Average product similarity score		Average same-industry indicator		Average mutual-mention indicator	
Intercept	0.05	0.00	0.37	-0.04	0.08	-0.10
	(9.26)	(0.25)	(5.11)	(-0.49)	(6.50)	(-6.13)
Relative self-positioning	-25.68	-24.80	-96.72	-77.56	-68.03	-34.23
	(-14.89)	(-12.28)	(-6.34)	(-4.67)	(-10.39)	(-5.83)
C-Rank		13.33		-61.61		446.60
		(0.68)		(-1.00)		(3.34)
Total product similarity score		0.06		0.26		-0.03
		(5.42)		(4.66)		(-2.70)
Log(size)		0.39		1.46		0.51
		(7.40)		(4.47)		(3.36)
Log(market-to-book)		-1.34		-2.31		-1.58
		(-3.03)		(-0.74)		(-1.11)
Adj R-square	0.062	0.089	0.094	0.096	0.042	0.118

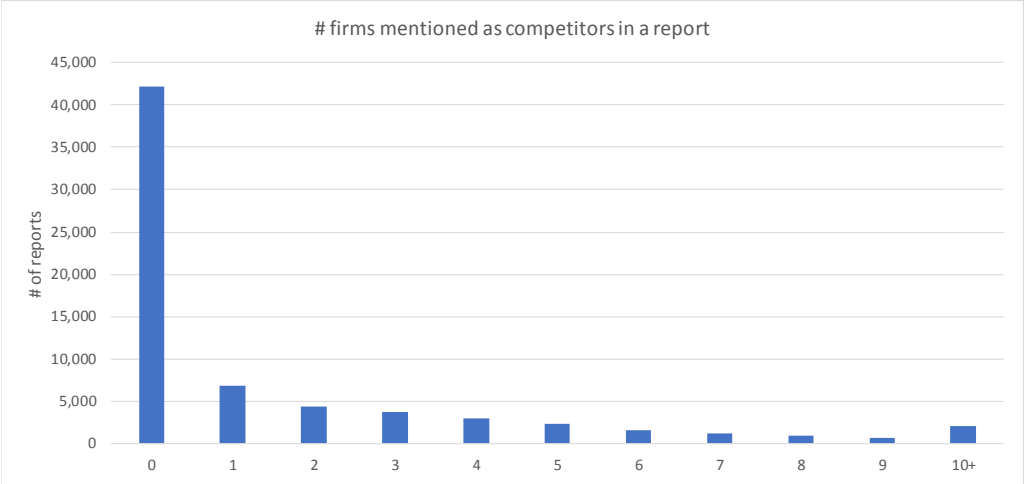
Table 9. Regression of firm self-positioning on changes in C-Rank

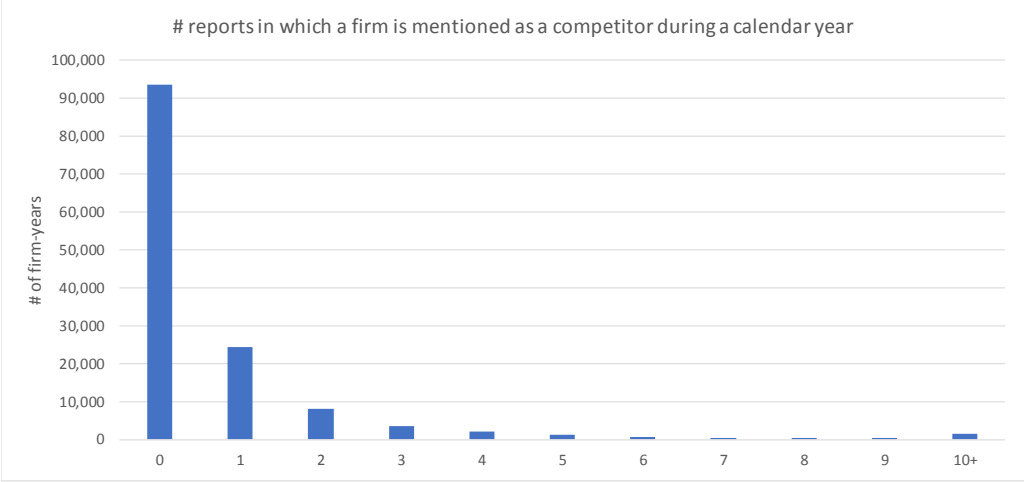
For each filing firm we measure the extent to which the firm "over-positions" its competition status as follows. We first calculate the average C-Rank of each of the firms that are mentioned as competitors in its report, referred to as 'own view C-Rank'. The relative self-positioning is then given by the difference between the firm's own view C-Rank and its actual C-Rank. We regress the firm's annual change in its C-Rank competition status from current report to next year report on its relative self-positioning in the current report. The changes in C-Rank are based on standardized values as C-Rank values are affected by the number of available reports in each year. We run the regression on subsamples of low and high values of the three measures of similarities described in Table 7: the average Hoberg-Phillips' product similarity score between the filing firm and its competitors; the average of the dummy variable indicating if the competitor operates in the same industry as the filing firm; and the average of the dummy variable indicating if the competitor also recognizes the filing firm in its report. The control variables include the filing firm's C-Rank, (log) size, and (log) market-to-book ratio. The coefficients' *t*-statistics are in parentheses. The sample contains 21,970 annual reports over 1995-2017.

	Average product similarity score		Average same-industry indicator		Average mutual-mention indicator	
	Low	High	Low	High	Low	High
Intercept	-2.86	-2.33	-4.89	-1.21	-2.05	8.16
	(-3.16)	(-2.37)	(-3.28)	(-1.51)	(-3.63)	(1.81)
Relative self-positioning	0.17	0.29	0.11	0.23	0.12	2.48
	(1.45)	(1.90)	(0.46)	(2.07)	(1.53)	(2.97)
C-Rank	-13.15	-18.13	-28.44	-10.95	-13.03	-38.03
	(-11.24)	(-19.39)	(-19.21)	(-13.25)	(-18.39)	(-13.14)
Log(size)	0.00	0.00	0.00	0.00	0.00	0.00
	(7.81)	(23.85)	(23.32)	(13.99)	(16.67)	(17.69)
Log(market-to-book)	0.13	-0.01	-0.19	0.13	0.20	-0.89
	(0.96)	(-0.09)	(-0.79)	(1.19)	(2.51)	(-1.52)
Adj R-square	0.042	0.075	0.124	0.036	0.042	0.190

Figure 1. Distribution of competitor mentions

The upper figure shows the distribution of the number of firms mentioned as competitors in a report for a total of 68,952 10-Ks with competition sections over the period 1995-2017. The middle figure shows the distribution of the number of reports in which a firm is mentioned as a competitor during a calendar year for a total of 135,921 firm-years. The bottom figure shows the joint distribution.



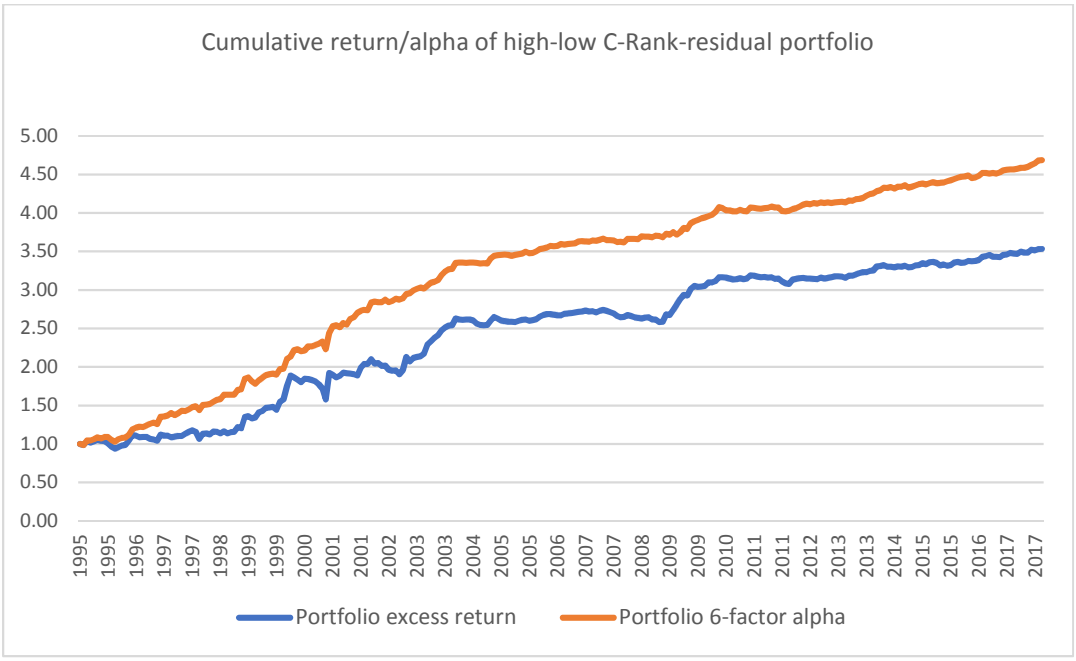


% reports in which a firm is mentioned

% competitors mentioned in a report	0	1	2	3	4	5	6	7	8	9	10+	Total
0	50.48	5.60	2.12	0.98	0.57	0.36	0.22	0.14	0.09	0.08	0.45	61.09
1	8.37	0.95	0.30	0.12	0.06	0.04	0.02	0.01	0.01	0.01	0.05	9.94
2	5.13	0.76	0.30	0.08	0.07	0.03	0.01	0.01	0.01	0.01	0.02	6.43
3	4.24	0.67	0.29	0.12	0.06	0.05	0.03	0.01	0.01	0.00	0.03	5.51
4	3.15	0.56	0.29	0.12	0.07	0.04	0.02	0.01	0.01	0.01	0.02	4.30
5	2.34	0.46	0.24	0.11	0.06	0.03	0.02	0.02	0.02	0.01	0.02	3.33
6	1.64	0.38	0.16	0.08	0.04	0.02	0.02	0.01	0.01	0.00	0.02	2.38
7	1.20	0.29	0.13	0.06	0.04	0.01	0.01	0.00	0.01	0.00	0.03	1.78
8	0.88	0.20	0.11	0.06	0.03	0.02	0.01	0.01	0.01	0.00	0.01	1.34
9	0.59	0.17	0.08	0.04	0.04	0.02	0.01	0.01	0.00	0.00	0.02	0.98
10+	1.50	0.52	0.33	0.17	0.14	0.09	0.06	0.03	0.02	0.00	0.08	2.94
Total	79.52	10.56	4.35	1.94	1.18	0.71	0.43	0.26	0.20	0.12	0.75	100.00

Figure 2. Cumulative effect

The upper graph plots the cumulative return and 6-factor alpha of the zero-investment portfolio of buying top quintile and selling bottom quintile of the C-Rank residual stocks, as described in Panel A of Table 4. The lower graph shows the cumulative coefficient of the standardized C-Rank from the Fama-MacBeth monthly cross-sectional regressions of stock returns, as described in Table 6.



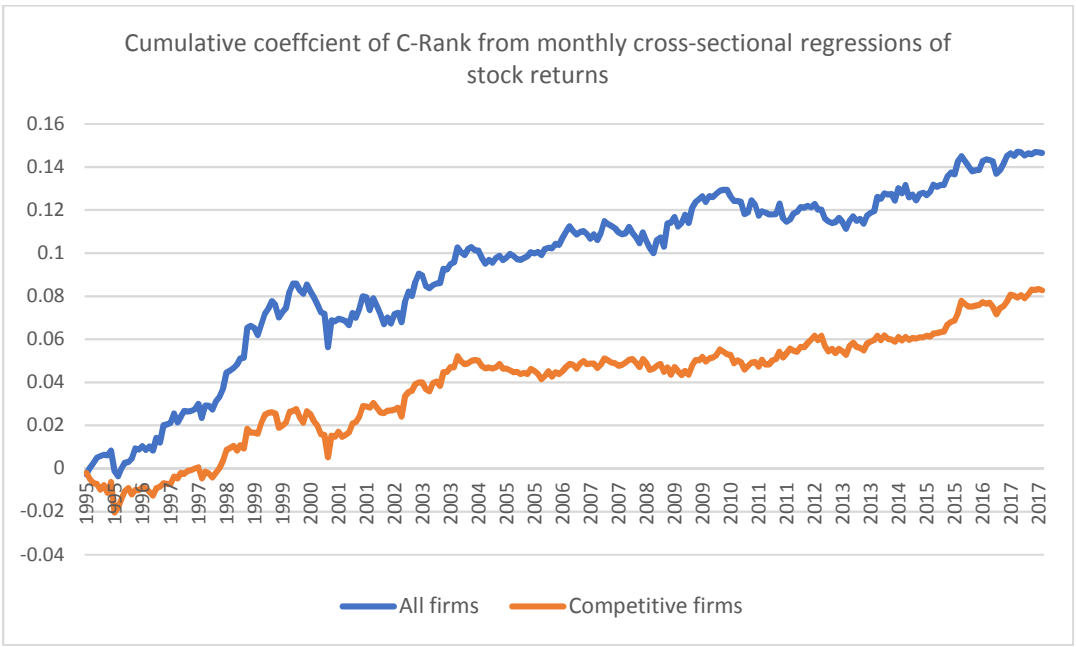


Figure 3. C-Rank coefficient from Fama-MacBeth regressions around statement release

We run the same Fama-MacBeth regression as in the left column in Table 6 (full market C-Rank for the full sample), where the C-Rank is measured in each month from three months before the month in which the most recent 10-K filings were used to calculate the C-Rank to nine months after. The coefficient of '+3m' is the one from the regression in Table 5 (as we assume that it takes three months to release the reports). The figure shows only the coefficient of the C-Rank. The sample period is 1995-2017.

