September 12, 2017

MACHINE LEARNING TAKEOVERS

Predicting Takeover Targets and Its Applications

- **Targeting Mergers and Acquisitions.** Major M&A transactions almost always make the headline news and capture investor attention. In a typical takeover deal, the acquirer offers a significant premium over the target firm's prevailing share price. The upside from accurately identifying takeover targets is significant. However, we have to be mindful that M&A transactions are rare events – on average, at any given point-in-time, only 0.5% companies are actual targets.
- **Alternative Big Data Factors.** In addition to standard quantitative factors, we develop and use a set of unconventional factors in our M&A model. In particular, we find that corporate filings (e.g., Forms 8-K, 4, 13D/13G, DFAN14A, etc.) from the EDGAR database contains useful information about future M&A transactions. Moreover, we use the TRSDC database to construct past M&A event count signals and our proprietary Sector M&A Momentum factor.
- **Machine Learning Takeovers.** To deal with the rare event/highly unbalanced sample nature of takeover prediction, we introduce matched sample and bagging (bootstrap aggregating) techniques. We also study the impact of factor (feature selection) via boosting and assess its impact on the logit model. Beyond conventional logit/logistic regression, we investigate a wide range of machine learning algorithms in takeover prediction: CART, random forest, SVM, ANN, and AdaBoost. Our final SMAP (Systematic Merger and Acquisition Prediction) model takes advantage of forecast combination, by blending seven machine learning models. The SMAP model exhibits materially higher predictive power for takeovers than prior standard models.
- **Application of the SMAP Model.** Discretionary managers can obviously use the SAMP model as a pre-screening tool before conducting more in-depth fundamental research. In this report, we also show two other applications. First, due to the superior accuracy of the SMAP model, buying stocks with the highest takeover probabilities result in higher returns. Next, we demonstrate how long/short managers can use the SMAP to avoid shorting potential takeover targets. The SMAP overlay on the short side not only boosts return/Sharpe ratio, but also reduce portfolio volatility/drawdown.

Source: gettyimages.com

Sheng Wang
SWang@wolferesearch.com

Yin Luo, CFA, CPA
YLuo@wolferesearch.com

Javed Jussa
JJussa@wolferesearch.com

Miguel Alvarez
MAlvarez@wolferesearch.com

Gaurav Rohal, CFA
GRohal@wolferesearch.com

QES Desk Phone: 1.646.582.9230
Luo.QES@wolferesearch.com

Table of Contents

A LETTER TO OUR READERS

Introducing our SMAP (Systematic Merger and Acquisition Prediction) Model

Major M&A transactions almost always make to the headline news and grab investors' attention. In a typical takeover deal, the acquirer offers a significant premium over the target firm's prevailing share price. As a result, the target firm's stock price tends to jump considerably on the merger announcement date. The upside from accurately identifying takeover targets is significant.

However, we have to be mindful that M&A transactions are rare events – on average, at any given point-in-time, only 0.5% companies are actual targets. Academic research generally concludes that predicting takeover is extremely difficult. Furthermore, we find that takeover targets tend to have poor fundamental characteristics, e.g., slow growth, low profitability, expensive valuation, and disliked by sell-side analysts. Although these traditional factors may help us to predict future transactions, the weak fundamental properties of these companies create one extra challenge. If we invest in companies with the highest takeover probabilities, because of the rare event nature, coupled with the negative drag from weak fundamentals, the economic profit might be minimal or even negative.

In this research, we once again show that statistical significance does not necessarily leads to superior investment performance. We first introduce a backtesting framework for event prediction. Then we shift our attention on whether traditional stock-selection factors can be used in takeover forecasting, as evidenced in the majority of academic research.

The first major contribution we make to the literature is to study how alternative Big Data can enhance our predictive power. In particular, we find corporate filings (e.g., Forms 8-K, 4, 13D/13G, DFAN14A, etc.) from the EDGAR database contains useful information about future M&A transactions. In addition, using the TRSDC database, we construct past M&A event count signals and our proprietary Sector M&A Momentum factor.

The predominant tool used in mainstream finance research for binary event classification problem is the logit or logistic regression. In this research, we setup a machine learning framework for event prediction, by introducing matched sampling, bagging (bootstrap aggregating), boosting, and forecast combination techniques. Then, we test a wide range of machine learning algorithms (e.g., CART, random forest, SVM, ANN, and AdaBoost) in takeover prediction. Finally, we introduce our SMAP (Systematic Merger and Acquisition Prediction) model, which demonstrates exceptional accuracy in out-of-sample prediction.

Discretionary managers can obvious use the SMAP model as a pre-screening tool before conducting more in-depth fundamental research. In this research, we also show two other applications. First, due to the exceptional accuracy of the SMAP model, buying those stocks with the highest takeover probabilities achieves superior return. Next, we demonstrate how long/short managers can use the SMAP to avoid shorting potential takeover targets. The SMAP overlay on the short side not only boosts return/Sharpe ratio, but also reduce portfolio volatility/drawdown.

Regards,

Yin, Sheng, Miguel, Javed, and Luo's QES team

OVERVIEW OF M&A TARGET

In the past 30 years, there have been over 5,000 takeover announcements involving target companies in the Russell 3000 universe, which averages to about 15 transactions per month. Usually the acquirer company offers a significant premium over the prevailing share price of the target company. Once a deal is announced, the share price of the target company usually jumps considerably, approaching to the offering price. How to trade after the announcement is what known as risk arbitrage. However, the upside potential from predicting takeover target can be significantly higher, albeit riskier and substantially more difficult.

In this research, we attempt to construct a takeover prediction model, i.e., identifying potential takeover targets before deals are actually announced. We study the common characteristics of M&A target companies; explore potential data sources; and more importantly; show how to apply such a model in practice.

M&A DATABASE

We use the Thomson Reuters SDC (TRSDC) database as our primary source for M&A announcements. For our research, we limit our analysis to takeover announcements that occur after 1985 and target companies in the Russell 3000 universe at the time of announcement.

Figure 1 shows the different types of corporate events that are defined as M&A transactions in the TRSDC database. The type “Merger” type is what investors normally considered as M&A’s. However, other M&A related events can increase our sample size, and contain useful information in predicting future takeovers. We will discuss this with more details in later sections.

Figure 1 Definition M&A Related Events in the TRSDC Database

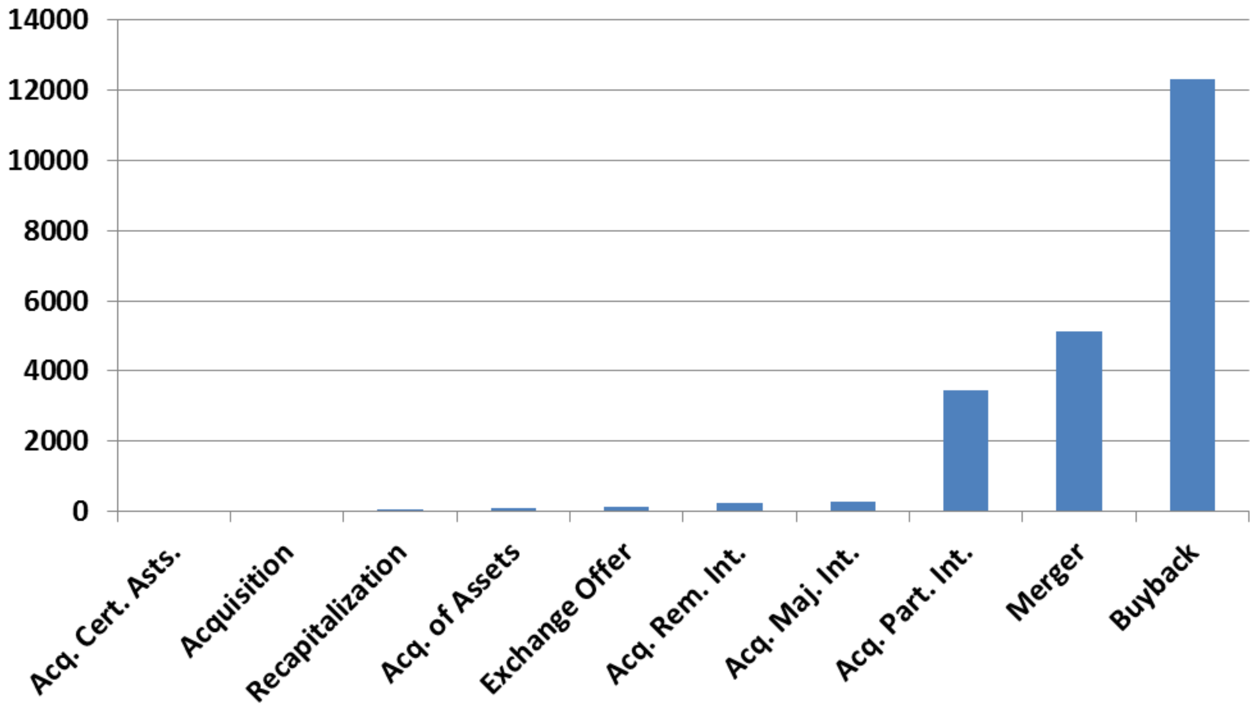
Event	Description
Acquisition	A 100% spinoff/equity carveout is considered an acquisition by shareholders
Acq. of Assets	The Acquisition of Assets is when all assets of a company, subsidiary, division, or branch are acquired. When the consideration sought for a company's acquisition is not given, it is also classified in this category.
Acq. Cert. Asts.	The Acquisition of Certain Assets occurs when certain assets of a company, subsidiary, division, or branch are acquired.
Acq. Maj. Int.	The Acquisition of Majority Interest is when the acquiror must have held less than 50%, and be seeking to acquire 50% or more, but less than 100% of the target company's stock.
Acq. Part. Int.	The Acquisition of Partial Interest is when the acquiror must have held less than 50%, and be seeking to acquire less than 50%, or must have held over 50% and be seeking to acquire less than 100% of the target company's stock.
Acq. Rem. Int.	The Acquisition of Remaining Interest is when the acquiror must have held over 50%, and be seeking to acquire 100% of the target company's stock.
Buyback	A company buys back its equity securities or securities convertible into equity, either on the open market, through privately negotiated transactions, or through a tender offer. Board authorized repurchases are included.
Exchange Offer	An Exchange Offer is when a company offers to exchange new securities for its equity securities outstanding or its securities convertible into equity.
Merger	A combination of businesses takes place or 100% of the stock of a public or private company is acquired. The acquiror must have held less than 50%, and be seeking to acquire 100% of the company's stock.
Recapitalization	A company undergoes shareholders' leveraged recapitalization in which the company issues a special one-time dividend (in the form of cash, debt securities, preferred stock, or assets) allowing shareholders to retain an equity interest in the company.

Sources: Thomson Reuters, Wolfe Research Luo's QES

Figure 2 shows the frequency of various M&A event types in the past 30 years. Obviously, "Buyback" has the highest number, followed by "Merger and acquisition of partial interest". The other event types have much lower frequencies. Therefore, based on the frequency and the nature of each event type, we re-group the events into three different categories: "Buyback", "Merger" and "Other":

- "Buyback" is the "Buyback" defined by TRSDC;
- "Merger" is the "Merger" category defined by TRSDC; and
- "Other" includes all the other event types.

Figure 2 Event Type Frequency



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Following the convention in event study, we compute event excess return as:

$$r_{i,t}^{Excess} = r_{i,t-1\to t+1} - r_{market,t-1\to t+1}$$

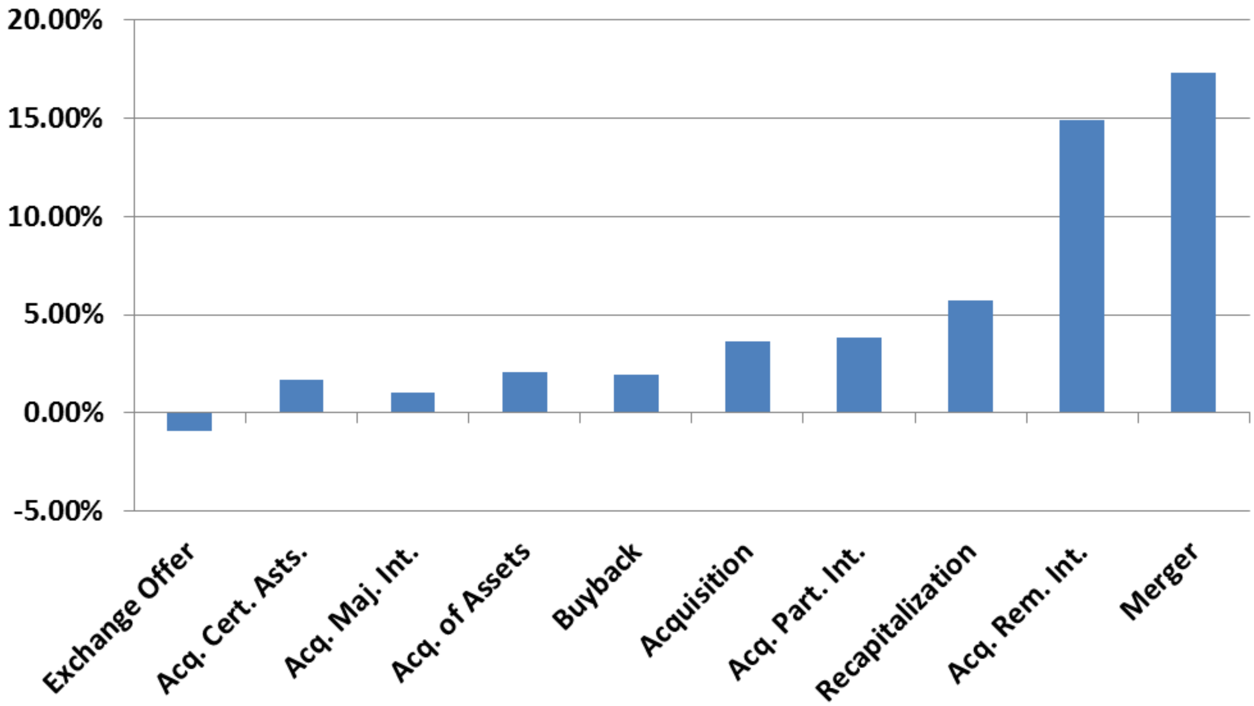
Where,

$r_{i,t-1\to t+1}$ is the return of stock i from day $t-1$ (i.e., one business day prior to the announcement) and $t+1$ (i.e., one business day post the announcement); and

$r_{market,t-1\to t+1}$ is market return during the same three-day window

As shown in Figure 3, "Merger" generates the highest excess return during the three-day event window, followed by "Acquisition of the remaining interest". Other than "Exchange offer", all other event types deliver positive excess returns on the announcement day.

Figure 3 Excess Return around the Event Date

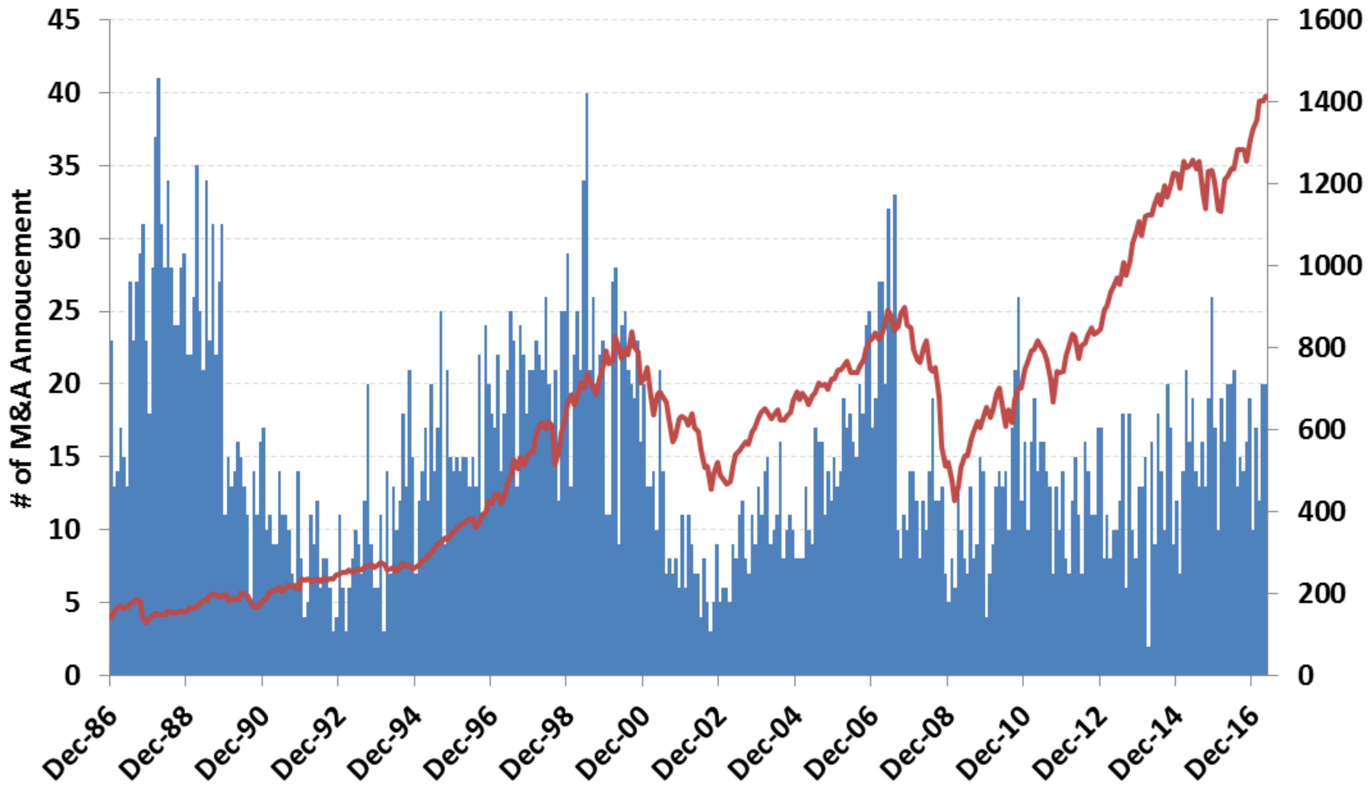


Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

FREQUENCY OF MERGER

We now shift our attention on the actual "Merger" transactions. As shown in Figure 4, the deal frequency tends to move in line with the overall equity market. The number of deals peaks at the top of the market and plunges during financial crises. In the long run, there are about 15 mergers per month or 0.5% of the universe.

Figure 4 Number of M&A Announcements



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

MOTIVATION OF TAKEOVER PREDICTION

The major motivation for predicting the M&A target is the huge excess return on the announcement date. Figure 5 shows the event study for the mergers before and after the announcement date. There are a few interesting findings:

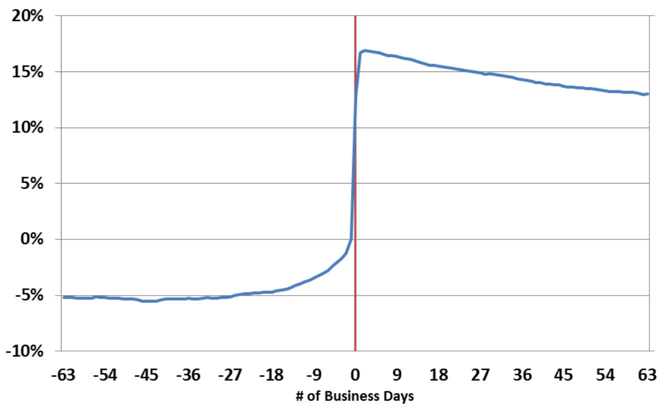
- The announcement day return is highly skewed. The average return on the announcement date is almost 20%, while the median is less than 10%.
- Prior to the announcement date, the median excess return is negative. This is because the merger target companies are on average underperforming firms. Therefore, unless we have high prediction accuracy, the drag from these underperforming stocks is likely to offset any gain from actual takeover targets.
- After the announcement date, the average (and median) excess return tends to be negative, mainly due to failed transactions. If a deal fails to go through[1], the stock price of the target firm tends to fall sharply (typically close to where it were before the announcement). This is why the plain vanilla risk arbitrage indices mostly do not live up to the expectation – simply investing in

[1] A deal can break down due to multiple reasons, e.g., blocked by a regulator, rejected by target company's shareholders, etc.

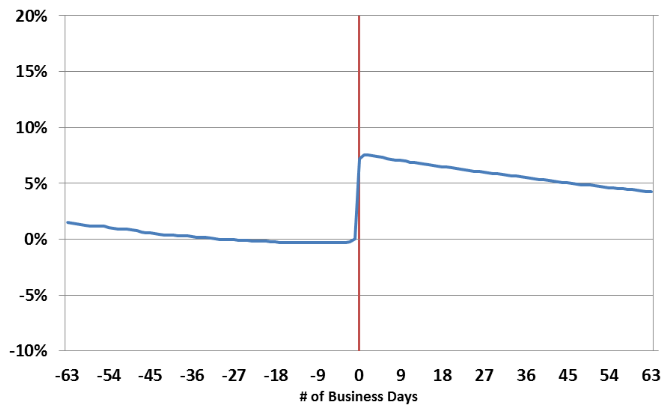
every single M&A transaction is unlikely to deliver superior return. How to systematically participate in M&A arbitrage will be covered in a forthcoming research paper.

Figure 5 Event Study

A) Average Cumulative Excess Returns



B) Median Cumulative Excess Returns



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Characteristics of Actual M&A Targets

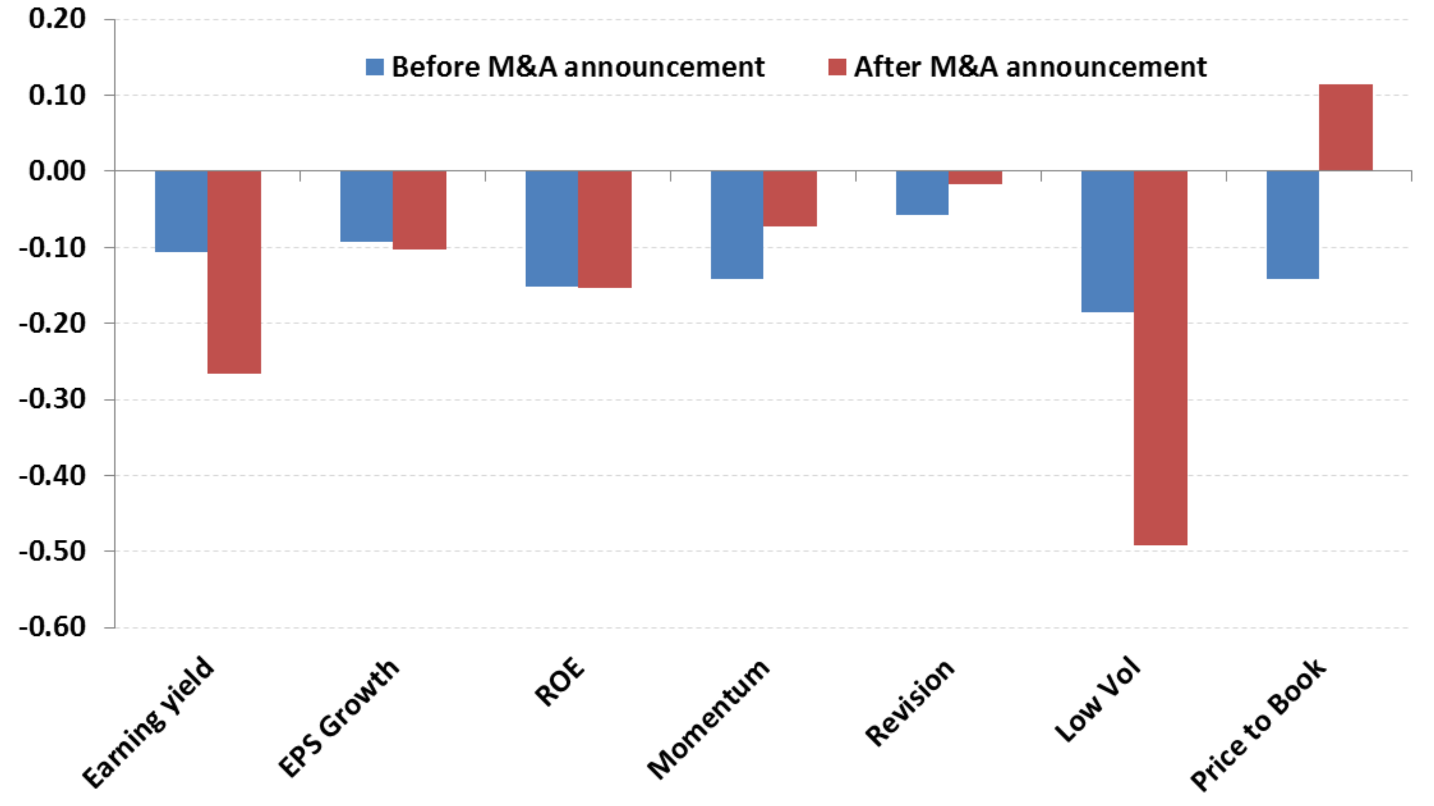
A natural starting point to identify takeover targets is from our common factors (see Luo, et al [2017b] for a detailed list and description of these factors). When we look at the factor exposure the month before and after the announcement (see Figure 6), we find that typically M&A targets are companies that active managers would typically bet against. The exposures are calculated based on sector neutral normalized factor score. Therefore, a negative exposure means the average value is lower than the sector median.

- Targets companies are more expensive based on earnings yield.
- They have poorer fundamentals, e.g., slower earnings growth and less profitable (lower ROE).
- They have undesirable trading characteristics, e.g., negative price momentum and higher volatility.
- They are less favored by sell-side analysts, as reflected in the negative earnings revision.

The poor average fundamental properties of takeover targets are as expected. This is why potential acquirers think they can either turn around these underperforming firms and/or capture some synergy from combining the two companies.

Furthermore, we also note that these fundamental factors further deteriorate after the announcement, especially for valuation ratios. Target companies are typically offered at a premium over their prevailing share prices.

Figure 6 Factor Exposure, before and after M&A Announcement



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

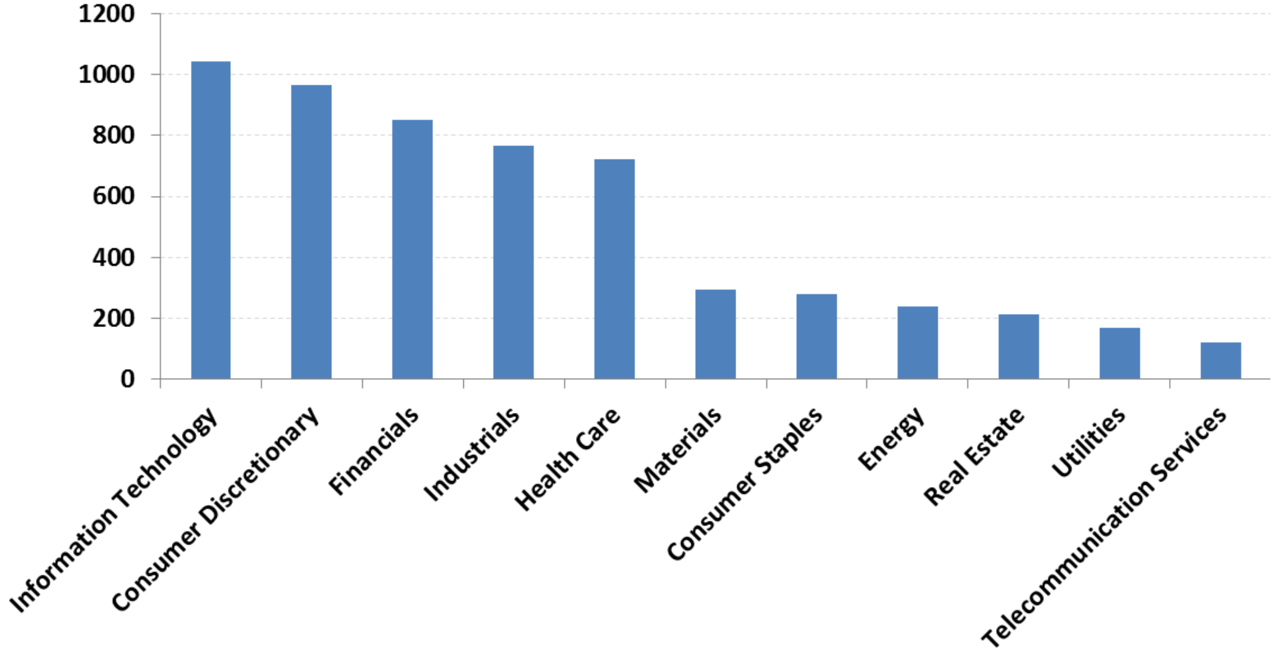
M&A Coverage by Sector

Different sectors also face different merger frequency. For example, information technology sector has more merger deals than other sectors (as shown in Figure 7). However, we also need to keep in mind that smaller sectors (defined by the number of companies in the sector) naturally tend to have fewer transactions. Therefore, to make a fair comparison, we adjust the frequency by the actual number of stocks in each sector:

$$\frac{Average\ \#\ of\ M\&A\ transactions}{Average\ \#\ of\ Companies}$$

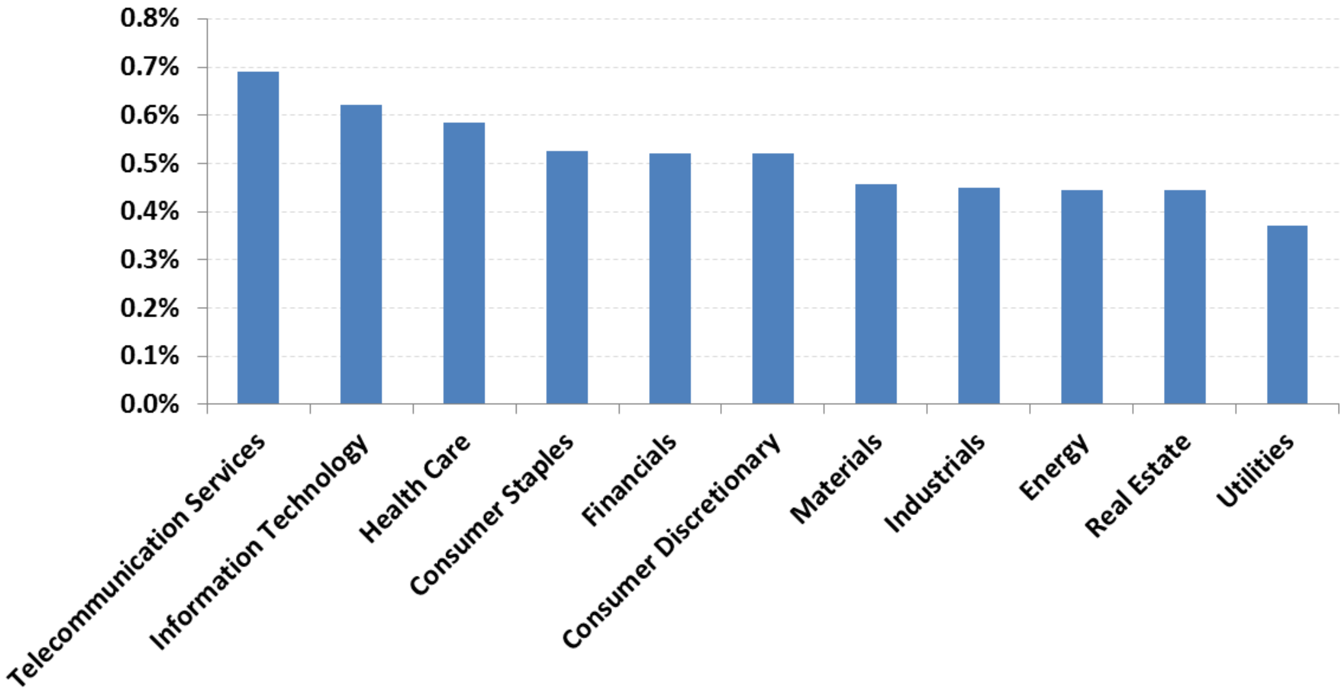
After the adjustment, telecom services and information technology sectors are now on the top see Figure 8).

Figure 7 M&A Frequency, by Sector



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Figure 8 Adjusted M&A Frequency, by Sector



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

TAKEOVER PREDICTION FACTORS

In this section, we study the potential factors that can help us to identify takeover targets. We start from a standard academic literature review. Then we establish a framework to measure the predictive power of each factor. Next, we shift our attention to traditional stock-selection factors. The most interesting section focuses on how an unconventional data source – SEC's EDGAR corporate filling database can be used to improve our forecasting accuracy. Lastly, we find the frequency of previous M&A transactions in the same industry also contain useful information about future takeover probability.

ACADEMIC RESEARCH ON M&A PREDICTION

There are a number of academic research papers on takeover prediction, but most of them conclude that it is an extremely difficult task.

Early works (see Hasbrouck [1985], Palepu [1986], Morck, et al [1988], Ambrose, et al [1992] and Shivdasani [1993]) mostly use accounting ratios (and to a lesser extent, ownership structure). Most of these papers suffer from a small sample bias, less than 500 companies.

Hoberg, et al [2010] and Routledge, et al [2013] find text mining of the SEC 10-K filings can be useful in predicting acquirers and targets. Routledge, et al [2013 find that the word and phrase features can improve the accuracy of predicting acquirers, above and beyond standard financial variables. However, for the more difficult and relevant task of predicting targets, the incremental benefit of text mining is much more modest.

Cocco, et al [2013] examine corporate pension plans. They suggest that firms sponsoring a DB (Defined Benefit) pension plan are less likely to be takeover targets. They argue that sponsoring a DB pension plan is subject to greater information asymmetry than similar firms without such plans, which deters potential acquirers. Because the complexity in evaluating pension liabilities, the sponsoring firm's managers are likely to have information advantage over outsiders.

The biggest issue with academic research is that researchers are primarily interested in understanding the variables differentiating takeover targets, rather than building an out-of-sample predictive model. None of the above mentioned research discusses how to translate from takeover prediction into an actual investment strategy, where we need to convert from takeover probability into cross-sectional rankings. Furthermore, most finance and economic research uses linear regressions exclusively, which ignores the potential higher predictive power offered by more sophisticated machine learning techniques.

Because the primarily purpose of our research is to eventually form an investable strategy on potential takeover targets – different from standard academic research, we need to first define our backtesting methodology of how to measure the out-of-sample predictive power of each factor.

FRAMEWORK OF M&A PREDICTION

Our goal is to predict the list of takeover targets for the following month. In a typical factor backtesting (see Luo, et al [2017b] on the detail of factor backtesting in stock selection), we compute rank IC (Information Coefficient), i.e., the Spearman rank correlation between: 1) the ranking of stocks based on a given factor; and 2) the ranking of returns of the same list of stocks in the subsequent month. If a factor has a perfect predictive power of future stock returns, rank IC should be 100% or -100%.

Similarly, if a factor has no ability to differentiate outperforming and underperforming stocks, rank IC should be close to zero.

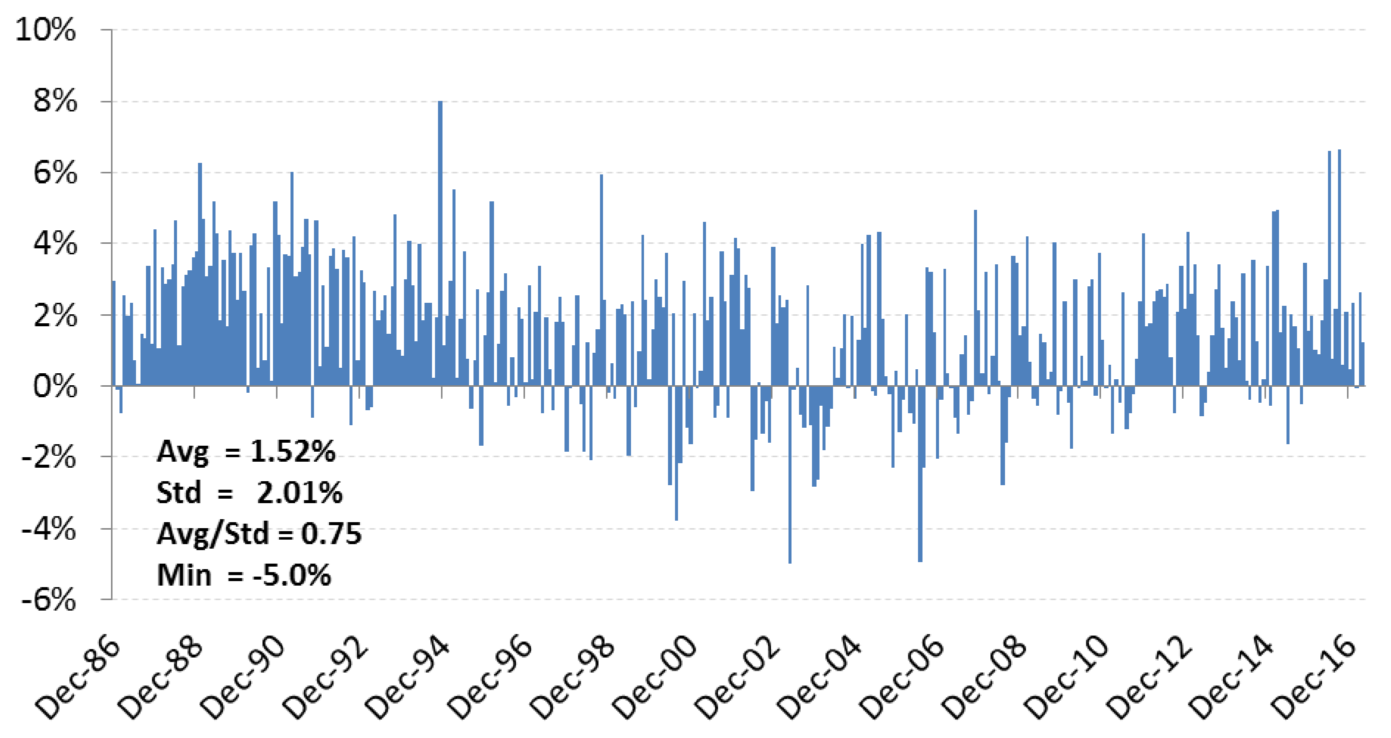
We define M&A IC (Information Coefficient) as the Spearman rank correlation between: 1) the ranking of stocks based on a factor; and 2) a vector of zero's and one's where ones represent actual takeover targets in the subsequent month. To adjust for the time variation of the predictive power, we also compute the risk-adjusted M&A IC as:

$$\frac{Average\ IC}{Standard\ Deviation\ of\ IC}$$

We need to keep in mind that takeover transactions are rare events. Even in the most prolific sector – telecom services, there is less than 0.7% companies become targets in a given month. As a result, the M&A IC is likely to be much lower than the stock-selection IC we use to measure return predictability.

For example, Figure 9 shows the realized one year daily volatility as a factor in takeover prediction. The average M&A IC is 1.52% (with a risk-adjusted IC of 0.75%) over the past 30 years. The M&A IC remains positive most of the times. More volatile stocks are more likely to become M&A targets. In this case, volatility has a decent predictive power over the subsequent month's takeover targets.

Figure 9 M&A IC for the Volatility Factor



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

TRADITIONAL STOCK-SELECTION FACTORS

We start our backtesting from traditional stock-selection factors. Because the length of history varies by factors. To be consistent, we limit our backtesting to those factors with at least 15 years of data

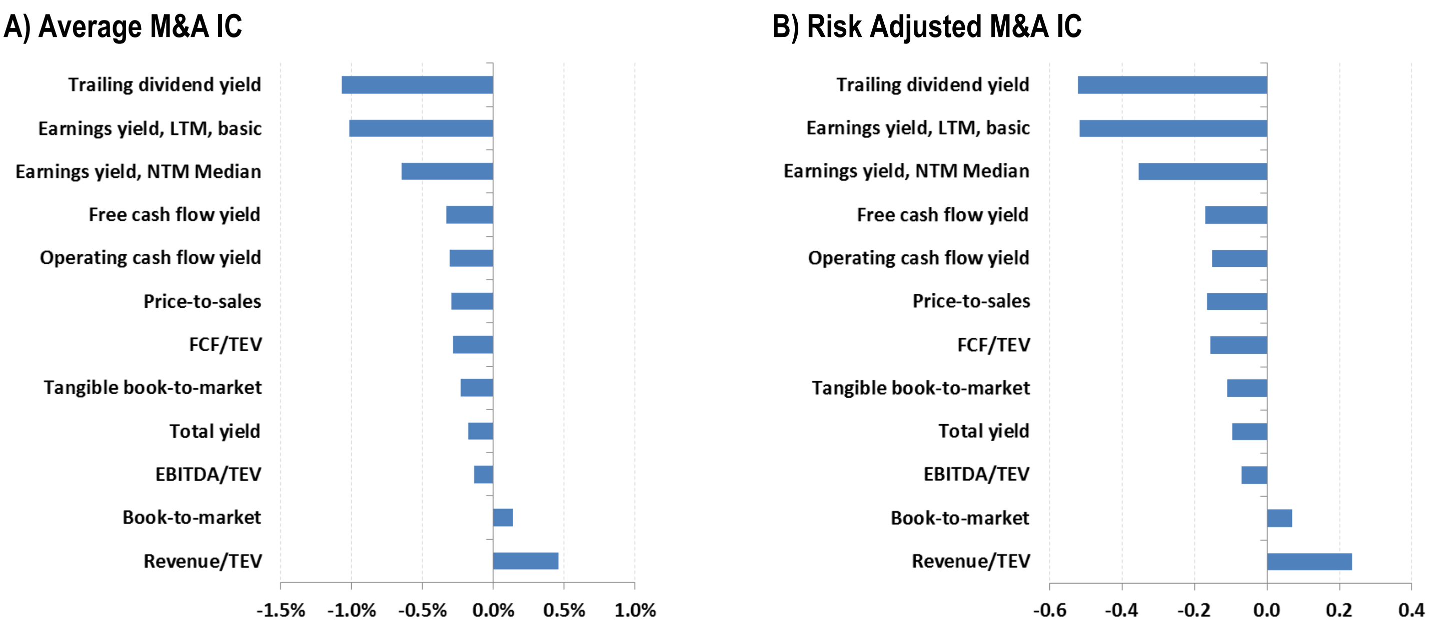
Value

Academic literature has a long history of documenting the value phenomenon. Value factors can also be based on other fundamental performance metrics of a company such as dividends, earnings, cash flow, EBIT, EBITDA and sales.

As shown in Figure 10, takeover companies to be:

- More expensive based on dividend yield, earnings yield, cash flow yield, tangible book-to-market, EBITDA/EV; but
- Cheaper based on price-to-sales, book-to-market, and revenue/TEV

Figure 10 Value Factor Performance



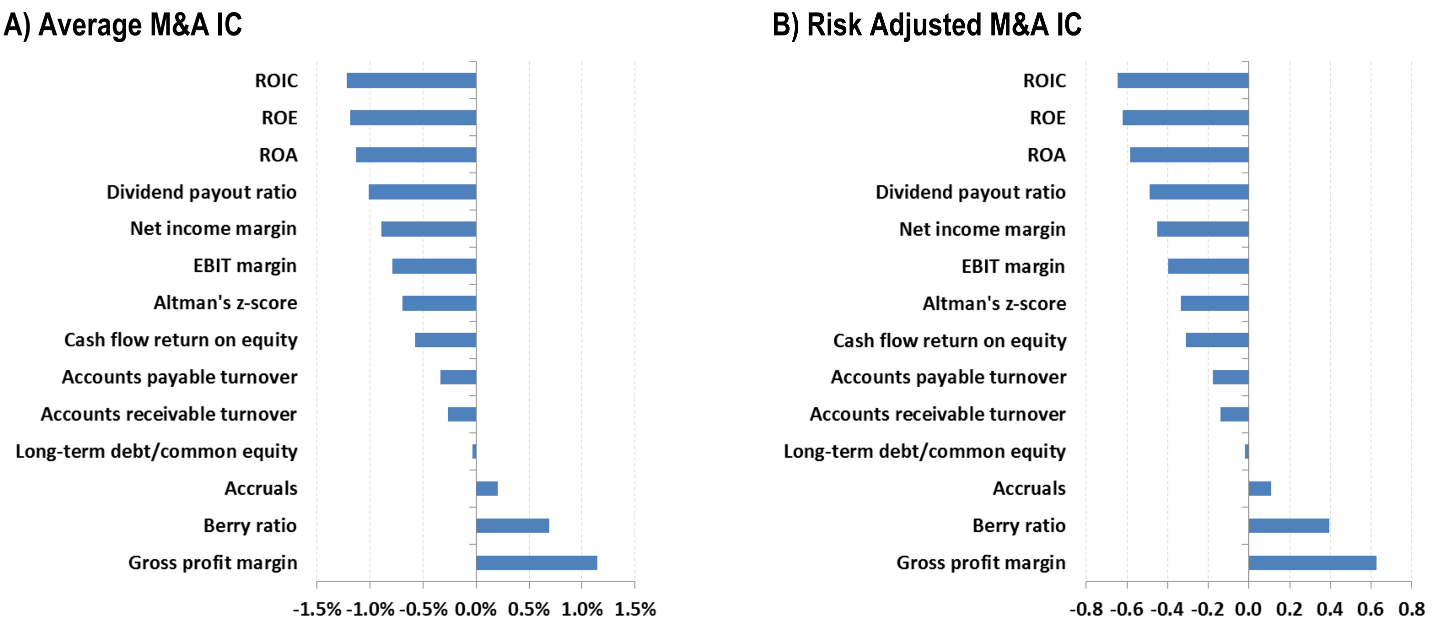
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Quality

On the quality front, Figure 11 shows a few interesting patterns about target companies:

- They are less profitable on most metrics, with the exception of gross profit margin. A positive gross margin and a negative net margin means target firms are particularly inefficient in general management and administrative activities.
- They have lower corporate governance standard but possible higher dividend paying sustainability, as reflected by the lower payout ratio.
- They have slightly lower financial leverage and lower bankruptcy risk.

Figure 11 Quality Factor Performance

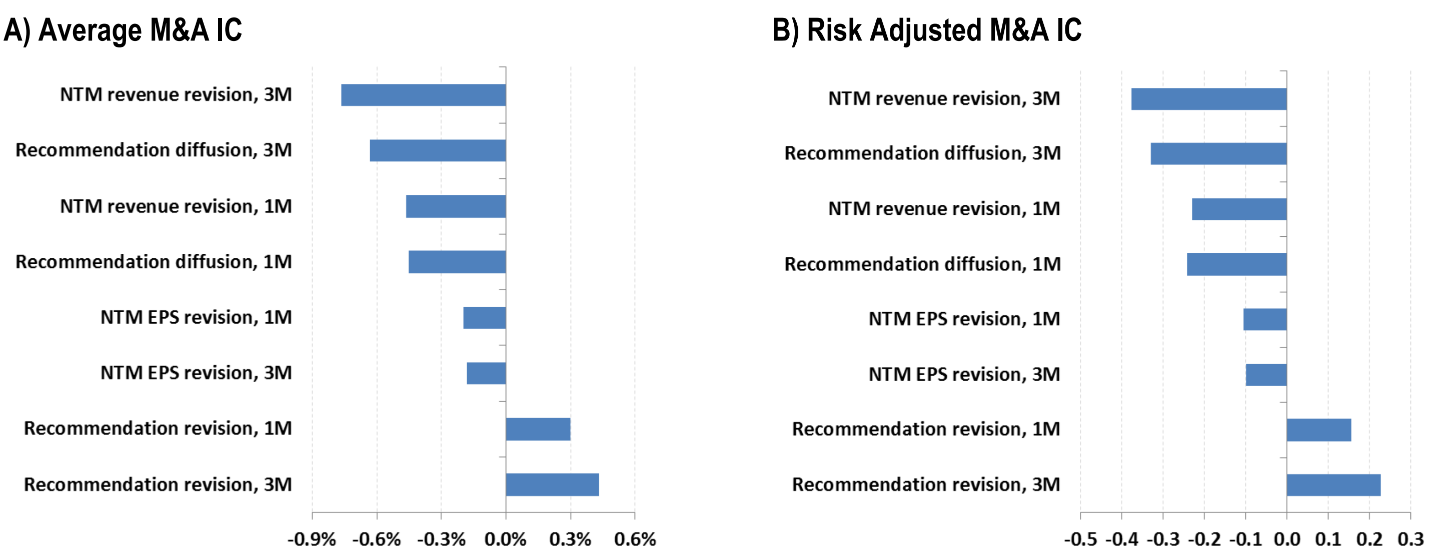


Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Sentiment

As shown in Figure 12, takeover companies are disliked by sell-side analysts by all common measures. It is important to note that consensus analyst revision is scored from one to five, where one represents "strong sell'. Therefore, a positive recommendation revision actually means analysts have downgraded a company.

Figure 12 Sentiment Factor Performance

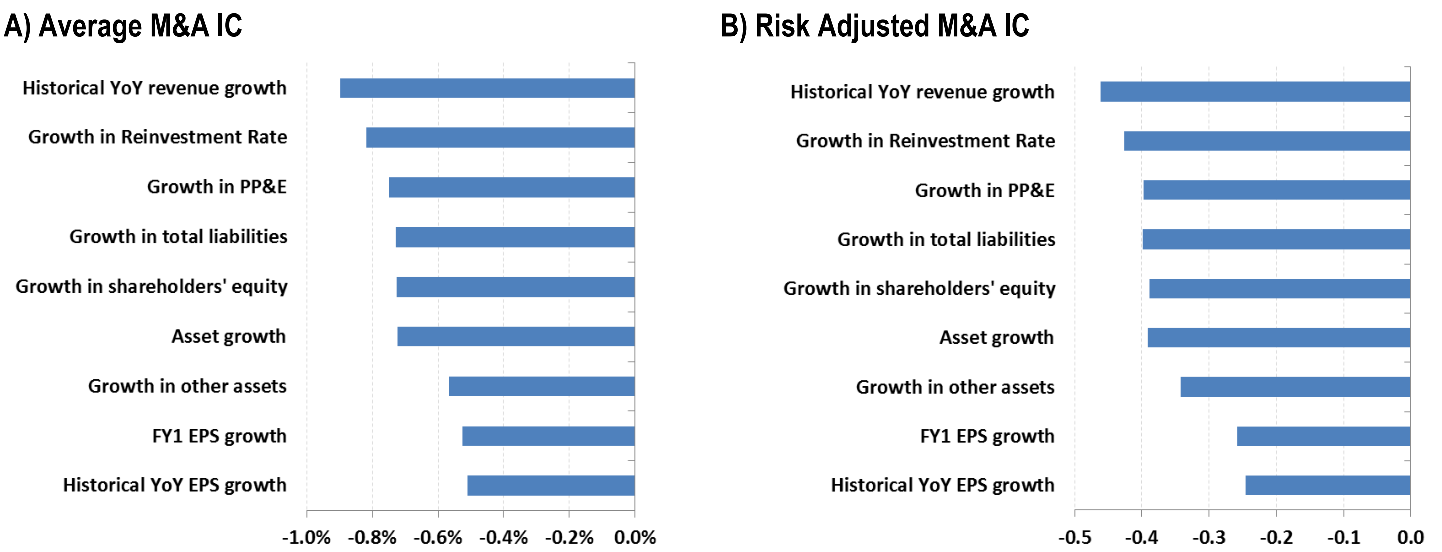


Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Growth

All the growth related factors show negative M&A IC (see Figure 13). One major reason that an acquirer wants to buy a target is to turn around a slow growth company, in order to generate merger synergy and outsized profit.

Figure 13 Growth Factor Performance



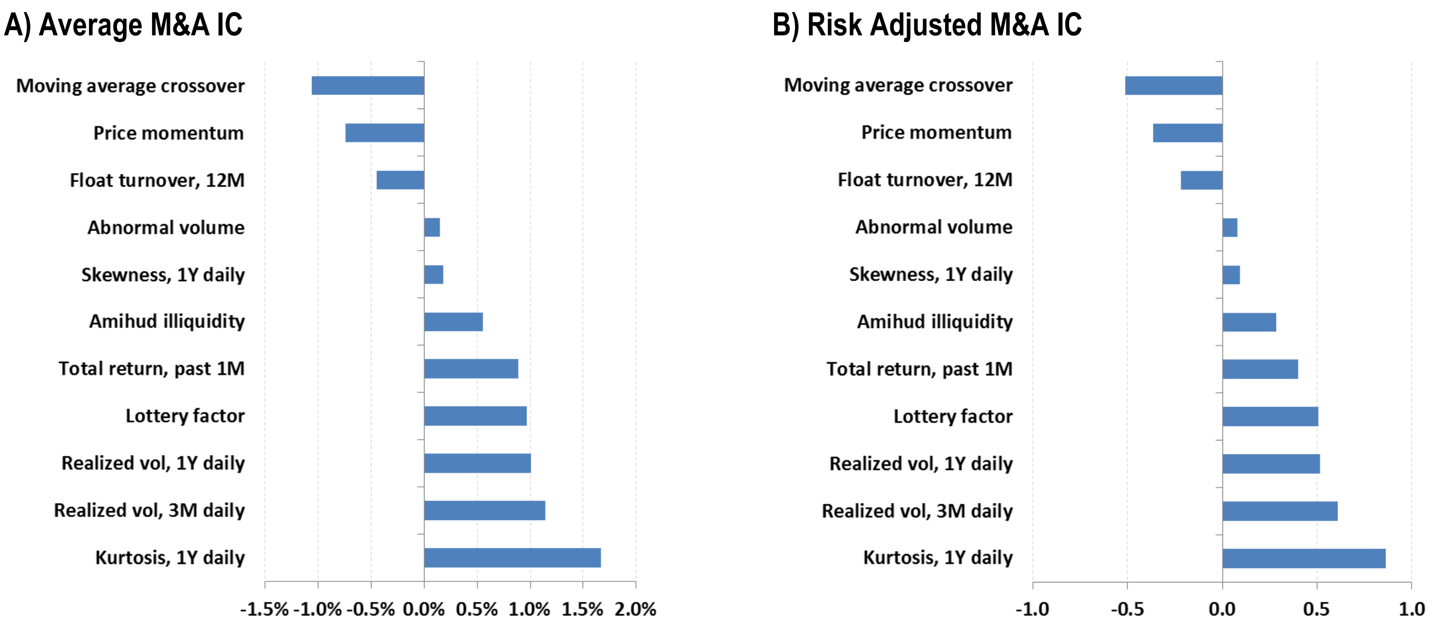
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Technical Factors

Lastly, based on technical factors (including price momentum and mean reversal signals), we observe the following patterns for the target firms (see Figure 14):

- They have poor price momentum.
- They have lower liquidity, e.g., float turnover, Amihud illiquidity.
- They are more volatile.
- The most interesting aspect is that target companies tend to rally sharply immediately before M&A announcements (positive one-month return), coupled by higher abnormal trading volume, and excess Kurtosis. This is possible due to the fact that M&A transactions are sometimes anticipated by certain market participants.

Figure 14 Technical Factor Performance



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

EDGAR Filing Signals

As hinted by Hoberg, et al [2010] and Routledge, et al [2013], according to SEC rules, public companies face intense filing requirements. Some of the filings might be related to the potential M&A intent and transactions.

In our past research (see Rohal, et al [2017]), we find text mining of corporate filing database, e.g., the EDGAR database, can discover important and orthogonal information. We have a well-developed technology system to script the EDGAR database and conduct NLP (Natural Language Processing) on corporate filing data. In this section, we explore opportunities of using corporate filing data to predict takeover targets. In particular, we expect the number of filings for certain forms may have useful information about future M&A transactions. To be conservative, all filing based signals are computed at the monthly frequency with a lag of three days.

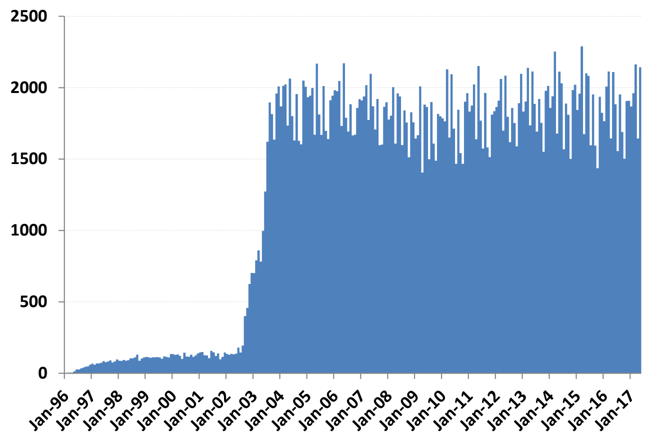
Form 3 and Form 4

Form 3 and Form 4 are SEC filings that relate to insider trading. Every director, officer or owner of more than 10% of a class of equity security registered must file with the SEC a statement of ownership regarding such a security. The initial filing is on Form 3 and changes are reported on Form 4. The forms contain information on the reporting person's relationship to the company and on purchases and sales of such equity securities.

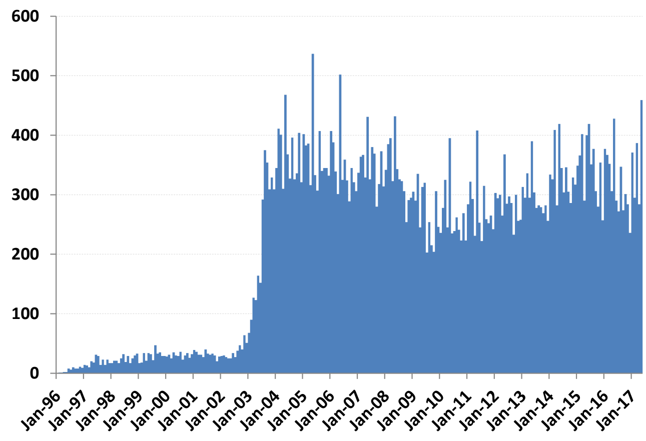
Figure 15 shows the number of the form 3 and form 4 filings related to Russell 3000 index constituents. We observe a significant jump in the number of fillings in 2002. Form 4 has a much higher number of filings than Form 3, because Form 3 is just the initial filing, while Form 4 includes all the follow up insider transactions.

Figure 15 Number of Filings, Form 4 and Form 3

A) Form 4



B) Form 3



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

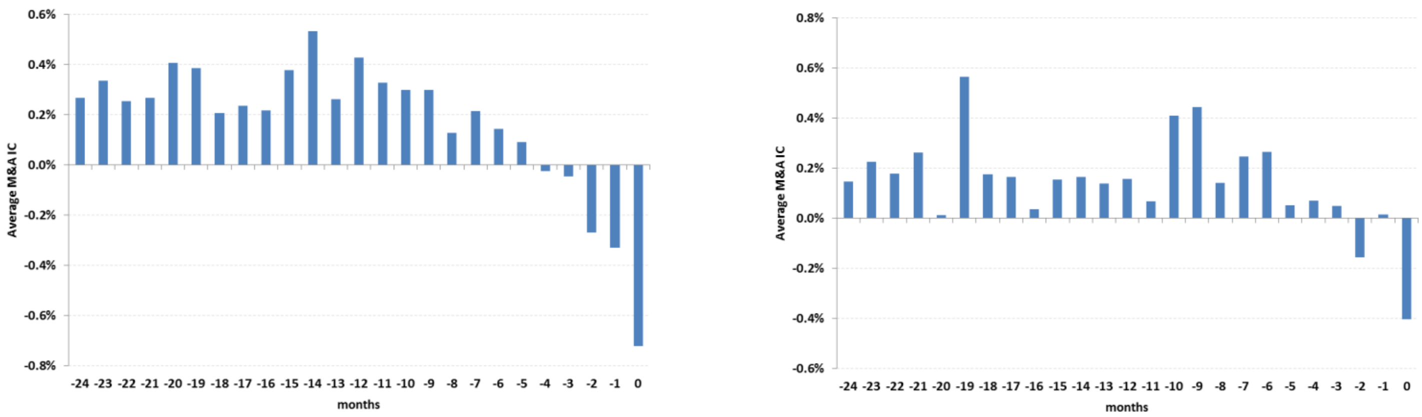
Due to the asymmetric information, company management and insiders should know their own firm better than outside investors. Therefore, for underperforming companies but expecting to be acquired, insiders may engage more active transaction of their own stocks. Figure 16 shows the M&A IC for the number of the filings of Form 3 and Form 4, from 24 months prior to the announcement to the announcement month. We find that there are heavier than normal insider transactions from two years prior to the announcement date until six months right before the deals are announced. The number of insider trades is actually much lighter one quarter before the announcement date, possibly because either lock-out period prevents insiders from trading right before transactions, or insiders want to avoid being perceived as trading on private information.

Therefore, we construct two signals based on the number of Form 3 and Form 4 filings in the past two and five year lookback windows, excluding the most recent quarter. We also apply an exponential decay with a half-life of one year and 2.5 years, respectively. In addition, we define two binary signals, which is one if there is a filing in the past two (or five) year lookback window, or zero otherwise. Below is a summary of our signals:

- Form 3 (and Form 4) two-year signal – total number of filings in the past two years, computed monthly, with an exponential decay (half-life of one year)
- Form 3 (and Form 4) five-year signal – total number of filings in the past five years, computed monthly, with an exponential decay (half-life of 2.5 years)
- Form 3 (and Form 4) two-year binary signal – equals to one if a company has had any filing in the past two years (zero otherwise)
- Form 3 (and Form 4) five-year binary signal – equals to one if a company has had any filing in the past five years (zero otherwise)

Figure 16 Average M&A IC, # of Filings, by Month Prior to M&A Announcement

A) Form 4 **B) Form 3**



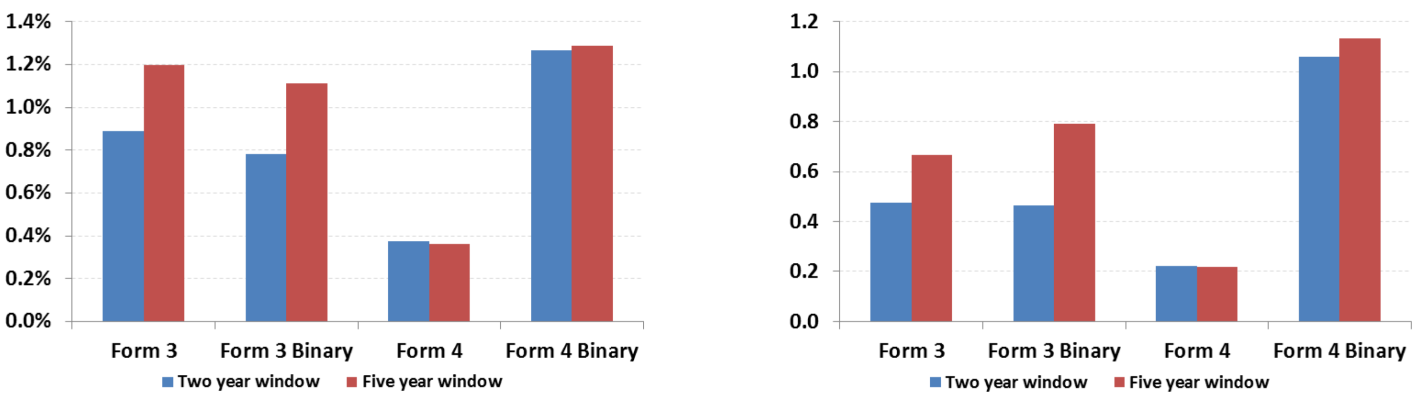
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Figure 17 shows the performance of the eight signals based on the Form 3 and Form 4. Signals based on the longer five-year window generally have better performance than the ones using a shorter two-year window.

In addition, for Form 4, the binary signals are significantly better than the ones based on actual number of filings. As a reminder, Form 4 represents subsequent transactions and there are many. Therefore, whether there is an insider transaction matters more than the frequency of insider trades.

Figure 17 Form 3 and Form 4 Factor Performance

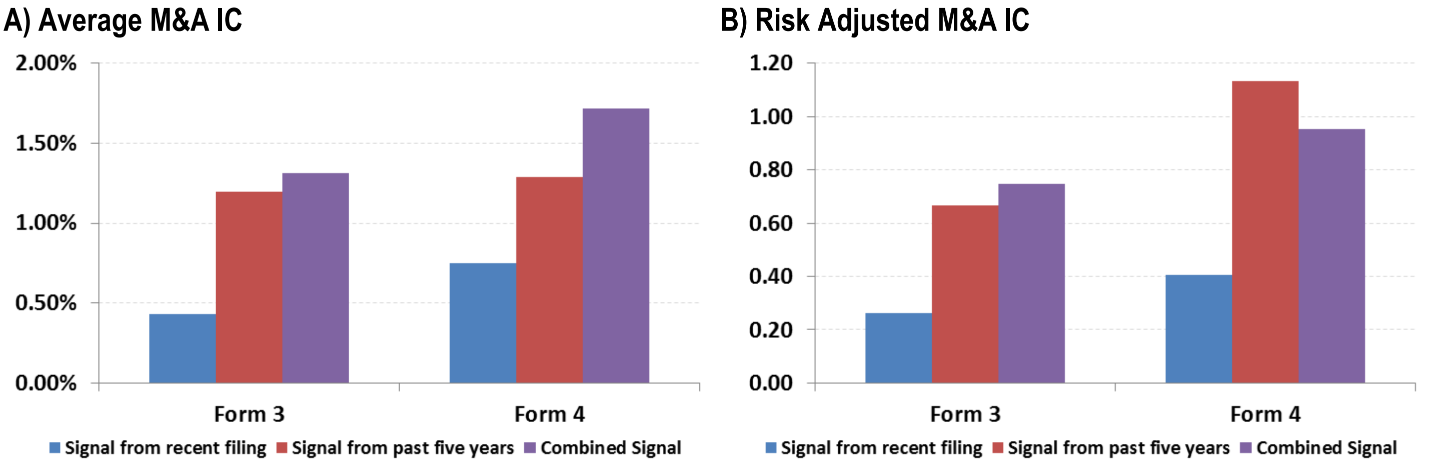
A) Average M&A IC **B) Risk Adjusted M&A IC**



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

In addition, since there are fewer filing during the month prior to the M&A announcement (see Figure 16). We can use the number of Form 3 (and Form 4) filings in the immediate prior month as a negative factor. As shown in Figure 18, the signal from the most recent month filing is less predictive than the ones using the past five-year data. However, the most recent month signals add potential diversification benefit.

Figure 18 Signal from Last Month Adds Diversification Benefit



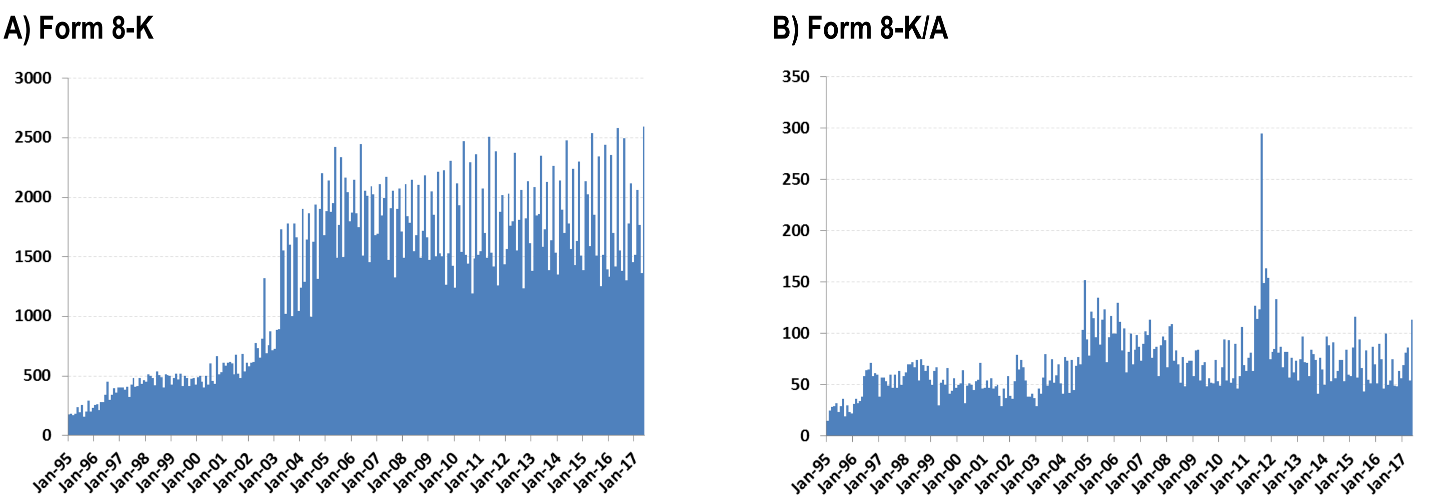
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Form 8-K and 8-K/A

Form 8-K is a very broad form used to notify investors of specified events that may be important to shareholders or the SEC. This is one of the most common types of forms filed with the SEC. The report notifies the public of events reported including acquisition, bankruptcy, resignation of directors and other material events. Form 8-K/A is the amendment of 8-K filing.

The popularity of the Form 8-K is witnessed in Figure 19 (A). As expected, the number of amendments in Form 8-K/A is much less frequent (Figure 19 B).

Figure 19 Number of Filings, Form 8-K and Form 8-K/A

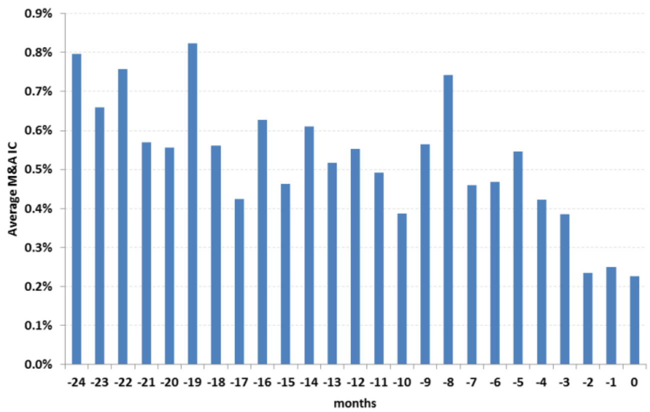


Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

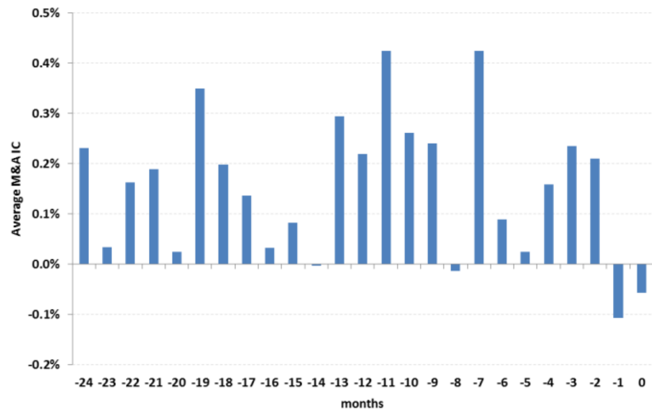
We expect companies involved in M&A activities (either as acquirers or targets), with corporate governance issues, or materials news are likely to future takeover targets. Empirical data supports our hypothesis. As shown in Figure 20 (A) and (B), a more frequent filing of Form 8-K (and to a lesser extent, Form 8-K/A) in the past two years is associated with a higher probability of being acquired.

Figure 20 Average M&A IC, # of Filings, by Month Prior to M&A Announcement

A) Form 8-K



B) Form 8-K/A



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

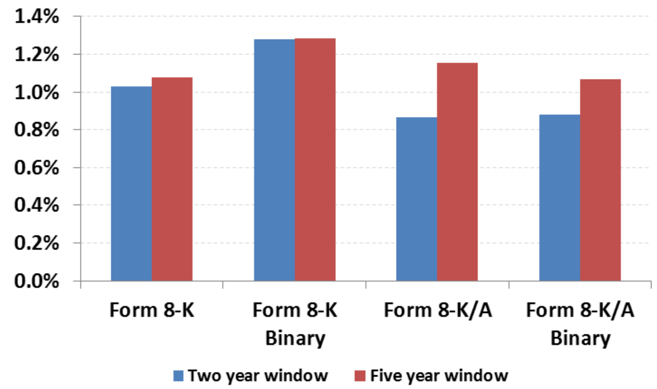
We construct the same set of eight factors as in the previous section.

- Form 8-K (and Form 8-K/A) two-year signal – total number of filings in the past two years, computed monthly, with an exponential decay (half-life of one year)
- Form 8-K (and Form 8-K/A) five-year signal – total number of filings in the past five years, computed monthly, with an exponential decay (half-life of 2.5 years)
- Form 8-K (and Form 8-K/A) two-year binary signal – equals to one if a company has had any filing in the past two years (zero otherwise)
- Form 8-K (and Form 8-K/A) five-year binary signal – equals to one if a company has had any filing in the past five years (zero otherwise)

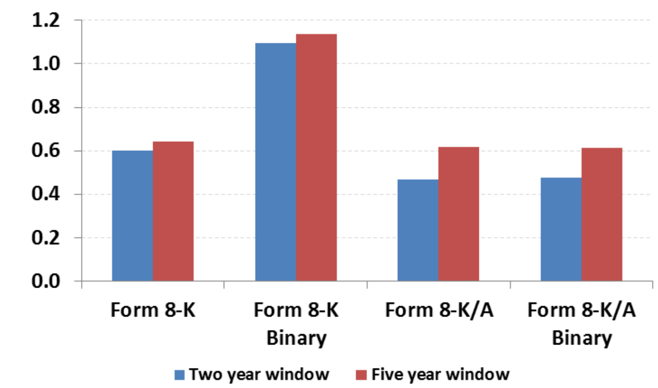
As shown in Figure 21, a longer five-year lookback window produces slightly better signals. Similar to what we see for Form 3/Form 4, for the more popular Form 8-K, binary signals perform better than frequency-based factors. As long as a company has ever filed a Form 8-K, it increases its takeover probability.

Figure 21 Performance

A) Average M&A IC



B) Risk Adjusted M&A IC



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Form SC 13D, SC 13D/A, SC 13G and SC 13G/A

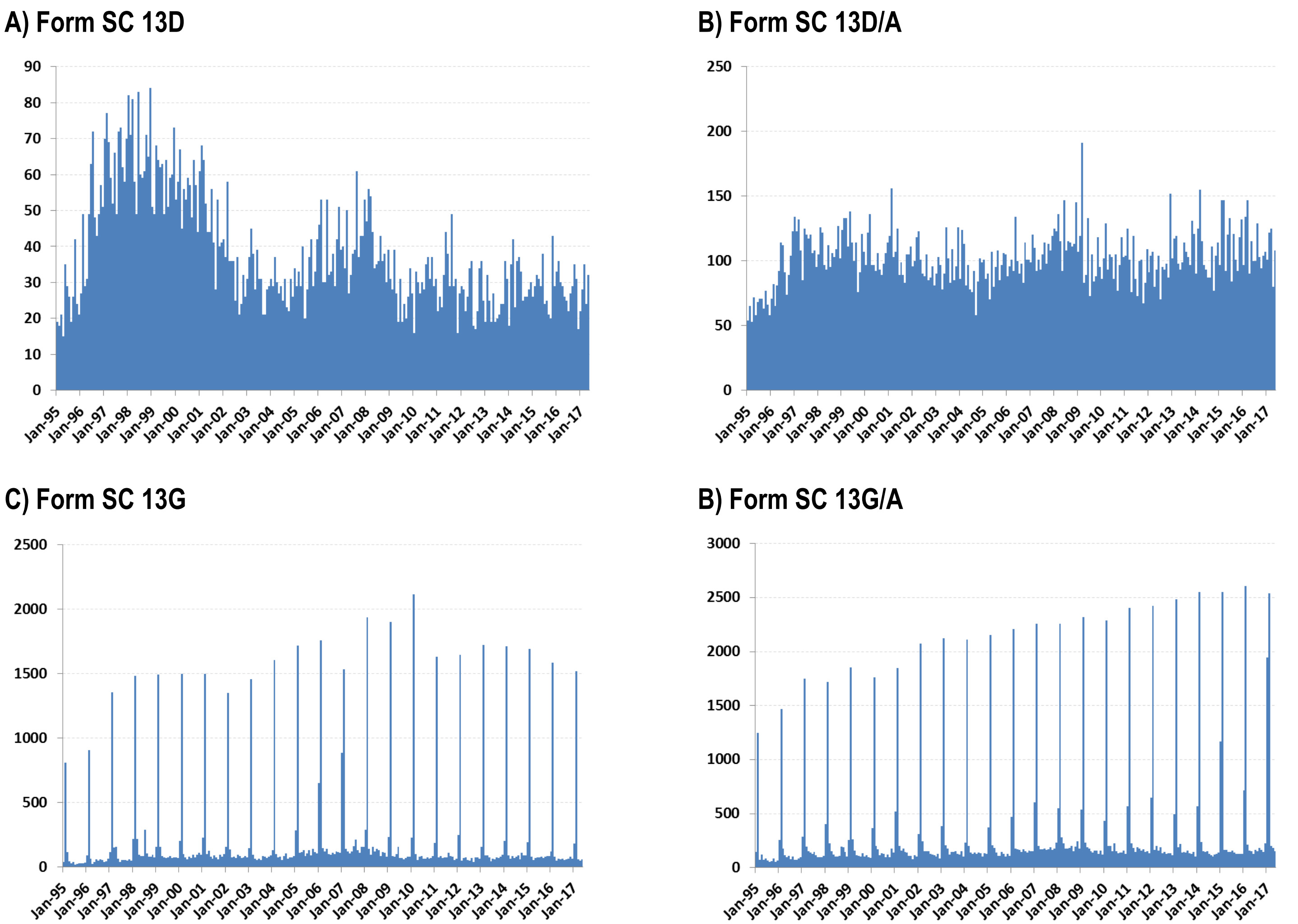
The Schedule (SC) 13D is a form that must be filed when a person or group acquires more than 5% of any class of a company's shares. This information must be disclosed within 10 days of the transaction. Rule 13D also requires the owner to disclose any other person who has voting power or the power to sell the security.

SC 13G is similar to SC 13D used to report a party's ownership of stock that is over 5% of the company. SC 13G is shorter and requires less information from the filing party. To be able to file SC 13G instead of SC 13D, the party must own between 5 and 20% in the company. The party acquiring the stake in the company must only be a passive investor and does not intend to exert control. If these criteria are not met or if the size in the stake exceeds 20%, a SC 13D must be filed.

SC 13D/A and SC 13G/A are the amendment of SC 13D and SC 13G, respectively.

Figure 22 shows the number of filing for ownership related forms. There are far more SC 13G and 13G/A than SC 13D and 13D/A, indicating that the majority of controlling shareholders are passive investors. We observe a very strong annual seasonal pattern for SC 13G and 13G/A filings, because passive controlling shareholders have to re-file these forms every year.

Figure 22 Number of Filings, SC 13D, SC 13D/A, SC 13G and SC 13G/A

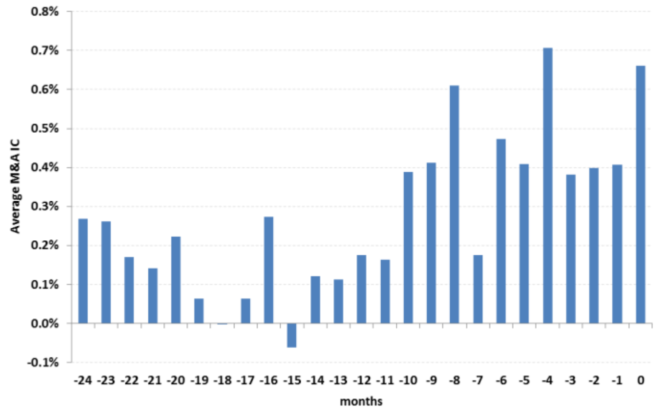


Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

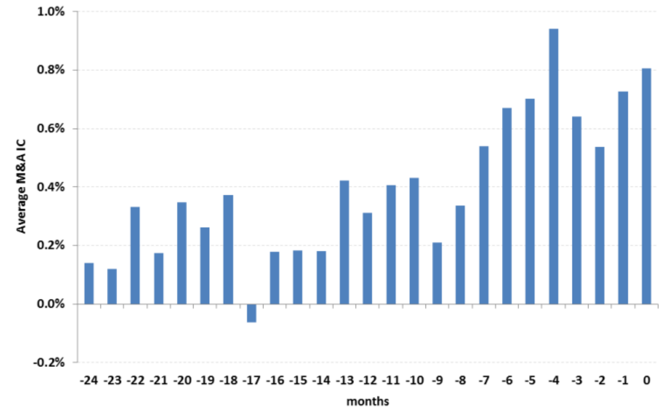
One of the commonly used techniques in M&A is to start from accumulating target firm's shares on the open market, and then launch a full-fledged takeover afterwards. It is also often used by activist investors, who may eventually push forward their portfolio companies to sell to potential acquirers. Therefore, we expect the frequency of SC 13D, 13D/A, 13G, 13G/A to be associated with higher takeover probabilities. Figure 23 confirms our hypothesis.

Figure 23 Average M&A IC, # of Filings, by Month Prior to M&A Announcement

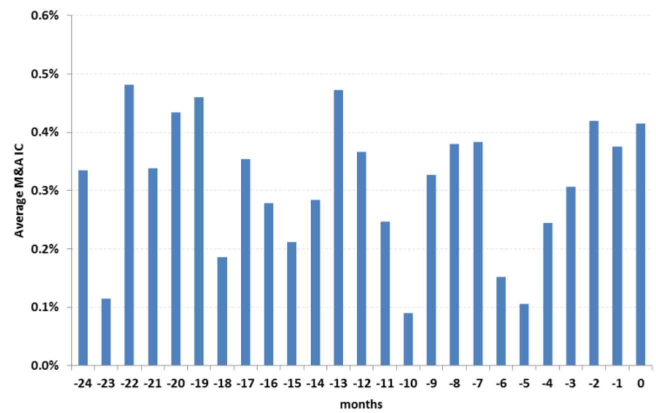
A) Form SC 13D



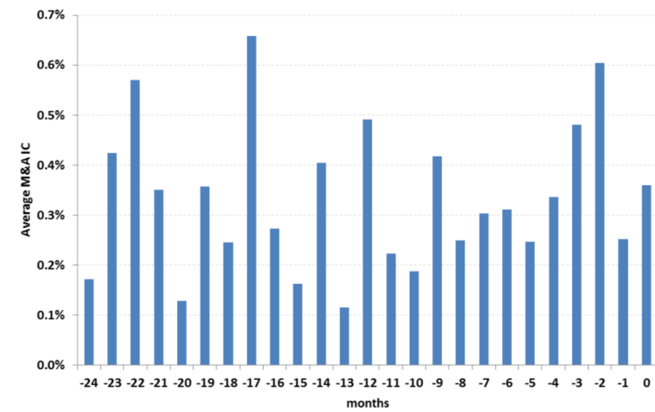
B) Form SC 13D/A



C) Form SC 13G



B) Form SC 13G/A



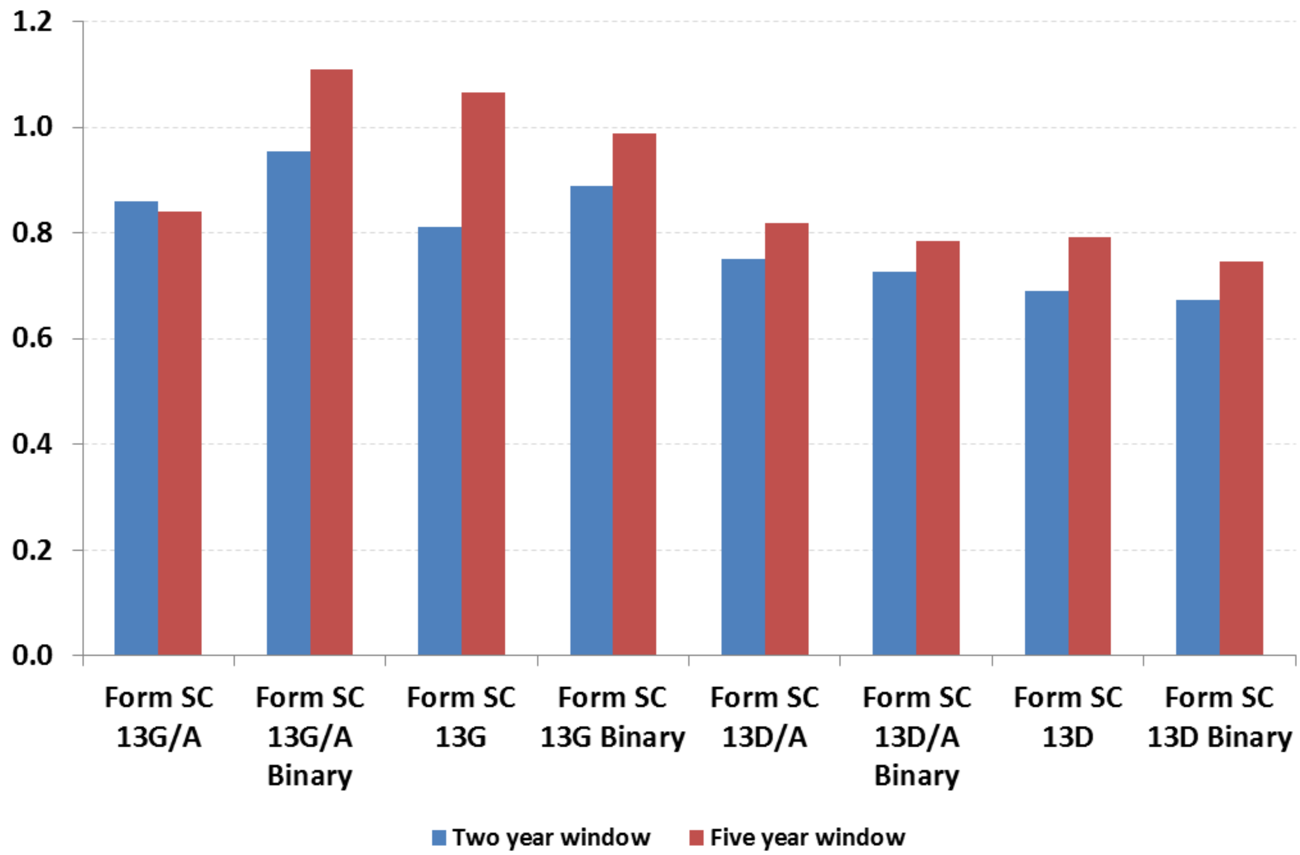
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

For each of the four forms (SC 13D, 13D/A, 13G, 13G/A), we again construct four signals:

- Two-year signal – total number of filings in the past two years, computed monthly, with an exponential decay (half-life of one year)
- Five-year signal – total number of filings in the past five years, computed monthly, with an exponential decay (half-life of 2.5 years)
- Two-year binary signal – equals to one if a company has had any filing in the past two years (zero otherwise)
- Five-year binary signal – equals to one if a company has had any filing in the past five years (zero otherwise)

Interestingly, all 16 factors based in this category show strong predictive power of future M&A activities (see Figure 24). A longer five-year window is generally preferred. Similarly, for the most popular form SC 13G/A, the binary signals have better predictive power than the frequency based ones.

Figure 24 Risk Adjusted M&A IC, SC 13D, SC 13D/A, SC 13G, SC 13G/A



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Form DEFA14A, DFAN14A, PRER14A, PREC14A

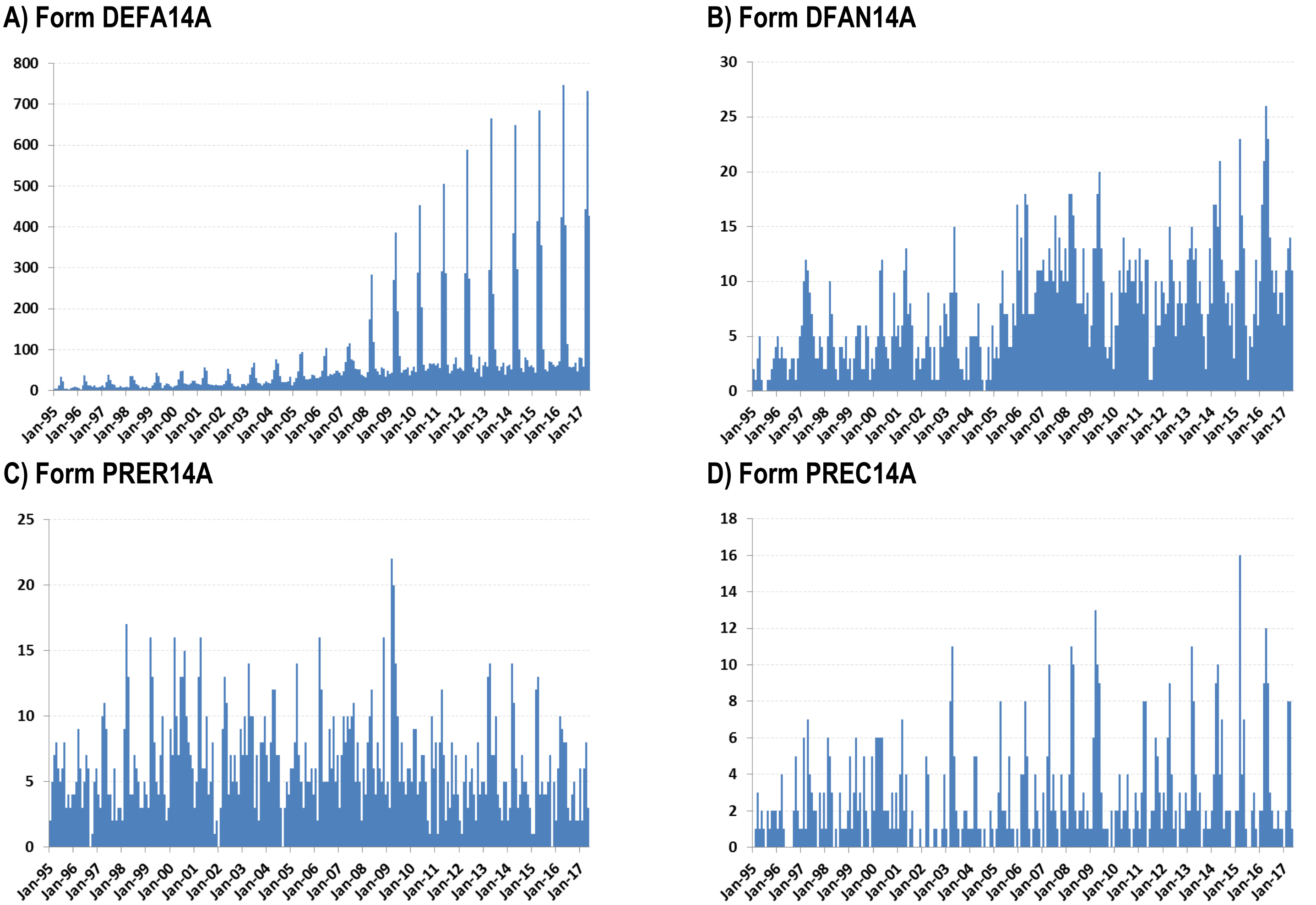
Certain filings must be produced when a company wants to provide additional materials related to an upcoming required shareholder vote. We study four types of filings in this section:

- Form DEFA14A is filed with the SEC, when additional definitive proxy materials are given to shareholders.
- DFAN14A covers additional definitive proxy soliciting materials filed by non-management.
- Form PREC14A is filed when a preliminary proxy statement is given to shareholders.
- Form PRER14A is filed when a revised preliminary proxy statement is given to shareholders.

These forms should provide security holders with sufficient information to make an informed vote at an upcoming security holders' meeting or to authorize a proxy to vote on their behalf. Key information includes the date, time and place of the meeting of security holders, revocability of proxy, dissenter's right of appraisal, persons making the solicitation, direct or indirect interest of certain persons in matters to be acted upon, modification or exchange of securities, financial statements, voting procedures; and other details.

As shown in Figure 25, these filings are less common and less frequent than the other forms covered in the previous sections. All four forms (in particular, Form DEFA14A) show strong annual seasonal patterns, coincided the annual shareholder's meeting (AGM).

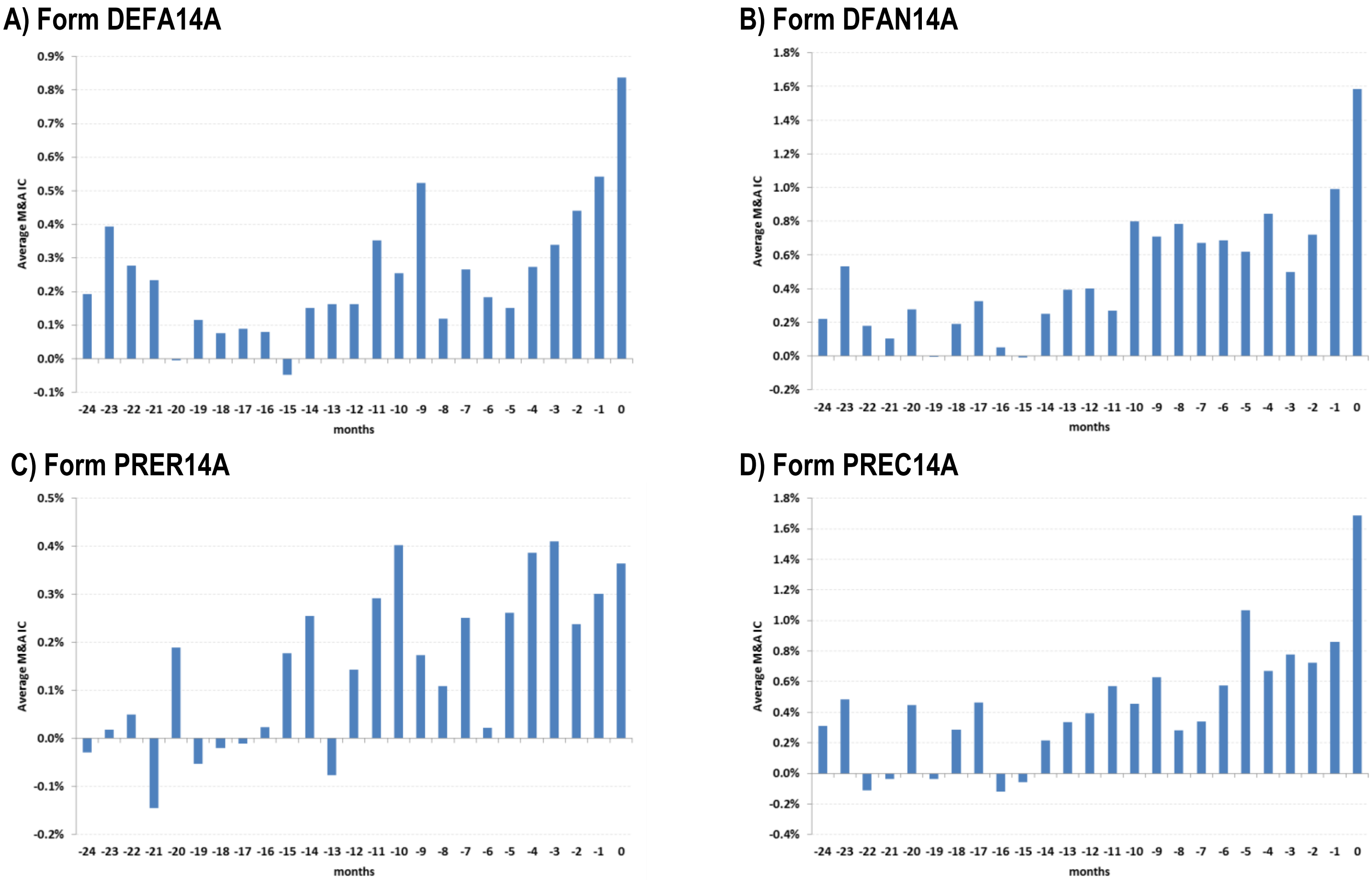
Figure 25 Number of Filings, DEFA14A, DFAN14A, PRER14A, PREC14A



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

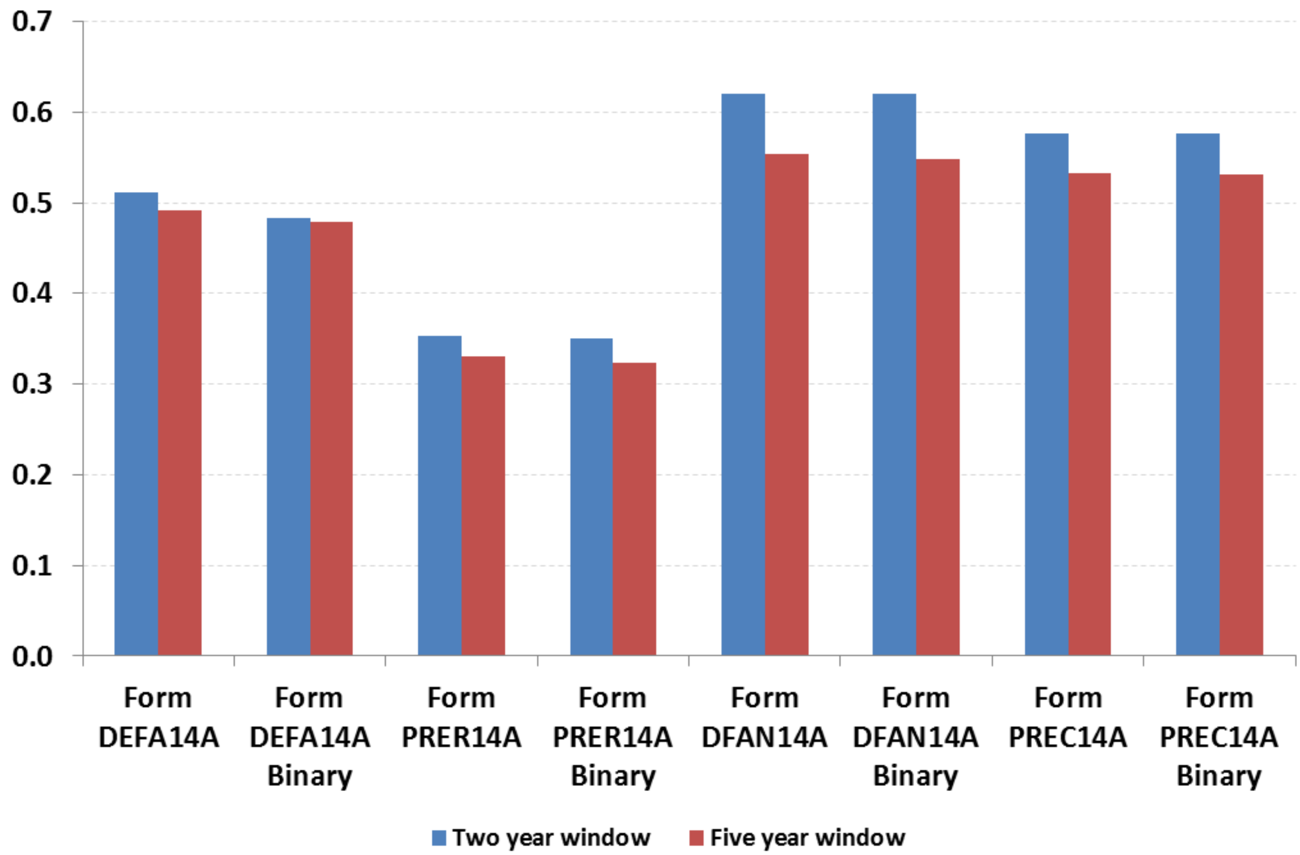
As expected, the frequency of these forms is also associated with future takeover transactions (see Figure 26). The key difference from signals based on other forms in the previous section is that the information in the proxy voting forms seems to decay faster. As shown in Figure 27, the two-year lookback window factors perform better than the five-year lookback window based.

Figure 26 Average M&A IC, # of Filings, by Month Prior to M&A Announcement



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Figure 27 Risk-Adjusted M&A IC, Proxy Voting Based Signals



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Other Relevant Filings

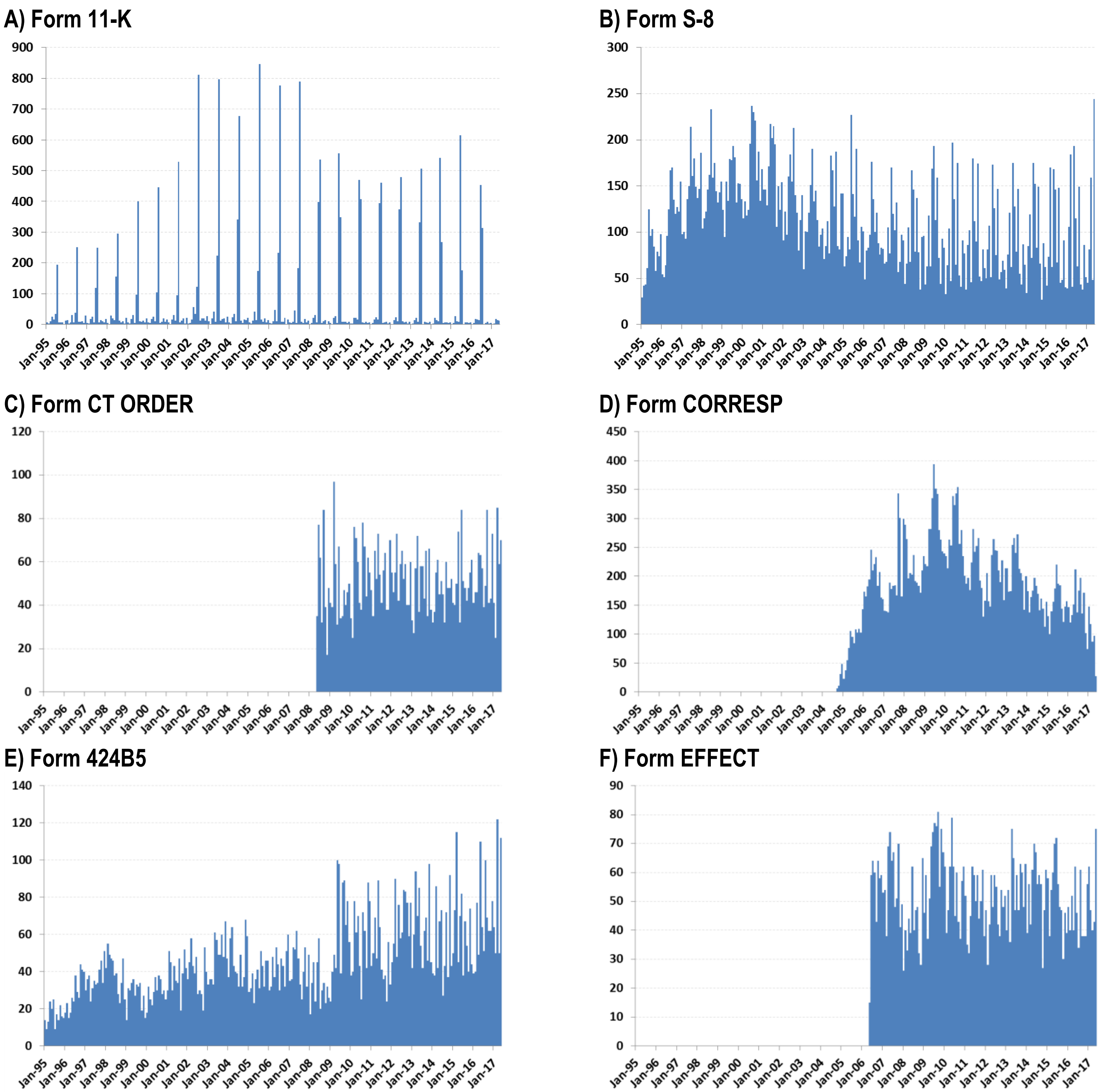
The above sections cover the most frequent and relevant SEC filings. In this section, we group together the other less frequent, but nonetheless still useful forms:

- Form 11-K is the annual report of the employee stock purchase, savings and similar plans.
- Form S-8 is the initial registration statement for securities to be offered to the employee benefit plans.
- Form CT ORDER stands for confidential treatment order. It is a form filled out in accordance to company's 8-K, 10-Q or 10-K. This form allows a company to keep confidential information as a secret for a certain period of time; otherwise such information would put the firm at a disadvantage. For example, if a company is seeking to be bought out, it might file CT ORDER to prevent the M&A related information leak to the market.
- Form CORRESP is an EDGAR form type that registrants can use to submit non-public information to SEC. The information in Form CORRESP is not disseminated immediately but may be released in part or in whole by the SEC, if the documents within the filing relate to the SEC's review process.

- Form 424B5 must be filed within two business days of the determination date of an offering price or the date first used following the effective date of a public offering or sale of securities by the company.
- Form EFFECT is the Notifications of Effectiveness. It is a public declaration by the SEC that a public company's registration statement has been accepted. For shares in the public companies to trade on the open market, they must be registered by the company and approved by SEC.

As shown in Figure 28, these filings tend to have low frequency, with shorter history. Some types of filings starts late in the 2000s (e.g., Form EFFECT), as the SEC filing requirements were only introduced at the time.

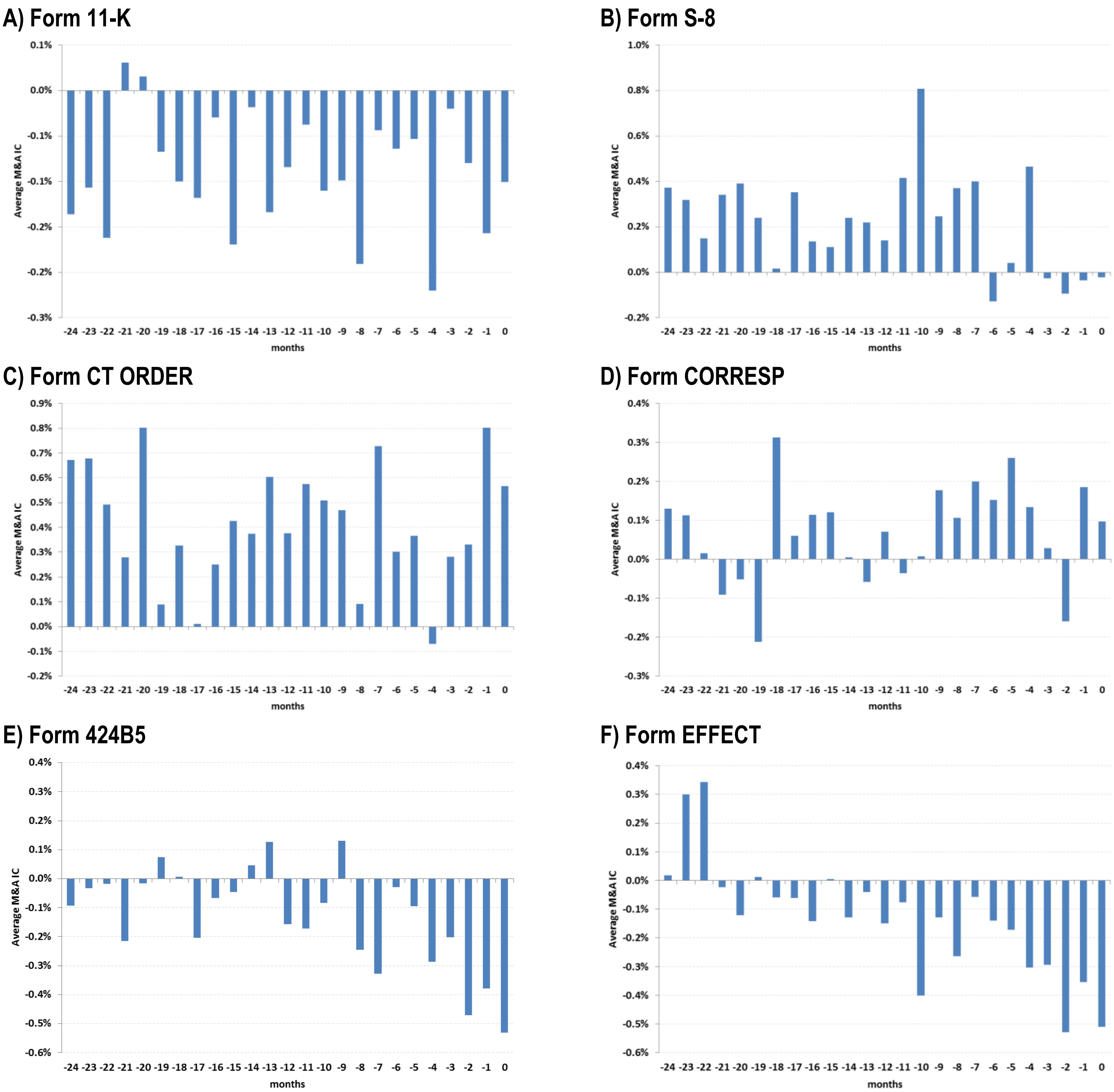
Figure 28 Number of Filings



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

As shown in Figure 29, most of these filings are predictive of future M&A transactions, with intuitive directions. For example, firms with heavy employee share purchase plans (Form 11-K) and new securities offerings (Form 424B5 and Form EFFECT) are less likely to be acquired. On the other hand, frequent filing of confidential information (e.g., Form CT ORDER and Form CORRESP) are associated with higher takeover probabilities.

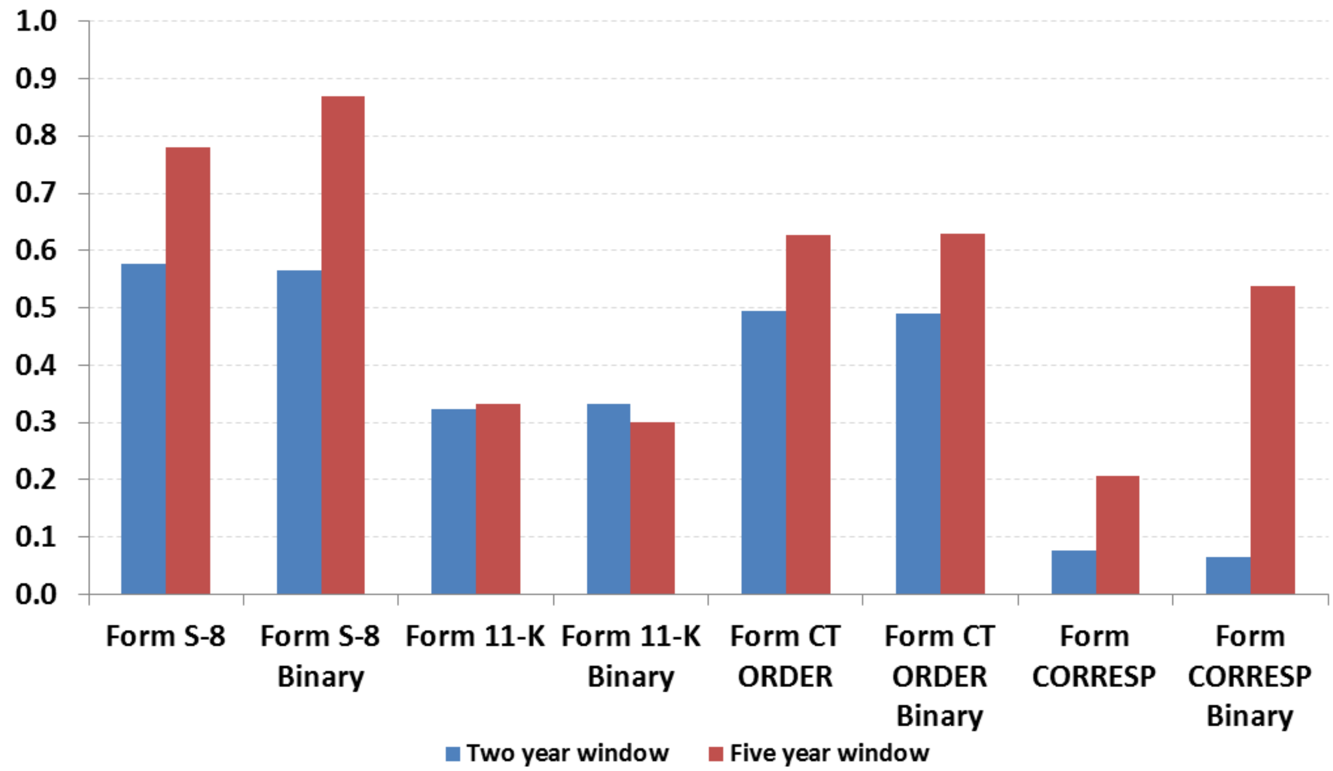
Figure 29 Average M&A IC, # of Filings, by Month Prior to M&A Announcement



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Most signals show a stronger predictive power with a longer five-year lookback window (see Figure 30). The exception lies in the two forms related to securities offerings (Form 424B5 and Form EFFECT). Because of the short timeline of these forms, a one-quarter lookback window has the best performance (see Figure 31).

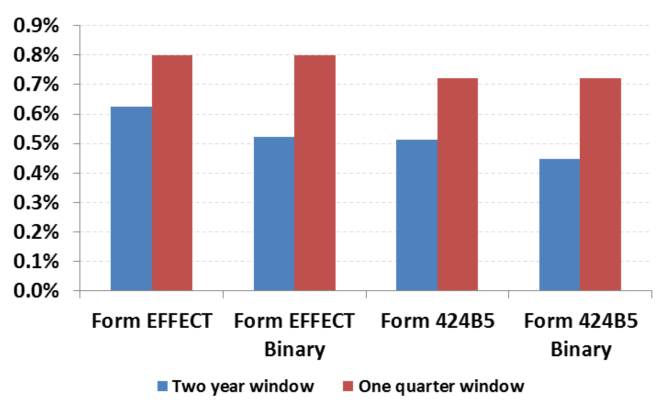
Figure 30 Performance



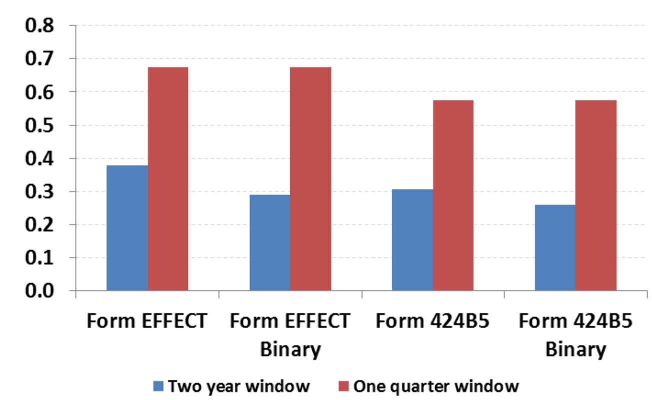
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Figure 31 Shorter Horizon for Form EFFECT and Form 428B5

A) Average M&A IC



B) Risk Adjusted M&A IC



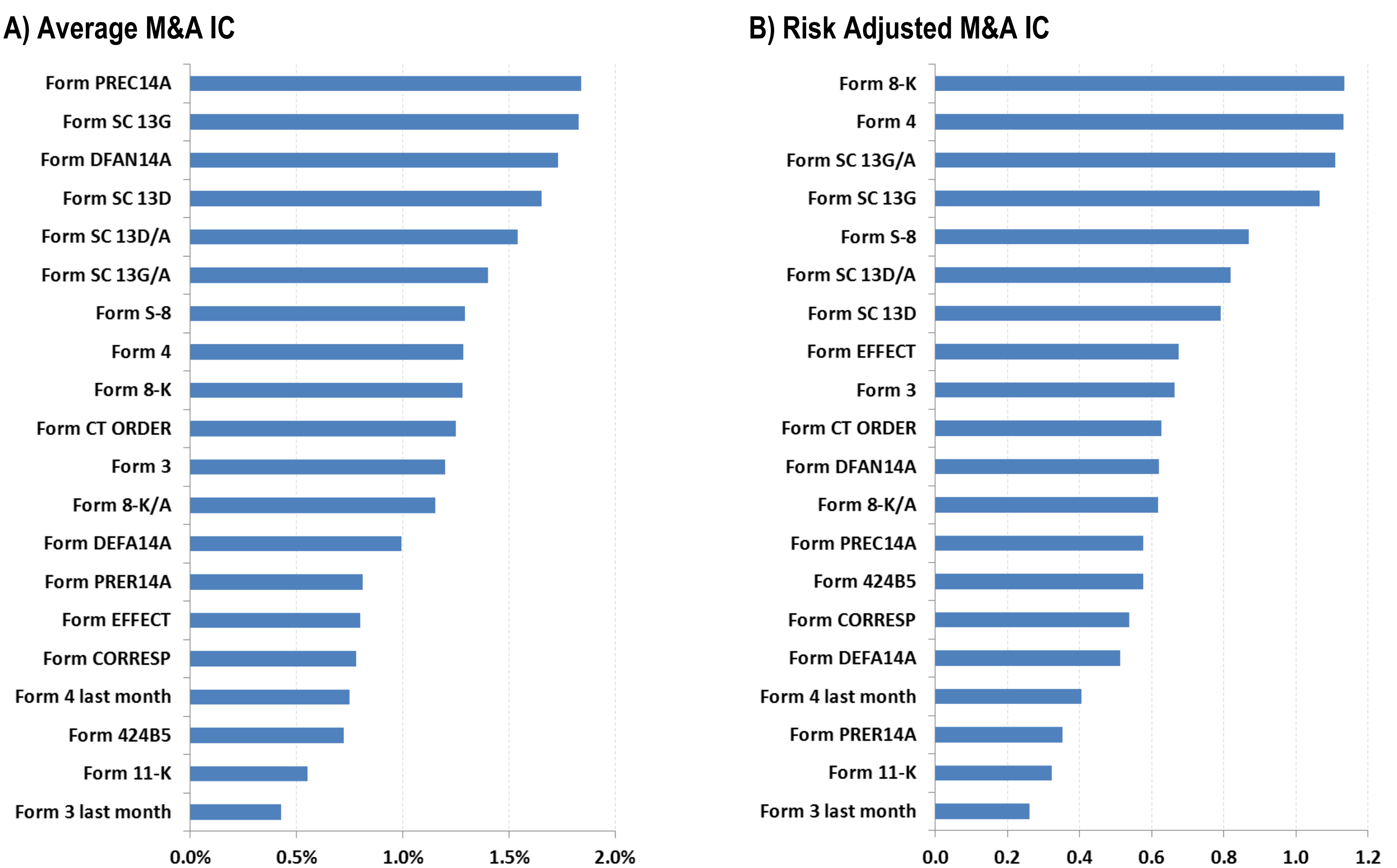
Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Performance Summary

For each of the major filing form, we pick one factor among the four possible choices (two-year or five-year lookback window, frequency-based or binary signal). For Form 3 and form 4, we choose two more short-term signals based on the number of filings in the previous month (in addition to the ones based on a five-year lookback window).

Figure 32 shows the performance of the EDGAR-based signals in predicting takeover targets. Most of them have stronger predictive power than traditional stock-selection factors (see Figure 10 to Figure 14). In some occasions, e.g., Form 8-K, Form 4, the risk-adjusted M&A IC's are almost twice as strong as the best traditional factors.

Figure 32 Performance, EDGAR Filing Factors



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

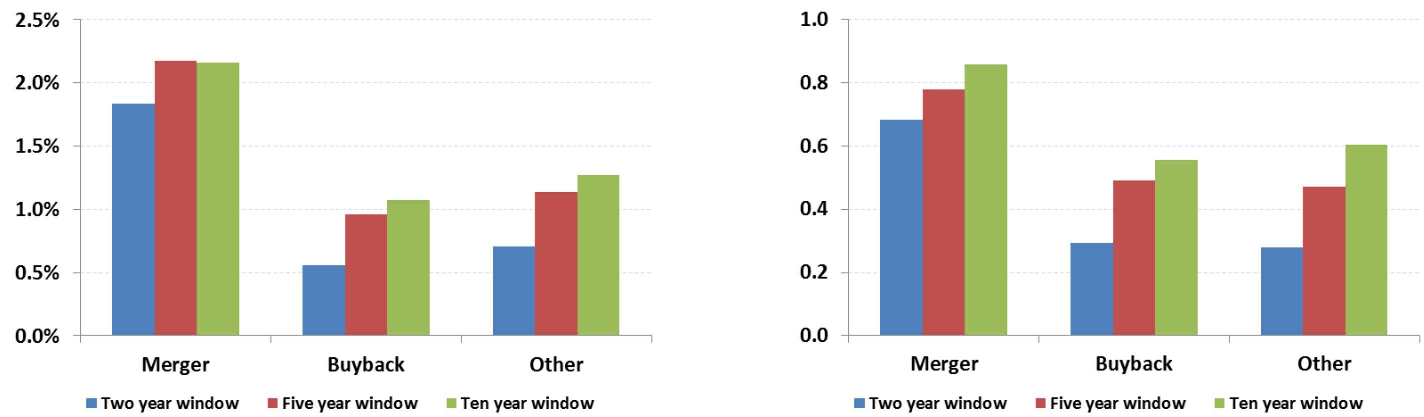
M&A Event Count Signal

Similar to the argument on counting filing frequencies in the EDGAR database, we are also motivated to examine whether the number of past M&A, buyback, and other related activities for a company makes it more or less likely to be targeted in the future. As a reminder, we group all event types in the TRSDC database into three categories: merger, buyback, and other.

Given the deep history available in the TRSDC database, for each of the three event types, we study three lookback windows: two-year, five-year, and 10-year. As shown in Figure 33, longer lookback window (e.g., five-year and 10-year) factors perform better than shorter window. Binary signals have comparable performance as frequency-based ones (see Figure 34).

Figure 33 Performance, M&A Event Count Signals

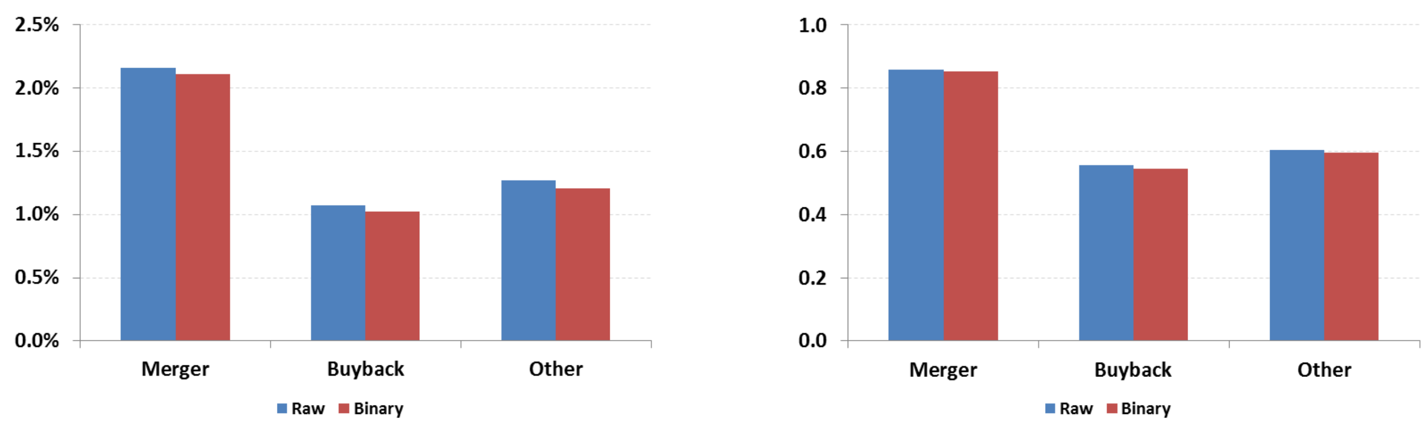
A) Average M&A IC **B) Risk Adjusted M&A IC**



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

Figure 34 Performance, Binary Signals

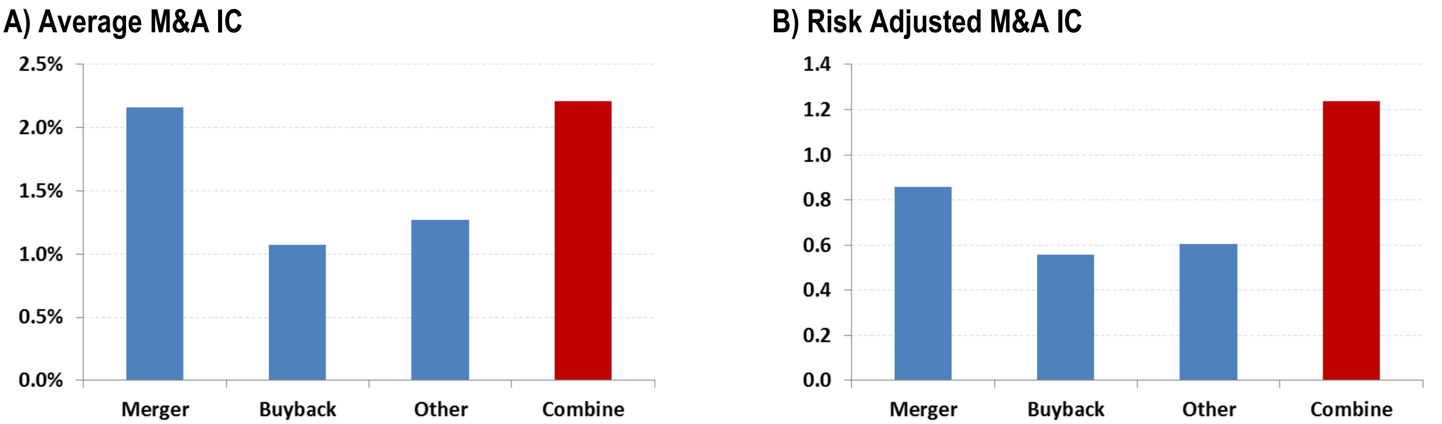
A) Average M&A IC **B) Risk Adjusted M&A IC**



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

The correlations among the signals based on the three event types are fairly low. Therefore, to take advantage of the diversification benefit, we further construct a simple equally weighted factor. As shown in Figure 35, the combined factor boosts the performance even further.

Figure 35 Combined Signal Offers Diversification Benefit



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

SECTOR M&A MOMENTUM

Anecdotally, M&A follows a strong industry cycle. If we see many takeover transactions in the commodities market, we are more likely to witness more deals in the same sector in the near future; and vice versa for other sectors. Figure 36 (A) plots the percentage of merger actions by each of the 11 GICS sectors. It is obvious that there is a strong positive serial correlation for all 11 sectors (see Figure 36 B).

As we discussed earlier, certain sectors have more M&A announcements. We calculate the sector hit rate based on the number of M&A target over the total number of stocks for each sector each year. As shown in Figure 36, the sectors have high hit rate tend to have hit rate going forward.

Now, let's introduce our sector M&A momentum factor, which is defined as:

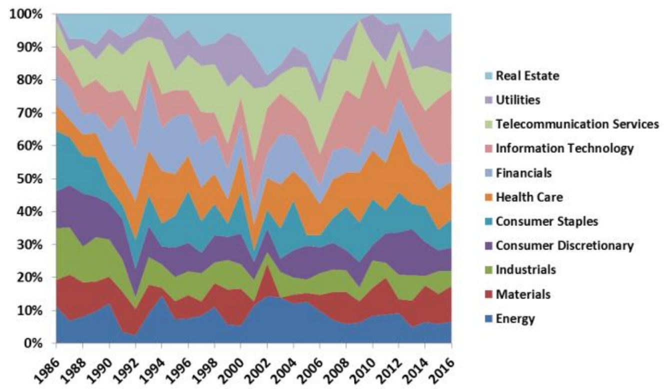
$$SectorM\&AMomentum = \sum_{t=-120}^{0} \omega_t \left(\frac{\#\ of\ M\&A\ Transactions\ in\ the\ Sector}{\#\ of\ Stocks\ in\ the\ Sector} \right)$$

Essentially, we compute the ratio of M&A transactions in the sector every month, by each of the 11 GICS sectors. Then we aggregate the ratio over the past 10 years (120 months), with an exponential decay function, with a half-life of five years.

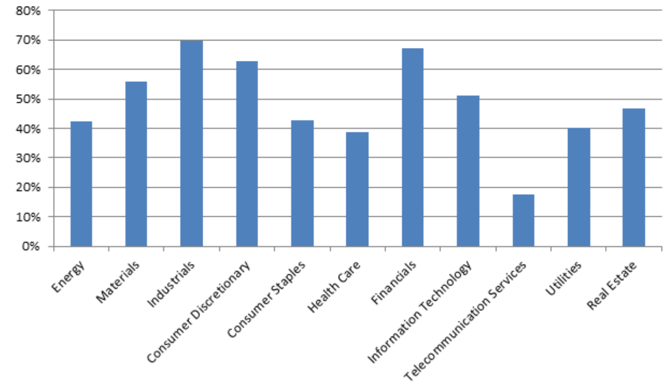
The sector M&A momentum factor is computed at the stock level, albeit the factor score is the same for every stock in that sector.

Figure 36 Percentage of M&A Transactions, by GICS Sector

A) Percentage of M&A Transactions, by GICS Sector



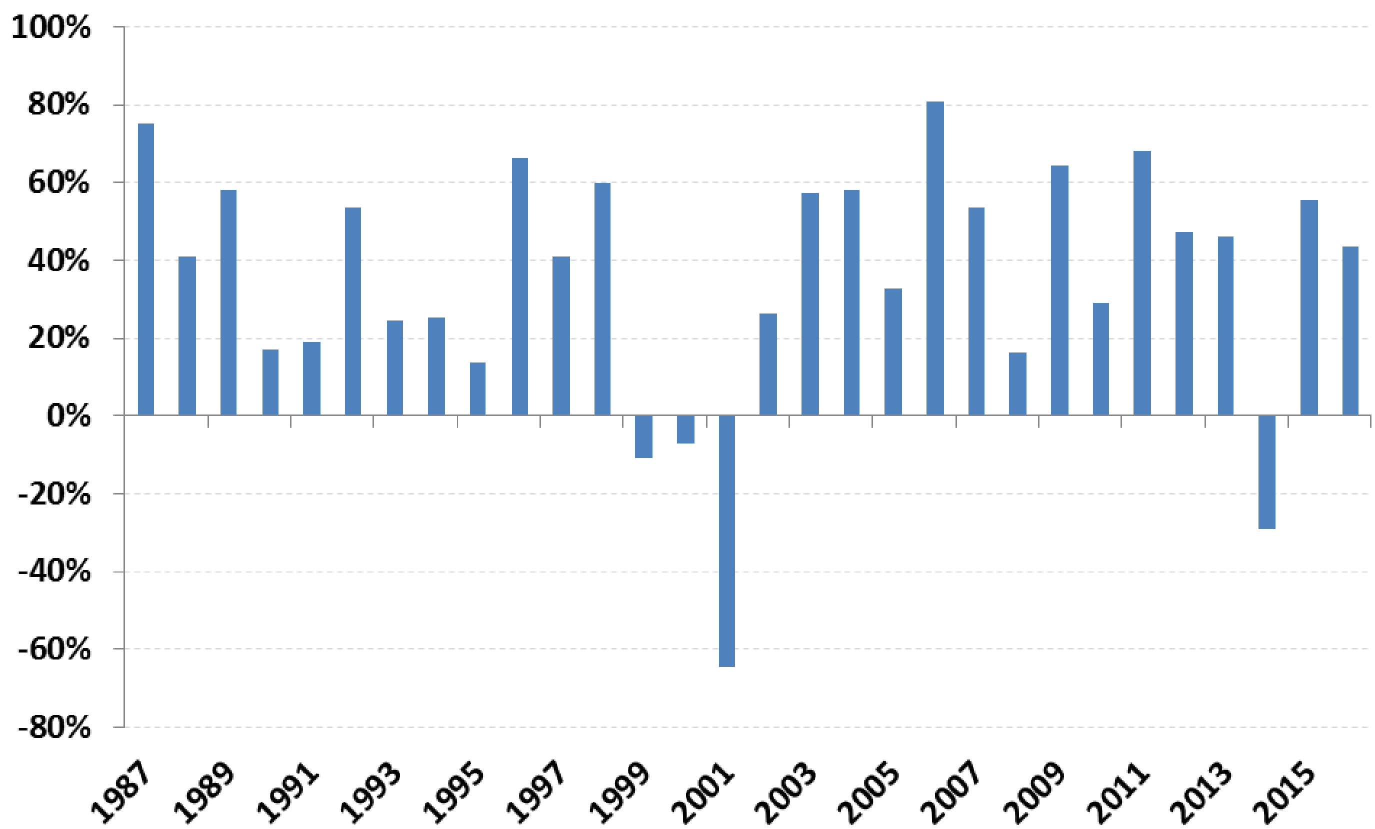
B) Autocorrelation



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

The autocorrelation of our sector M&A momentum factor is high and positive in most years (see Figure 37), suggesting that sectors with intensive past M&A transactions tend to see heavy deals in the next year as well.

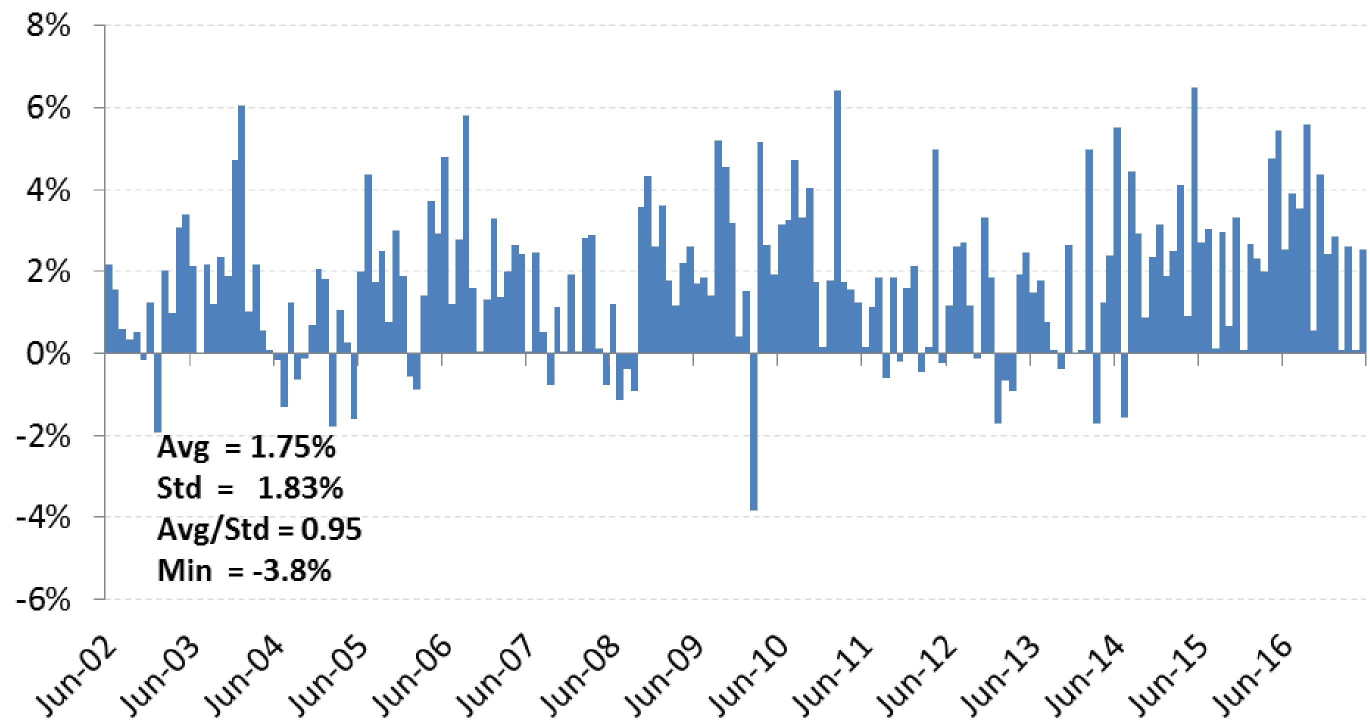
Figure 37 Autocorrelation of the Sector M&A Momentum Factor



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

As shown in Figure 38, the performance of our sector M&A momentum factor is extremely strong and stays positive most of the time.

Figure 38 Sector M&A Momentum Signal, M&A IC

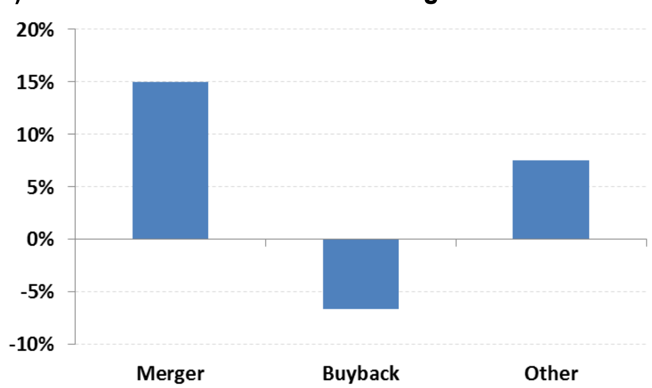


Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

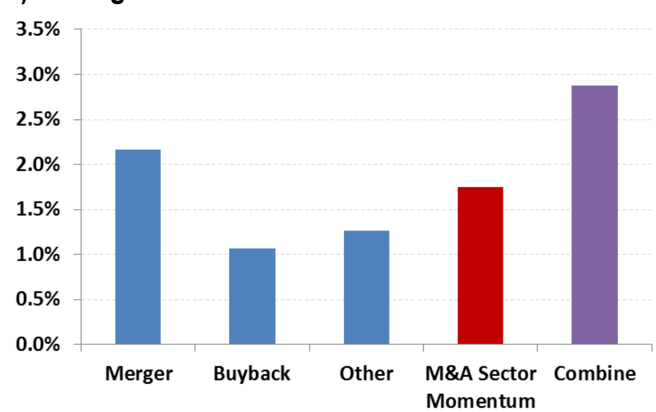
Furthermore, our sector M&A momentum factor is uncorrelated to the three M&A event count signals introduced in the previous section (see Figure 39 A). Combining the four factors together significantly improves the model performance (see Figure 39 B).

Figure 39 Uncorrelated to other Factors

A) Correlation with M&A Events Signals



B) Average M&A IC



Sources: FTSE Russell, S&P Capital IQ, Thomson Reuters, Wolfe Research Luo's QES

MACHINE LEARNING TAKEOVER

As argued in Luo, et al [2017a and 2017d], the future of active investing lies in either having access to better information and/or better ways to analyze the information than peer portfolio managers. After the introduction of alternative Big Data (e.g., EDGAR filing data) in takeover prediction, in this section, we want to see whether sophisticated machine learning algorithms can help us to identify takeover targets.

In our past research, we have shown the effectiveness of the machine learning models in different settings, such as stock selection, event prediction, and global macro forecast. Our flagship stock selection model: LEAP (see Luo, et al [2017c]) uses a modified random forest algorithm to select factors, a modified CART model for nonlinear modeling, and then a panel regression to extract linear factor premia. Our accounting fraud detection model (see Jussa, et al [2017]) employs our proprietary TS-boosting boosting model to predict companies that are likely to re-state their financial statements. In Rohal, et al [2017], we apply a penalized regression technique (LASSO) as a factor weighting tool. Lastly, in Luo, et al [2017d], we study advanced econometric models (e.g., state space Kalman filter, Markov regime switching, Bayesian VAR) along with our proprietary time-series machine learning algorithms (based on the Bayesian model average philosophy) in global macro investing.

In this section, we study and compare a wide range of machine learning algorithms in rare event prediction – identifying potential takeover targets. Our SMAP (Systematic Mergers and Acquisitions Prediction) model achieves great prediction accuracy and economic profit.

MACHINE LEARNING FRAMEWORK FOR RARE EVENT PREDICTION

Takeover prediction is a classic binary forecasting problem. At each month end, among the 3,000 stocks in the Russell 3000 universe, we want to know which companies are most likely to be acquired in the subsequent month. In academic research, logit or logistic regression[2] is the most common technique for binary prediction. In standard academic research, the number of factors considered is typically small. Factors are selected by researchers based on economic intuition.

As argued in Luo, et al [2017b, 2017c], traditional factor selection approaches based economic intuition are sensitive to in-sample variable selection bias. In this paper, rather than handpick a fixed set of factors based on past performance, we rely on machine learning feature selection algorithms to select and weigh each factor, at each given point-in-time. Therefore, our models are far more conservative and avoid look-ahead bias.

Since we have hundreds of factors including both traditional stock-selection factors and innovative new signals (e.g., EDGAR filing based, M&A event count signals, and sector M&A momentum factor), dimension reduction or factor selection is the first issue that we need to tackle. Machine learning techniques are well designed for this purpose (see Luo, et al [2017b] on machine learning feature selection).

In this section, we use all the factors covered in the previous sections and group them into three categories:

[2] Probit models are also often used. In our experience, the performance of logit and probit tends to be very similar. Therefore, in this paper, we will only use logit model to represent traditional regression techniques.

- Traditional stock-selection factors, primarily accounting ratios, valuation multiples, and technical indicators
- Non-traditional EDGAR filing and M&A event count factors
- Sector M&A momentum signal

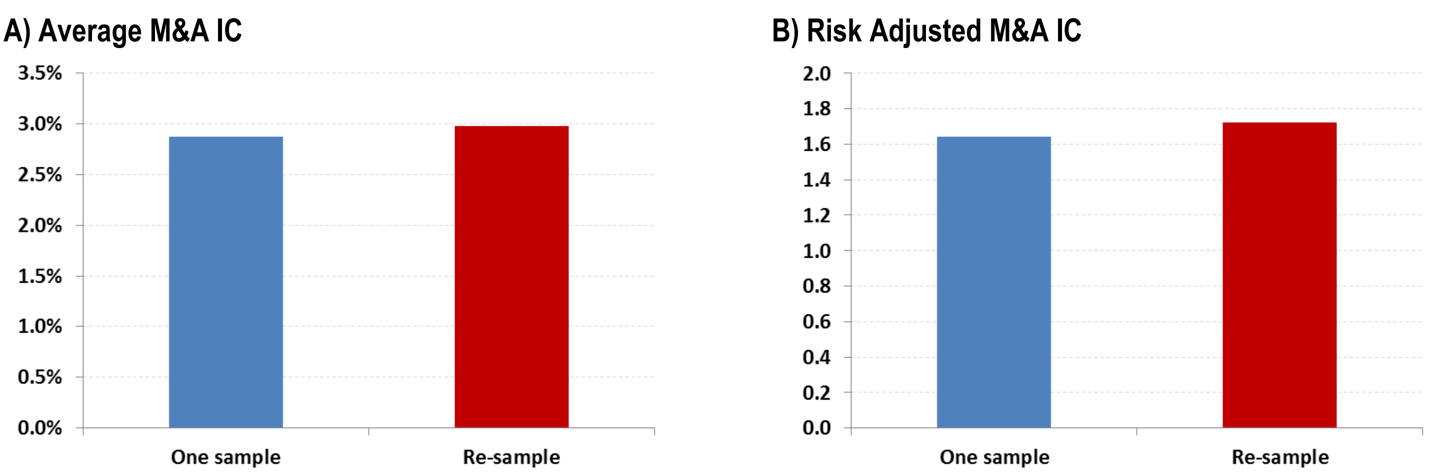
Because M&A's are rare events, at any given point-in-time, typically less than 0.5% of companies (about 15) are takeover targets, while 95% firms are in the comparable control universe. Most classification algorithms can't effectively handle such highly imbalanced samples. Therefore, following convention (see Jussa, et al [2017]), we also use a randomly selected match-sample for our study.

Our machine learning framework uses a 10-year rolling window of monthly data to train the model. At each month end, we normalize all factors[3], match each target with a randomly selection comparable stock, and stack together all observations in the past 10 years as our training sample.

Following Breiman's [1996] bagging (bootstrap aggregating) philosophy, at each point-in-time, we repeat the same random sampling procedure above 10 times. The average predictions from the 10-fold bagging forms the basis of our final forecast.

As shown in Figure 40, the re-sample bagging procedure boosts the performance by 5% for a standard logit model (using only non-traditional factors) in our takeover application. The bagging algorithm is added to all machine learning models in the paper.

Figure 40 Performance Improvement from Bagging (Logit Model with Non-Traditional Factors)



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

SECTOR NEUTRALIZATION

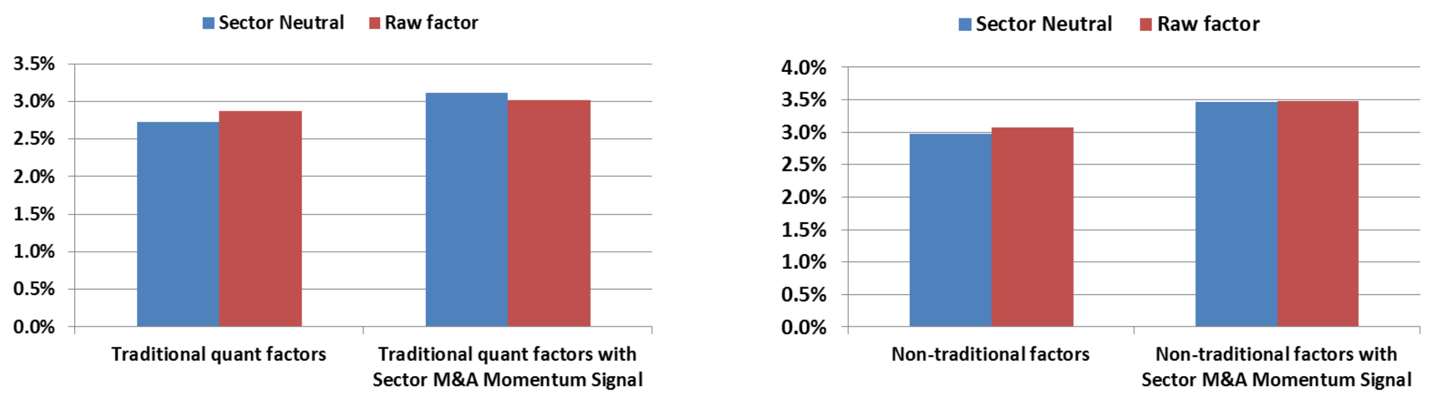
As argued in Luo, et al [2017b], most common factors are better suited for stock selection within a sector rather than sector timing. Therefore, neutralizing or normalizing each factor within sectors generally improves a model's risk-adjusted performance.

[3] See Luo, et al [2017b] for our proprietary data normalization techniques

As shown in Figure 41 and Figure 42, adding our sector M&A momentum factor boosts the performance of both traditional and non-traditional factors. In addition, sector neutralization modestly improves the performance of traditional factors, especially combined with our sector M&A momentum factor. For non-traditional factors, sector neutralization makes no difference. Therefore, for the rest of the research, we only sector neutralize all traditional factors.

Figure 41 Average M&A IC, Logit Model

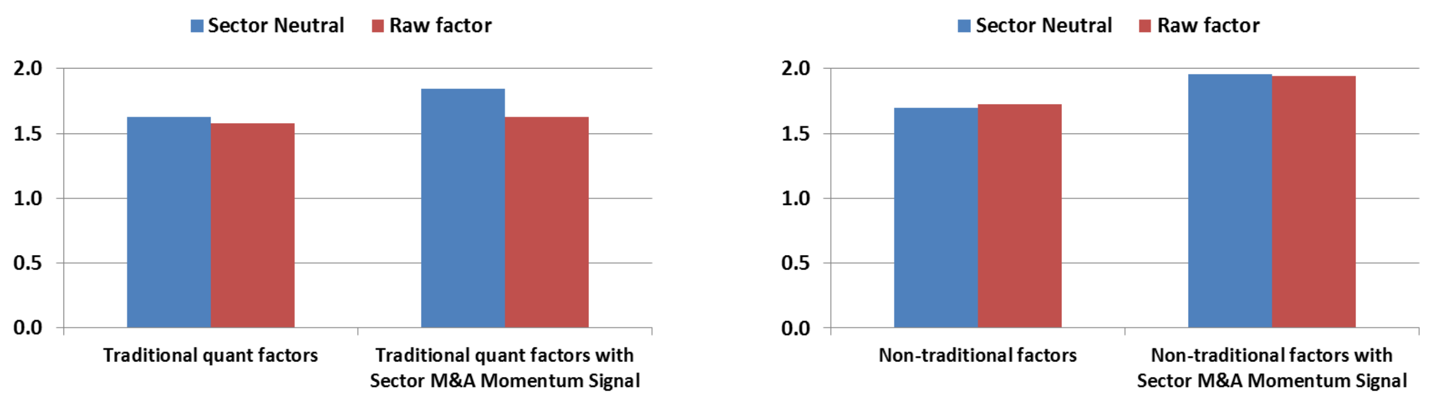
A) Traditional Quant Factors **B) Non-traditional Factors**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Figure 42 Risk Adjusted M&A IC, Logit Model

A) Traditional Quant Factors **B) Non-traditional Factors**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

A Primer on Machine Learning Algorithms

In this section, we provide a brief primer for the six machine learning algorithms to be used for takeover prediction:

- Logit and logistic regression
- CART

- Random forest
- SVM
- ANN
- Adaboost

Logit

Logit model (also known as logistic regression) is one the most traditional classification algorithms. It was developed alongside and is very similar to linear regression. Because the dependent variable in a binary (zero and one) instead of a continuous variable, we need to make a simple transformation.

For binary classification problems, there are two key concepts – odds and odds ratio. Odds represent the ratio of the probability of an event over the probability of non-event. One way to quantify the predictive ability of a binary predictor is the odds ratio – the ratio of the odds when the predictor takes on one value versus the odds when the predictor takes on the other value. If p is the probability of an event, the odds of the event are then $p/(1-p)$. Logistic regression models the log odds of the event as a linear function:

$$log\left(\frac{p}{1-p}\right) = \beta_0 + \sum_{k=1}^{K} \beta_k f_k$$

Linear regression and logistic regression both belong to the broad generalized linear model (GLM) family (see McClullagh and Nelder [1989] for an introduction). Similar to linear regression, we can conduct formal statistical inference and the models are generally easy to interpret.

CART Model

CART (Classification and Regression Tree) is among the most commonly used classification algorithms. We have introduced and applied CART model in our LEAP global stock selection model (see Luo, et al [2017]).

One major problem with the CART model is that it may suffer from overfitting. In practice, we typically prune the trees using cross-validation to reduce the size of the model. A typical pruned tree only has a few splits, each split corresponding to a factor. CART model has the advantage of transparency as well as computational speed. A typical CART model can be interpreted fairly intuitively. The downside of the CART model is that prediction accuracy is generally weak, compared to other more sophisticated machine learning algorithms.

Random Forest

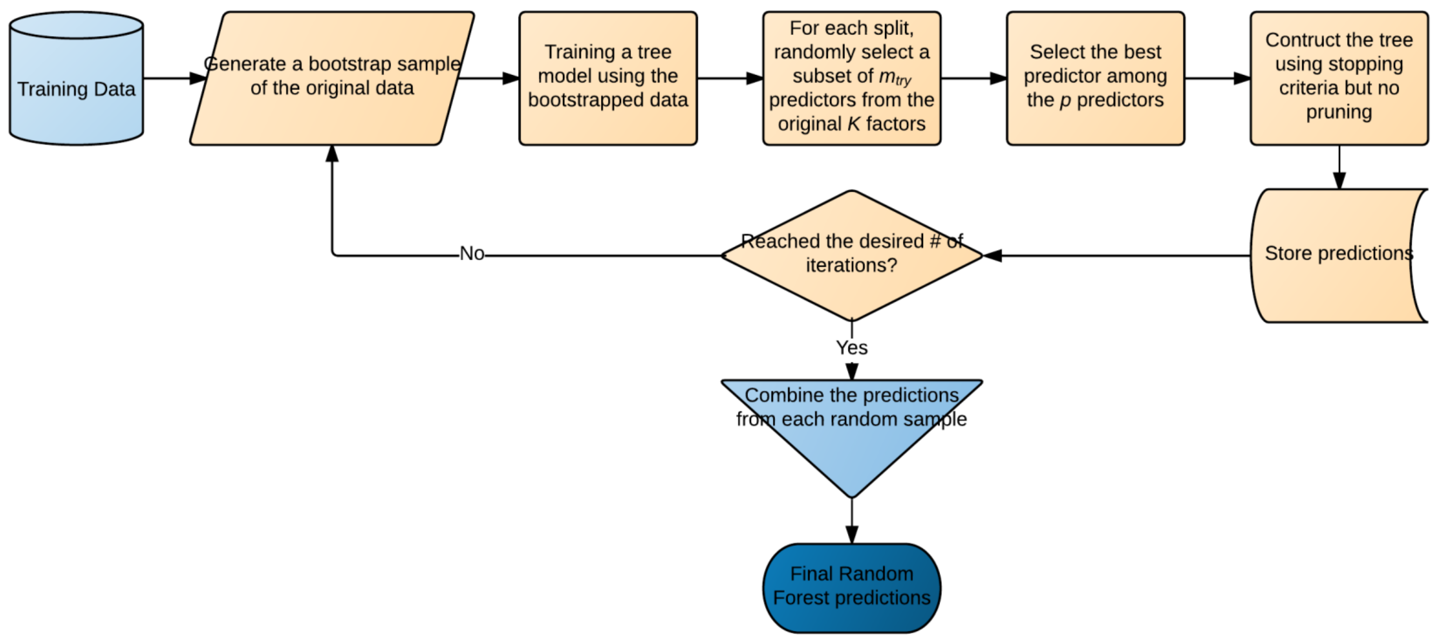
Random forest is a natural extension of the CART model. The random forecast model adds to important dimensions to the CART algorithm – randomly selecting the training data and randomly selecting factors, typically thousands of times. The average prediction from the thousands of trees forms the foundation of the random forest model. In our LEAP global stock selection model (see Luo, et at [2017]), we use a modified version of the random forest model as a feature selection tool.

Figure 43 shows the random forest model flowchart. In summary, the random forest algorithm uses bootstrap to select a sub-sample of data and a subset of factors, fit a CART model[4] with a non-

[4] Although CART is commonly used, it can be other machine learning algorithm too.

pruned tree. Then it repeats the same procedure many times[5]. Finally, we take the average from each CART model's prediction to derive our final forecast. The performance of the random forest model is typically much stronger than the CART algorithm.

Figure 43 Random Forest Flowchart



Sources: Wolfe Research Luo's QES

SVM

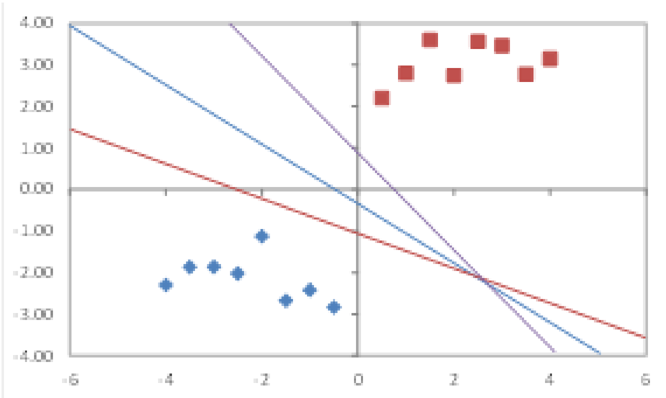
Support Vector Machines (SVMs) were first developed in the 1960s by Vladimir Vapnik (see Vapnik [2010] for a more recent and comprehensive treatment). The SVM model has gained great popularity in a number of fields, including finance (see http://www.svms.org/finance/ for a long list of papers applying SVMs in finance).

The intuition of SVMs can be shown from a simple two-class and one predictor example. In this example, we assume the two classes can be completely separated by a straight line, as shown in Figure 44(a). The problem is that there are infinite numbers of linear boundaries that can perfectly classify the data. The SVM starts from a *margin* concept. The margin is the distance from the classification boundary to the closest data point (see Figure 44(b)). The dashed lines on both sides of the boundary are at the maximum distance from the line to the closest training data points (equal distance from the boundary line).

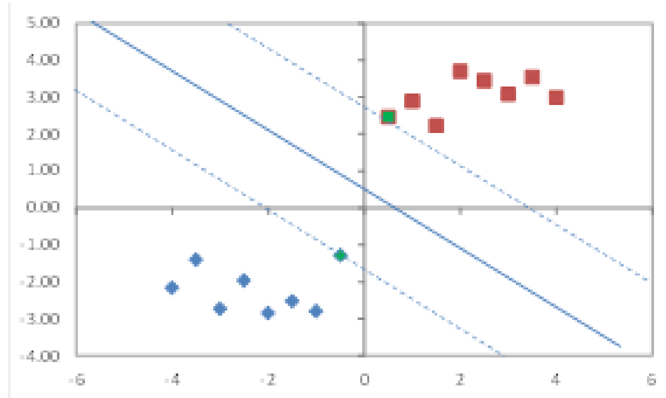
[5] In practice, many researchers fit 1,000 trees, as a standard rule of thumb. However, the model performance is typically not very sensitive to the number of trees.

Figure 44 Example of SVM

A) The growth of # of papers on SSRN



B) Seasonality, # of papers posted by month



Sources: Wolfe Research Luo's QES

In this simple two-class example, we use a linear function for the boundary line:

$$D(\boldsymbol{u}) = \beta_0 + \beta'\boldsymbol{u} = \beta_0 + \sum_{k=1}^{K} \beta_k u_k = \beta_0 + \sum_{i=1}^{N} y_i \alpha_i \boldsymbol{x}'_i \boldsymbol{u}$$

Where, if $D(\boldsymbol{u}) > 0$, we would predict a sample to be in one class and otherwise in the other class, for an unknown sample $\boldsymbol{u}$, the vector $\boldsymbol{x_i}$ contain the predictor data for a training set sample, and $\alpha_i \geq 0$. It can be shown that $\alpha_i = 0$ for all samples that are not on the margin. Therefore, the predictor equation is a function of only a subset of the training samples that are closest to the boundary, i.e., the prediction equation is supported solely by these data points; therefore, the maximum margin classifier is called the support vector machine.

In practice, the classes are almost never completely separable. The functional form does not have to be linear. We typically use a kernel function:

$$D(\boldsymbol{u}) = \beta_0 + \sum_{i=1}^{N} y_i \alpha_i K(\boldsymbol{x}_i, \boldsymbol{u})$$

Where, $K(.,.)$ is a kernel function. Commonly used kernel functions include polynomial (linear, quadratic), radial basis, and hyperbolic tangent, etc.

ANN (Artificial Neural Networks)

Neural networks (see Bishop [1995] and Ripley [1996]) were once popular in finance, but then lost their appeals, mostly due to their lack of interpretability. Neural networks are often labelled as "black-box", because there is often not enough intuition about the hidden variables (or hidden units). On the surface, neutral networks are very similar to PLS (Partial Least Squares), but they are not estimated in a hierarchical manner. The linear combination of input variables is typically transformed by a nonlinear function, such as the logistic (i.e., sigmoidal) function:

$$h_j = g\left(\beta_{0j} + \sum_{k=1}^{K} f_k \beta_{kj}\right)$$

Where, $g(u) = \frac{1}{1+e^{-u}}$, β_{kj} is the coefficient of the kth factor on the jth hidden unit.

The number of hidden units J is a tuning parameter. Once J is defined, each unit must be related back to the output, via a linear combination:

$$f(x) = \gamma_0 + \sum_{j=1}^{J} \gamma_j h_j$$

The NNs, there are often many coefficients to be estimated of $J \times (K+1) + J + 1$. Therefore, if we have 10 factors and use 5 hidden units, we need to estimate $5 \times (10+1) + 5 + 1 = 61$.

In recent years, deep learning (sometimes perceived as the new generation of NNs) has been popularity in many fields of science and engineering. One popular form of deep learning is called deep neural networks (DNNs). Compared to the simple NNs above, DNNs have more than one hidden layer and can proxy very complex nonlinear relationship, but of course, potentially even more likely to suffer from overfitting.

AdaBoost

AdaBoost is an effective and efficient machine learning algorithm for classification. It was first introduced by Schapire [1998], it has been widely used for computer vision problems such as face detection (see Wu, et al [2004]) and human pose estimation (see Wang, et al [2010]), and has been used in stock return prediction (see Creamer, et al [2006, 2010]) in recent years. We have applied the AdaBoost algorithm extensively in our stock-selection models since 2012.

The main idea behind the AdaBoost is that it adaptively builds a sequence of weak classifiers to construct the strong classifier that is working well across different months. In each step of the training, it emphasizes on the misclassified stocks thereby gradually improving the accuracy of classification. Although, each classifier can be weak, as long as their performance is not random, the performance of final strong classifier improves.

In the case of stock level classification, each weak classifier is typically a tree stump defined by a factor. It can be binary or multi-split tree stump, in our implementation; we prefer the ladder because that can capture better for the non-linear behavior of the factors.

The main advantage of AdaBoost algorithm is that it's robust, relatively transparent, and very efficient for both training and predicting.

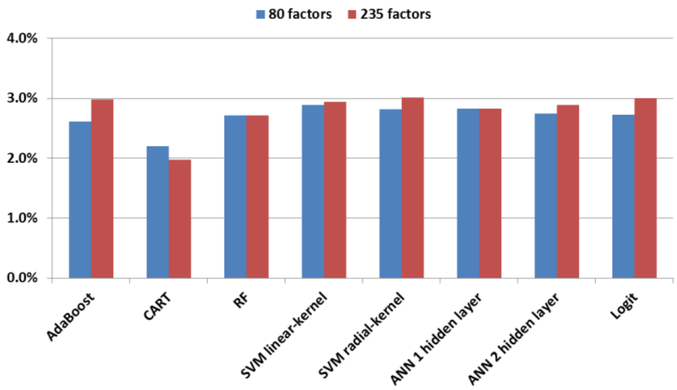
A Horserace of Machine Learning Algorithms in Rare Event Prediction

Before we compare the performance of the six machine learning models, we need to decide how many factors to feed into each algorithm. Some models (e.g., random forest, SVM, ANN, and AdaBoost) automatically conduct feature selection, while traditional algorithms such as logit do not handle large dimensionality problem very well. Therefore, for the logit model, we first apply a feature selection procedure. Only the top ranked factors are used by the logit model in fitting and forecasting.

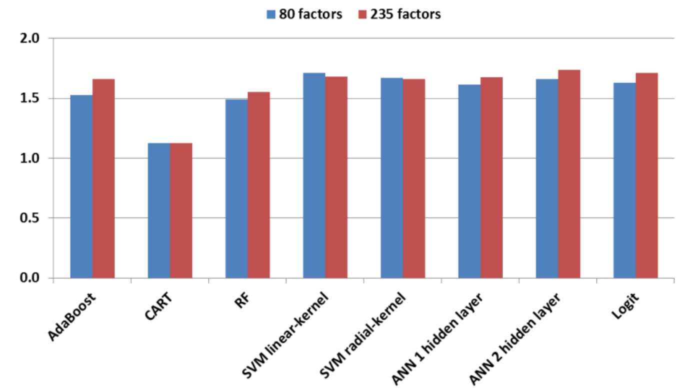
Once feature selection is added to the logit model, the performance rises as we add more factors to the process. Most of the other machine learning algorithms (other than CART) are also fairly robust to dimensionality. As shown in Figure 46, as we move from an 80-factor model to a 235-factor model, predictive power improves. The 80 factors are a subset of the 235 signals. These 235 factors form the basis of our "traditional" factor library in this research.

Figure 45 Increasing the Number of Traditional Quant Factor

A) Average M&A IC



B) Risk adjusted M&A IC

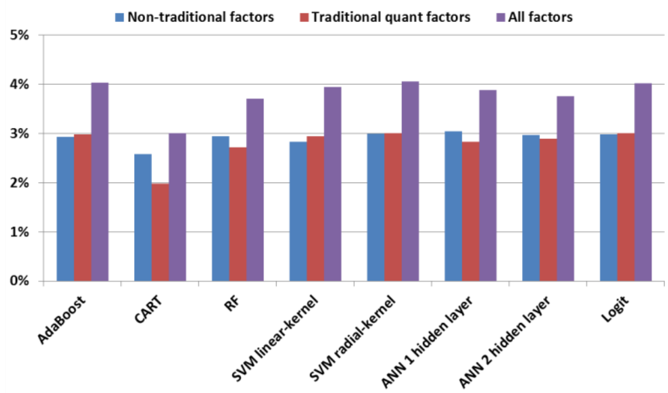


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

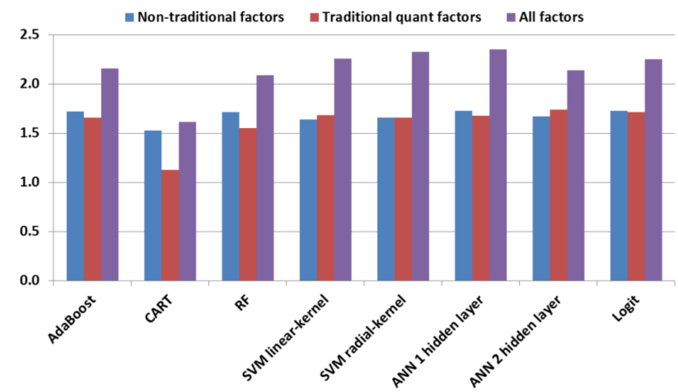
To demonstrate the potential value add from alternative Big Data, we add the other two categories of non-traditional factors (i.e., EDGAR & event count signals and our sector M&A momentum factor) to each of the machine learning models. As shown in Figure 46, interesting, the performance for models based on only traditional or only non-traditional factors is similar. However, due to the low correlation between tradition and non-tradition factors, adding both sets to the models boosts performance substantially.

Figure 46 Combining Traditional Factors with Non-traditional Factors

A) Average M&A IC



B) Risk Adjusted M&A IC

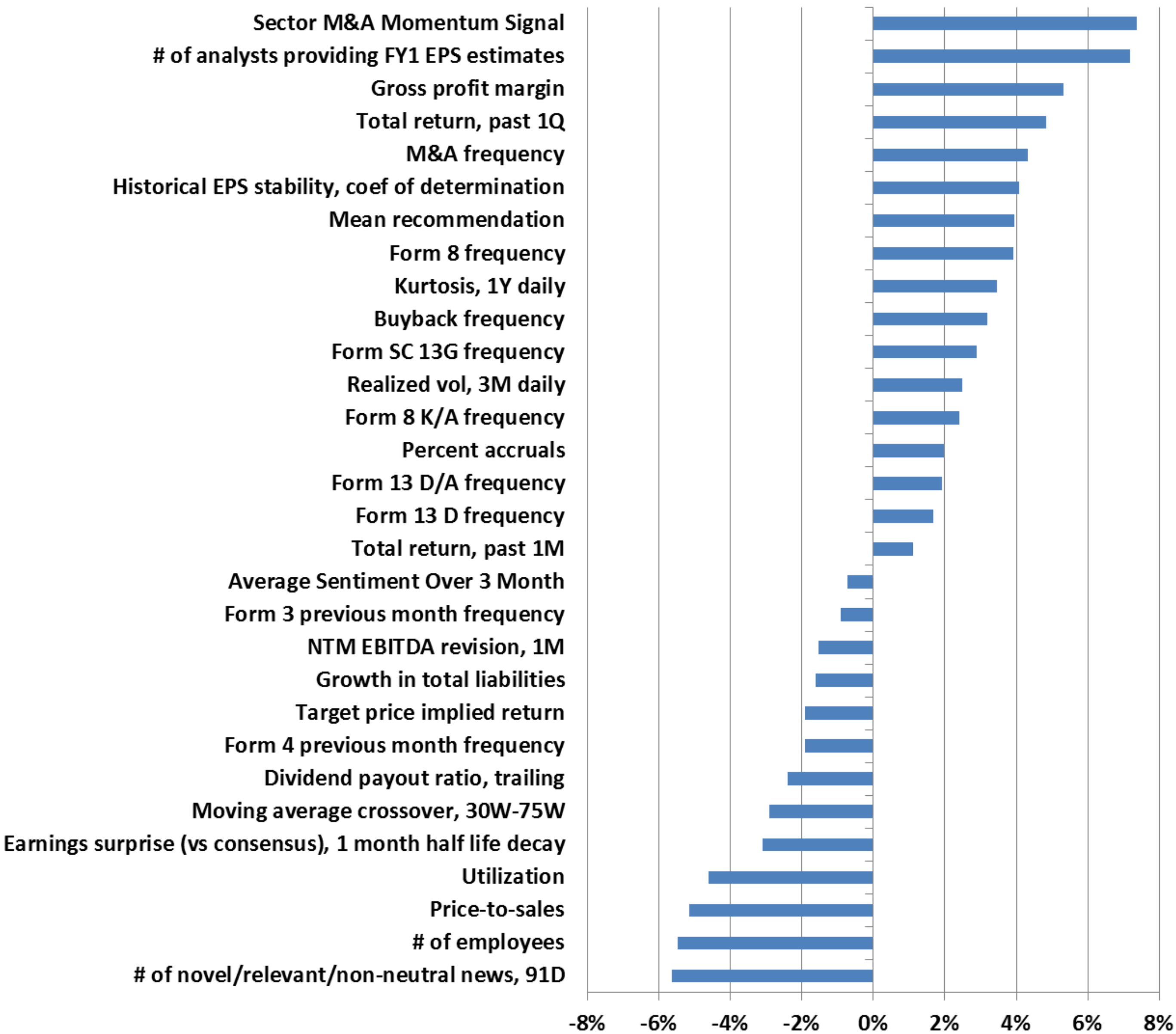


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Factors Selected by the Model

Among the six machine learning algorithms, the logit model is the most well-known and easiest to interpret. To help our readers better understand the intuitive behind our forecast, Figure 47 shows the current factors and weights in the logit model, as of the end of May 2017. There is a nice mix of both traditional and non-traditional factors been selected by the model. Most factors show intuitive signs. For example, our sector M&A momentum has the highest positive weight, followed by the # of analysts covering the company.

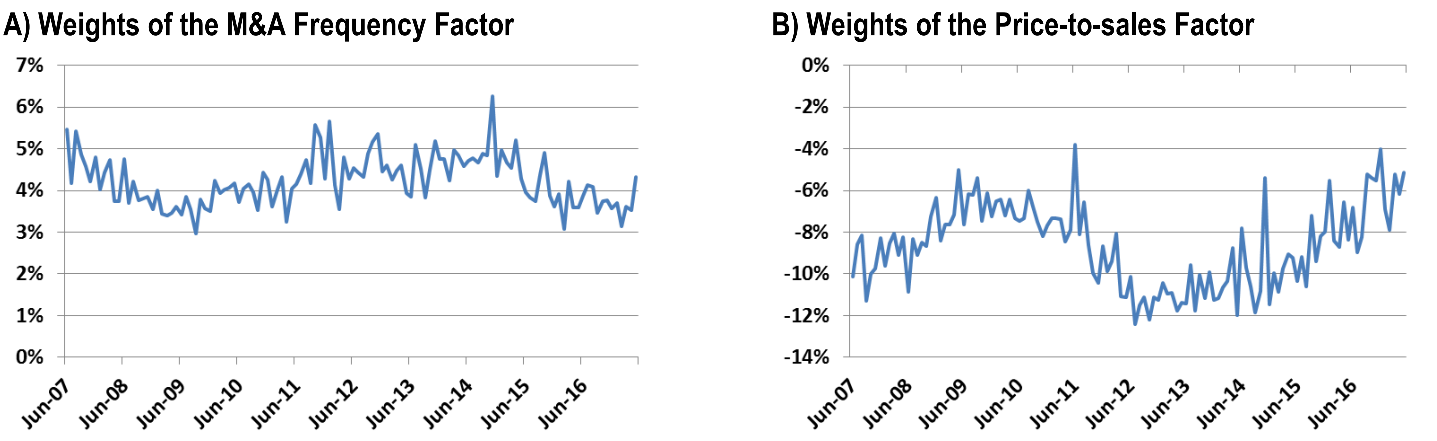
Figure 47 Factor Weights, Logit Model at May 31, 2017



Sources: Bloomberg Finance LLP, FTSE Russell, Ravenpack, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Because we use a 10-year rolling window to train the model, the weights of most factors are quite stable over time. For example, as shown in Figure 48 (A) and (B), the weight for the sector M&A momentum factor has been consistently positive, around 3% to 5%, while the weight for the price-to-sales factor has always been negative at -4% to -12% range.

Figure 48 Dynamic Weighting



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

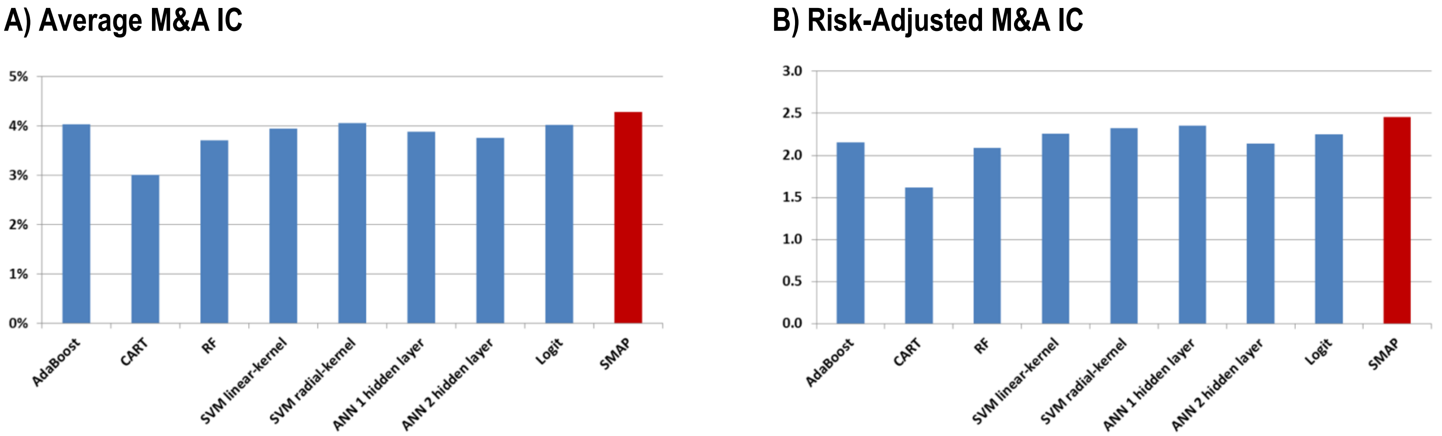
Final M&A Prediction Model

The six machine learning algorithms (and eight models) are only modestly correlation. Clement [1989] finds that combining multiple forecasts (i.e., forecast combination) generally beats each individual prediction. Clement [1989] further states that the simple averaging is an exceedingly robust procedure that often outperforms more complicated weighting schemes.

In this section, we introduce our final SMAP (Systematic Merger & Acquisition Prediction) model, which equally weighs the prediction from seven of the eight models above. We exclude the CART model, because our random forecast model encompasses the CART model and offers significantly better performance.

As shown in Figure 49 (A) and (B), the SMAP offers considerably higher M&A IC and better risk-adjusted accuracy than any of the underlying models in the past 15 years of out-of-sample backtesting. In the rest of this paper, we will analyze the SMAP model in more details and show potential applications.

Figure 49 Performance, Takeover Prediction Model



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

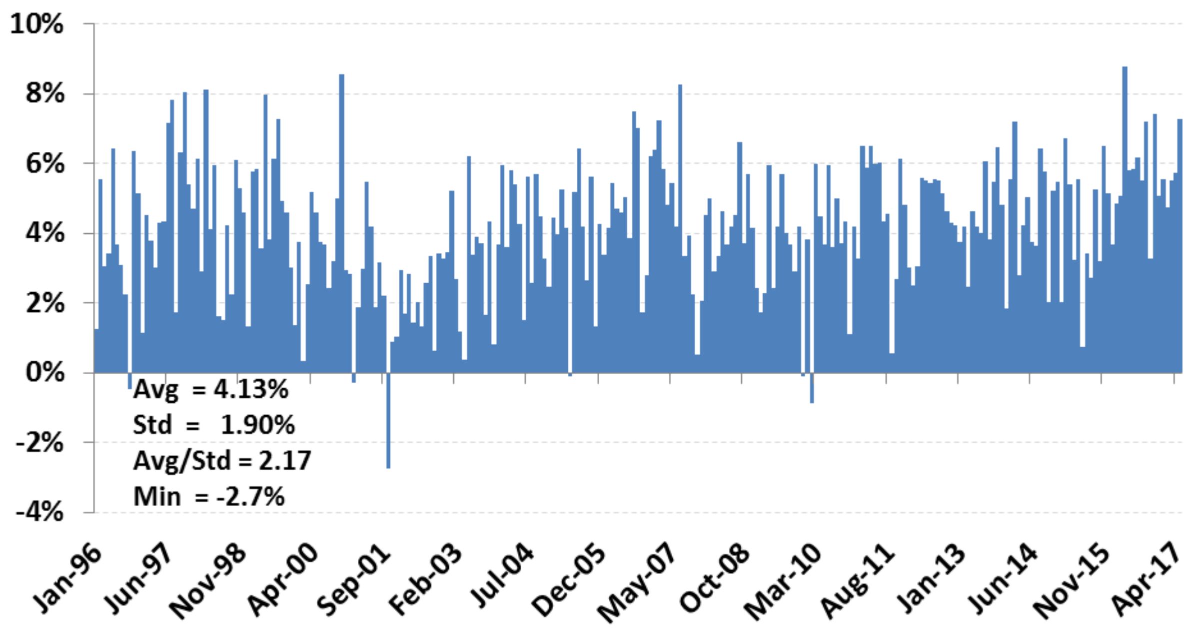
ASSESSING THE MODEL

So far, we have been focusing almost exclusively using our M&A IC framework. In this section, we evaluate our SMAP model with other more conventional accuracy measures and show other model properties.

PREDICTIVE POWER

Most of our common factor data starts from 1986. With a 10-year rolling window to train our models, we pure out-of-sample backtesting starts from 1996. As shown in Figure 50, the SMAP model's M&A IC stays positive most of the time.

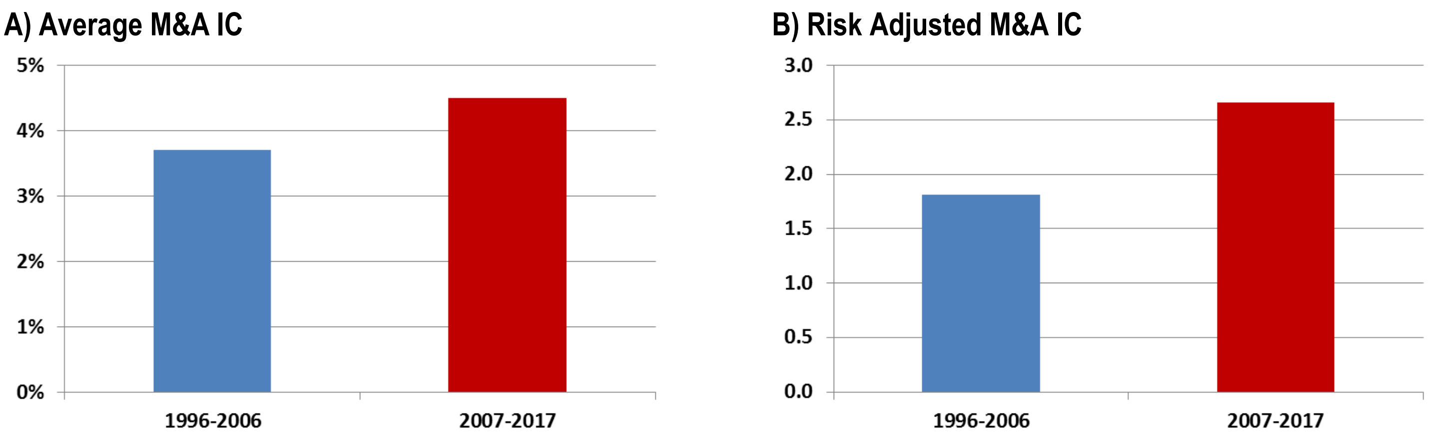
Figure 50 M&A IC, SMAP Model



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

As detailed in Luo, et al [2017a], the performance of most common factors decayed significantly in recent years. To further assess the time variation and decay, we break down our 20-year out-of-sample period into two 10-year sub-periods: 1996-2006 and 2007-2017, Unlike most other common factors and models, the SMAP actually delivers a strong performance in the recent 10-year period. This is mostly due to the non-traditional factors from the EDGAR filing database. Our EDGAR analytics system starts from mid 1990s. With a 10-year training window, these factors are added to the SMAP model only in recent 10 years.

Figure 51 SMAP Model Performs Better in Recent Years



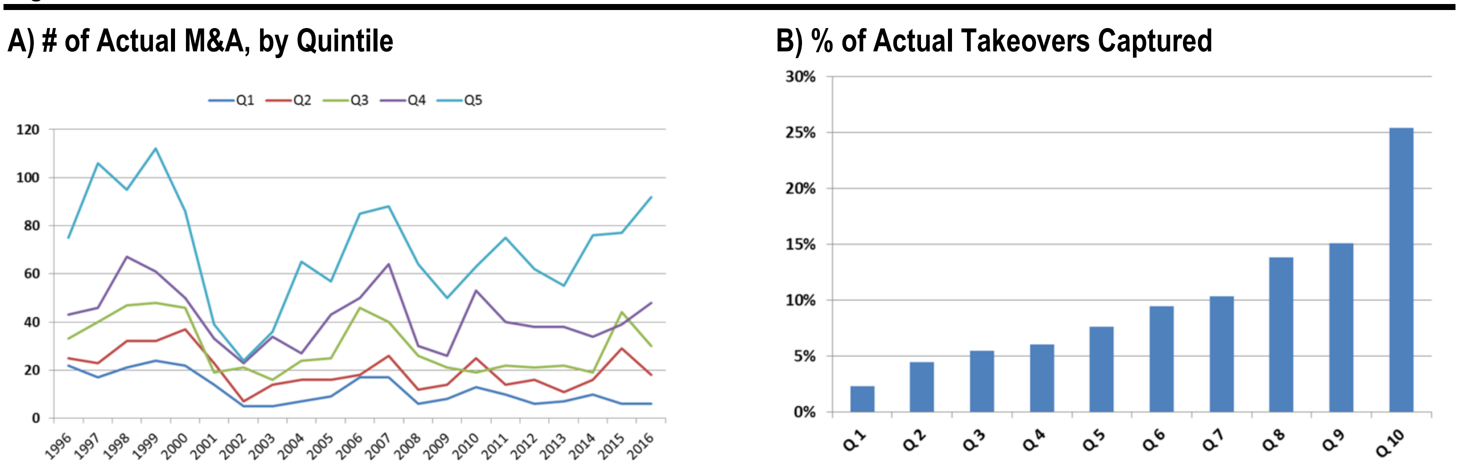
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

OUT-OF-SAMPLE HIT RATE

The output of the SMAP model is essentially the probability of being acquired in the subsequent month. Grouping all companies into five quantiles based on the SMAP probabilities, Figure 52 (A) shows the actual number of M&A transactions, by each of the predicted quintiles form the SMAP. As expected, higher quintiles (higher takeover probabilities) almost always have greater realized M&A transactions than the lower quintiles. The spread between Q5 (the highest quintile) and Q1 is especially striking.

Figure 52 (B) shows the percentage of takeovers captured by each predicted decile. Higher deciles/higher takeover probabilities catch greater percentages of actual takeover transactions than lower deciles. In particular, Decile 10 (decile 10) captures 25% of actual M&A deals before they are announced.

Figure 52 SMAP Model Hit Rate

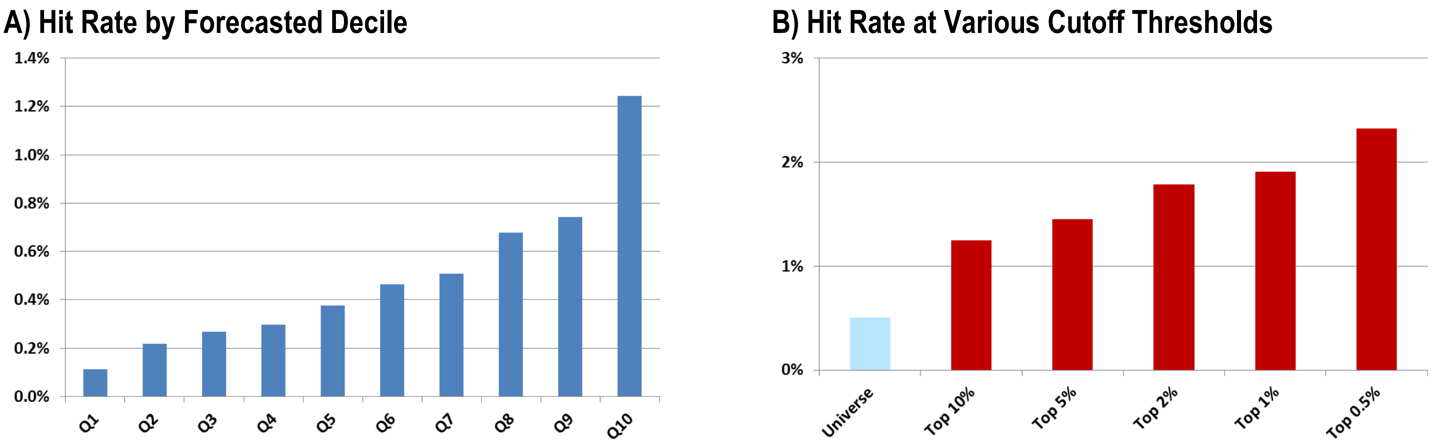


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Figure 53 (A) shows the average hit rate over the next month for each forecasted decile. M&A's are rare events and furthermore, to get the precise timing correct is even more difficult. The SMAP model again shows a monotonic pattern. However, even in the top decile, only 1.2% of the stocks become actual takeover targets in the subsequent.

If we further limit our investment universe to higher and higher probabilities, the hit rate does go up monotonically (see Figure 53 B). The top 0.5% bucket based on the SMAP model has a hit rate almost five times as high as the naïve forecast[6].

Figure 53 Average Hit Rate



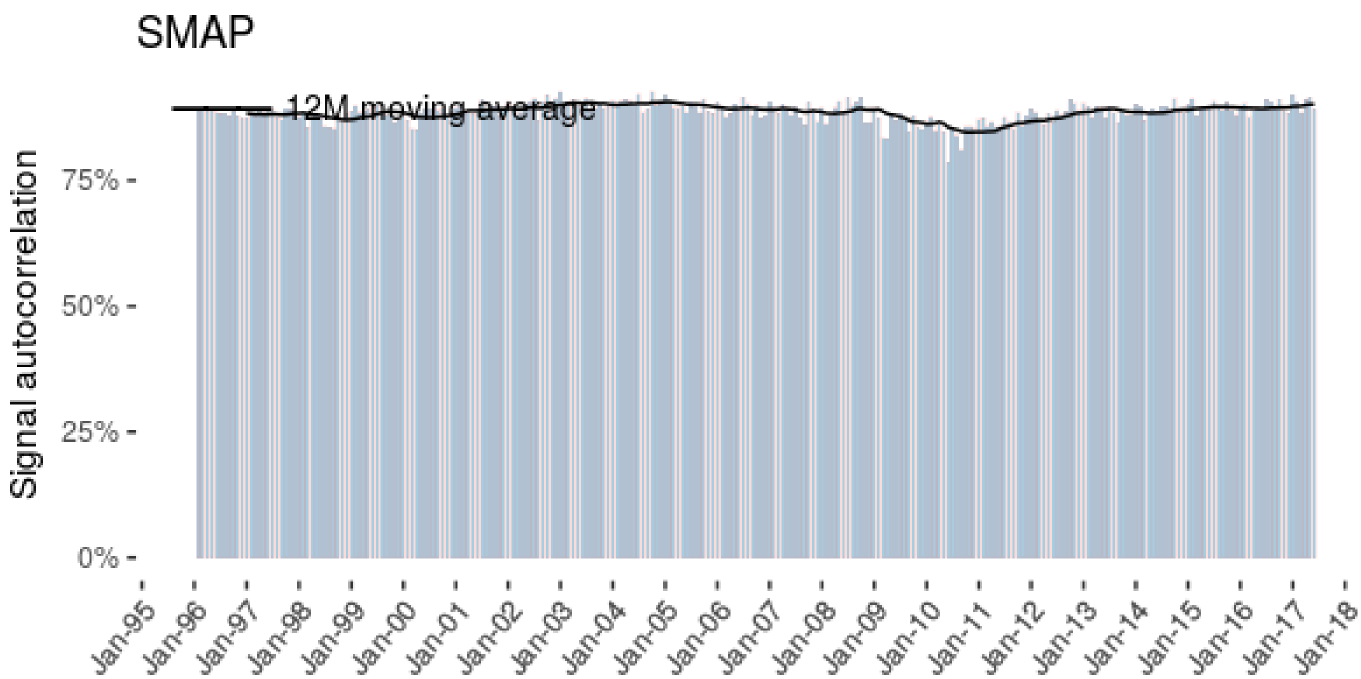
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Forecasting for Different Horizon

Because of the long rolling window (10 years of data) behind the SMAP model, we expect the model to have high autocorrelation and low turnover. As shown in Figure 54, the SMAP model monthly autocorrelation is over 90%.

[6] The naïve forecast is just the average percentage of takeovers, which is around 0.5%. If we do not make any attempt to predict M&A, we can simply use the average takeover ratio as our naïve estimate.

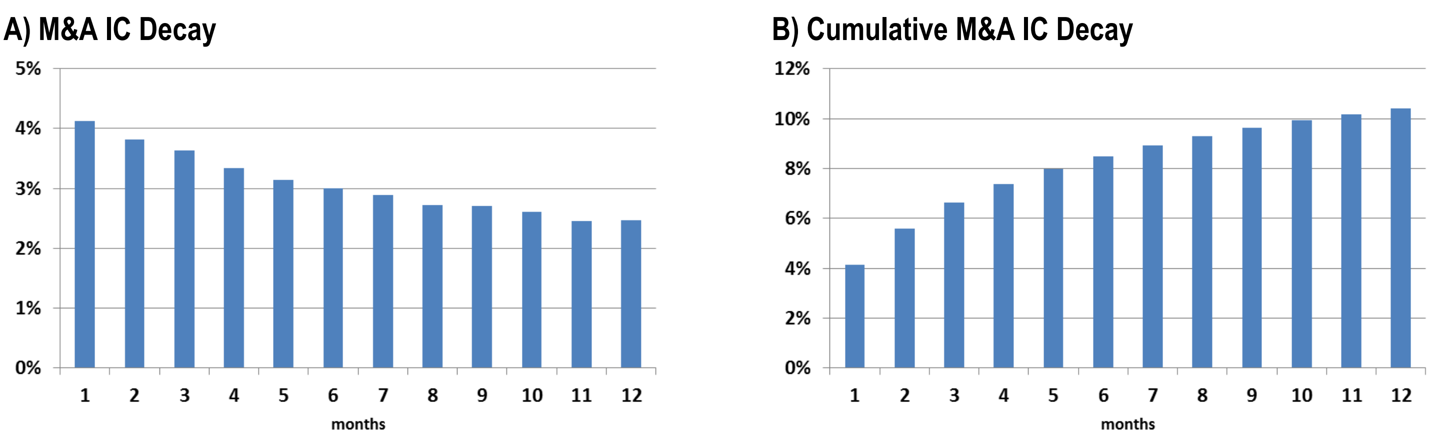
Figure 54 Signal Autocorrelation



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Because of the low turnover nature of the SMAP, it also demonstrates slow decay (see Figure 55 A). The SMAP model still shows strong predictive power a year later. The cumulative M&A IC as shown in Figure 55 B reveals the fact that, if a company has a high predictive probability based on the SMAP model, the chance of actually being taken over in the next year should be fairly high.

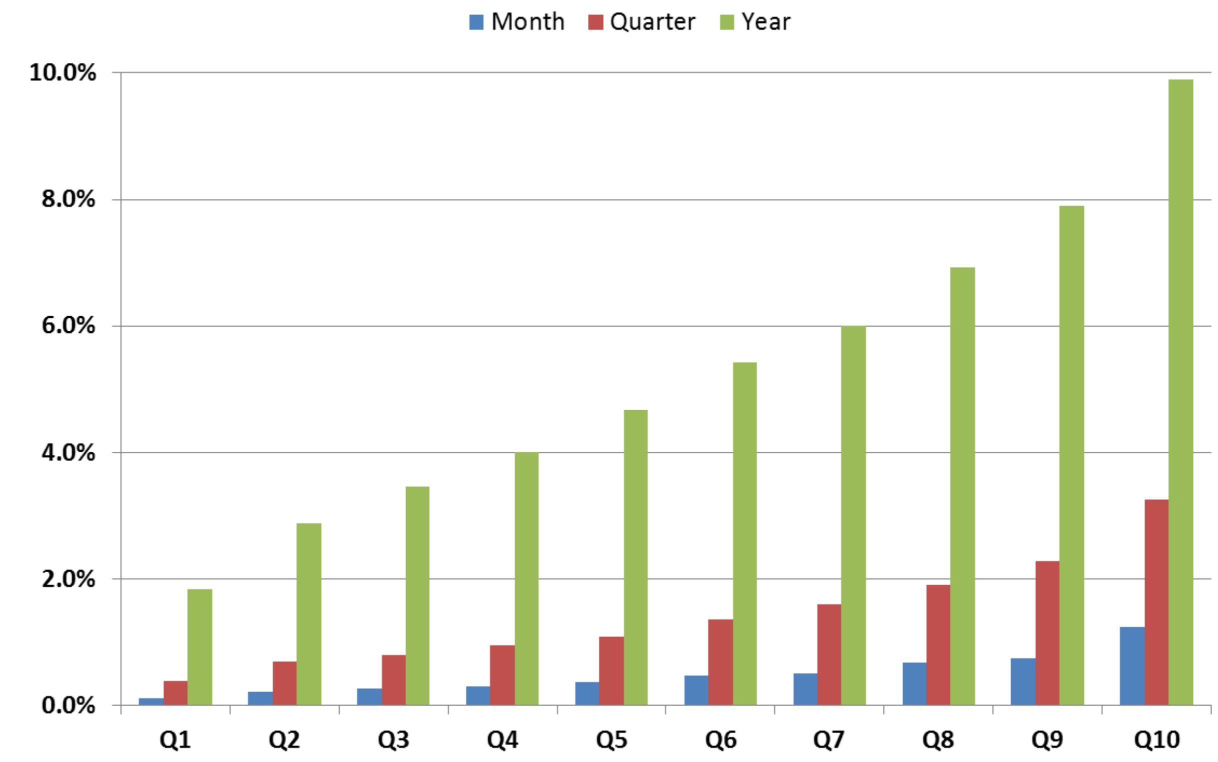
Figure 55 Signal Decay



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Figure 56 shows the hit rate for each SMAP deciles over different forecasting horizons. Although only 1.2% companies in the top decile were acquired in the subsequent month, almost 10% of them turned out to be takeover targets in the following year. Therefore, the SMAP model should be of equal interest to investors with longer holding horizons.

Figure 56 Average Hit Rate at Each Decile for Different Horizon



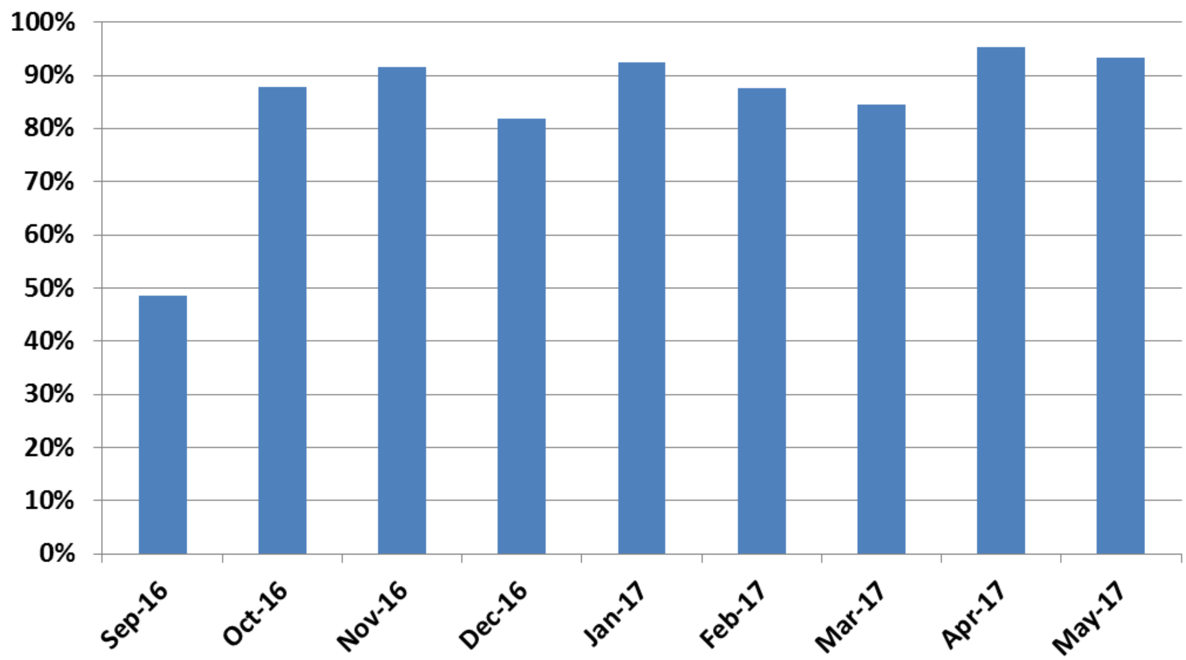
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

A Case Study

In this section, let's show a concrete example of how the SMAP model can be used in practice. On June 16, 2017, Amazon announced its plan to acquire Whole Foods Market. Interestingly, the SMAP model had already put Whole Foods in the top 10th percentile (i.e., the highest takeover probability decile) eight months before the announcement (see Figure 57).

As shown in Figure 57, the SMAP score of Whole Foods Market was less than the 50th percentile at the end of September 2016, but the probability goes up to almost 90% percentile at the end of October 2016. In April 2017, SMAP score further spiked to 95% percentile.

Figure 57 SMAP Percentile for Whole Foods Market before Amazon's Takeover Announcement

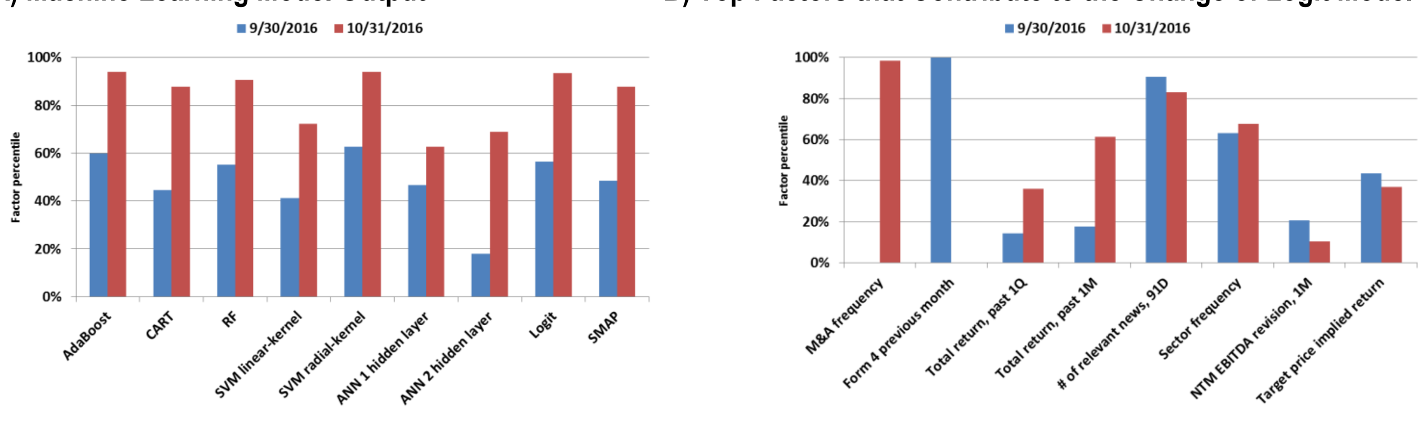


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

As shown in Figure 58 (A), every model witnessed a spike in takeover probability for Whole Foods, from September to October 2016. If we drill down to the logit model, as shown in Figure 58 (B), we can clearly see that M&A frequency and the absence of the Form 4 filing were the top two contributors to the changes in takeover probability, coupled with a jump in share price and a modest increase in sector M&A momentum score. Past M&A frequently was mostly driven by the news that Kroger was making or mulling a bid for Whole Foods on October 6, 2016.

Figure 58 Whole Foods Market M&A Prediction Changes in October 2016

A) Machine Learning Model Output **B) Top Factors that Contribute to the Change of Logit Model**

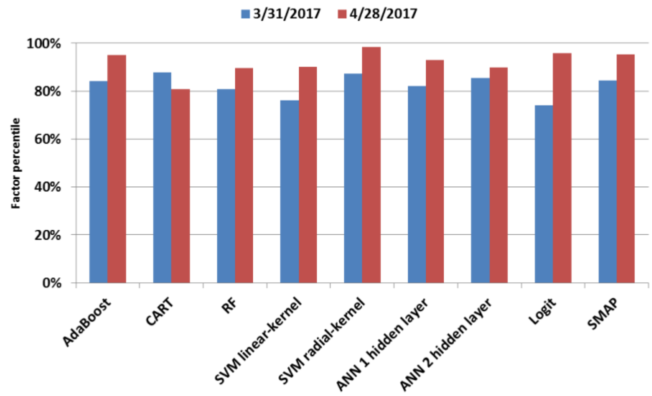


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

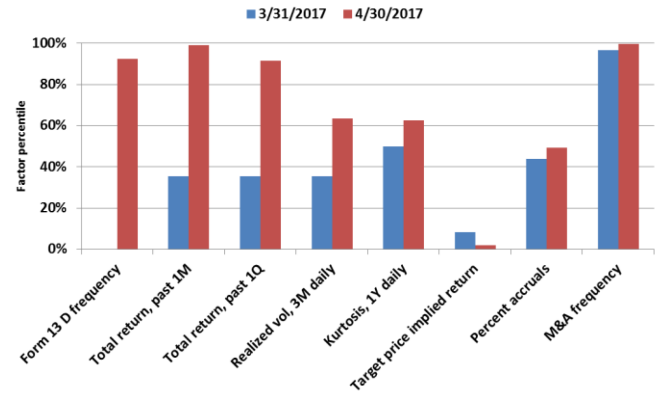
Similarly, as shown in Figure 59 (A), every signal underlying model behind the SMAP also saw an increase in takeover probability for Whole Foods from March 2017 to April 2017. Figure 59 (B) shows that the biggest contributor to the spike in probability for the logit model is the recent Form 13D filing by the activist Jana Partners[7].

Figure 59 Whole Foods Market M&A Prediction Changes in April 2017

A) Machine Learning Model Output



B) Top Factors that Contribute to the Change of Logit Model



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

[7] Jana Partners took a 9% stake in Whole Foods in April 2017.

Practical Applications

Given the high accuracy of the SMAP model, it should be useful in a wide range of applications. Obviously, for discretionary risk arbitrageurs, event-driven managers, and fundamental investors, the SMAP model, given its high accuracy should be a great tool for pre-screening potential takeover targets.

In this section, we want to show two simple applications using the SMAP model:

- Buying potential M&A targets before deal announcements to capture the large positive returns on the M&A announcement date, and
- Avoiding shorting potential M&A targets to reduce the drawdown and volatility on the short side

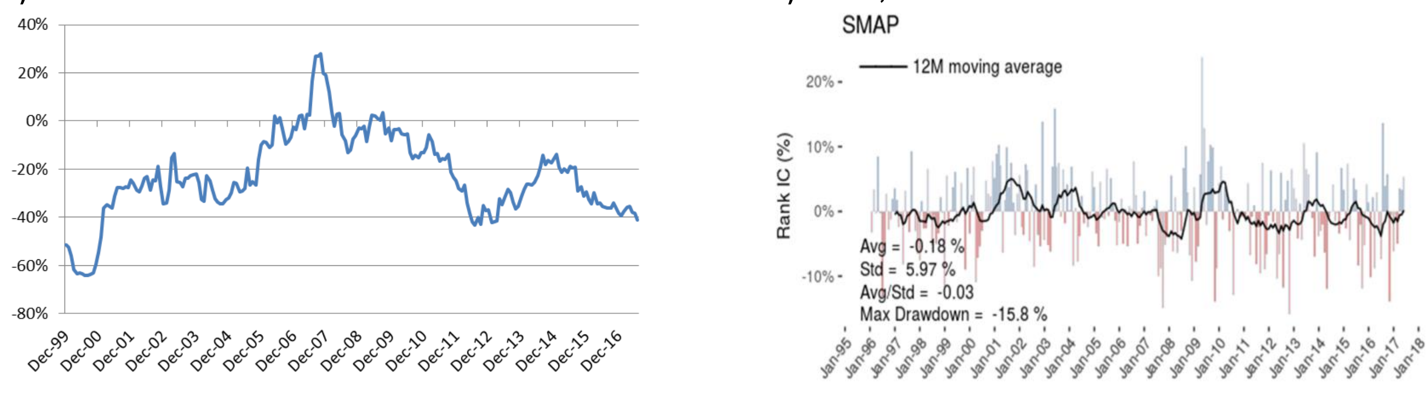
Buying Potential M&A Targets

First of all, we want to remind investors that the SMAP is not a typical stock-selection model. The goal of the SMAP is to predict winners and losers. Rather, the design of the model is to identify potential takeover targets. As shown in Figure 6, on average, merger target stocks suffer from slow earnings growth, expensive valuation, poor price momentum, and disliked by sell-side analysts. Although on the announcement date, if we are able to capture the actual target and its share price spikes, the drag from the other high takeover probability firms can still offset most of the profit.

As shown in Figure 60 (A), the correlation between the SMAP and our flagship global stock-selection model – LEAP is actually negative most of the time. Figure 60 (B) shows the performance of the SMAP as a stock-selection tool (measured by rank IC). On the surface, we can't seem to use the SMAP naively as a stock selection tool. However, we have to keep in mind that rank IC, by construction is robust to outliers. For M&A transactions, the upside is all in the outliers, i.e., those actual takeover target companies. In the next section, we will examine the return profile of the SMAP model, which is more relevant and approximate.

Figure 60 The Nature of the SMAP Model

A) Correlation between the SMAP and LEAP **B) SMAP, Rank IC**



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

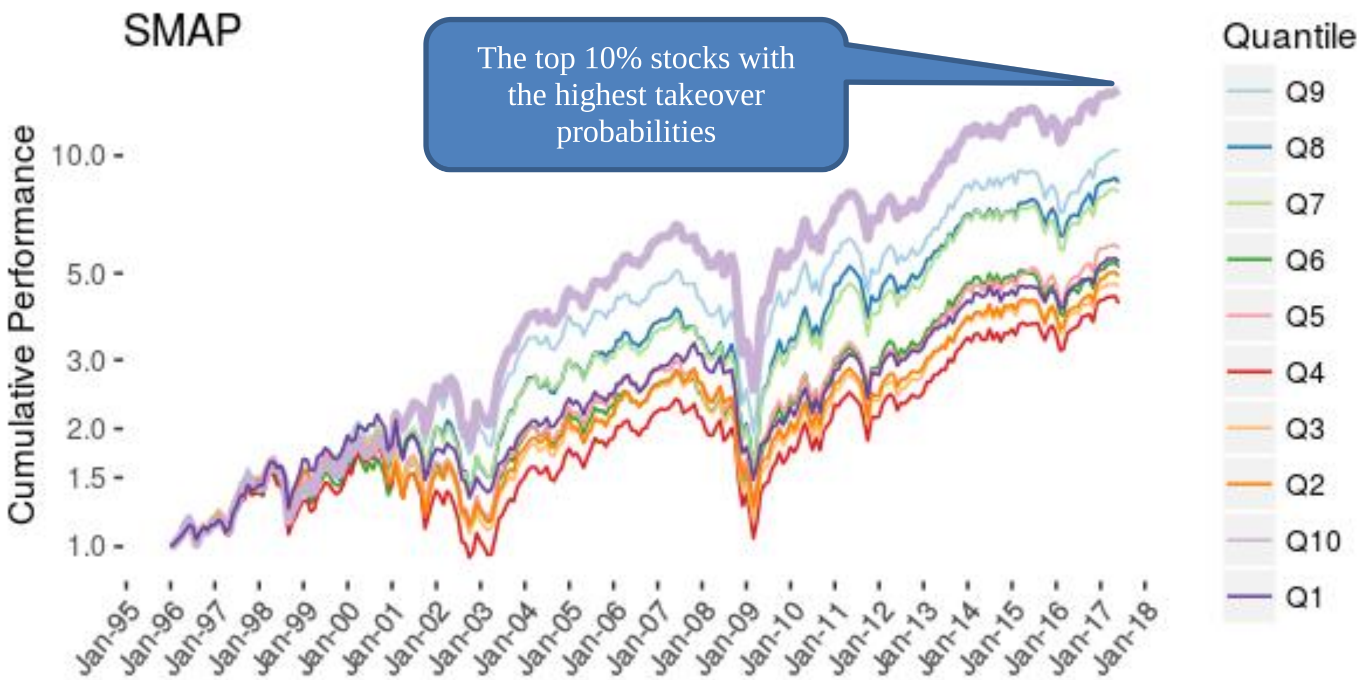
Whether investing in potential M&A targets depend on a number of factors. Not only the model's accuracy (in predicting takeovers) matters, but also we need to measure the potential negative drag from those non-target firms. A model could be highly effective in M&A prediction, but still delivers negative returns.

As shown in Figure 61 (A), the top decile (Q10 or stocks with the highest takeover probabilities) based on the SMAP model clearly outperforms the market and all other deciles. Figure 61 (B) gives a clearer picture of the annualized return for each decile portfolios. The payoff pattern is not linear. The returns from Decile one to Decile six are almost identical, and then the payoff pattern shifts a monotonic upward slope from Decile seven to Decile ten. Stocks with the highest takeover probabilities produce outsized high returns.

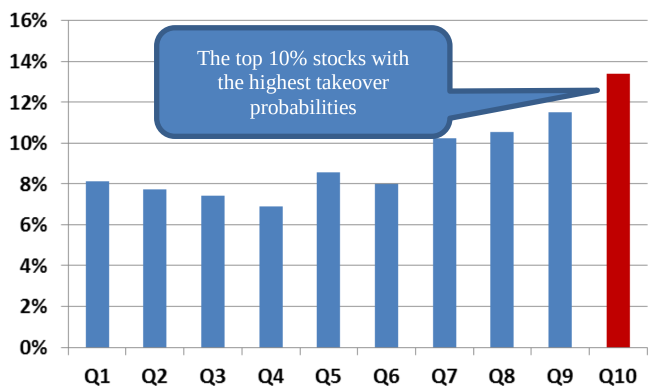
The significant outperformance of the top decile is mainly due to the extremely high returns from actual M&A targets. As shown in Figure 61 (C), the skewness is much higher in the top decile compared with other deciles.

Figure 61 Performance of the SMAP Model as a Stock-Selection Tool

A) Cumulative Performance, by Deciles



B) CAGR, by Decile



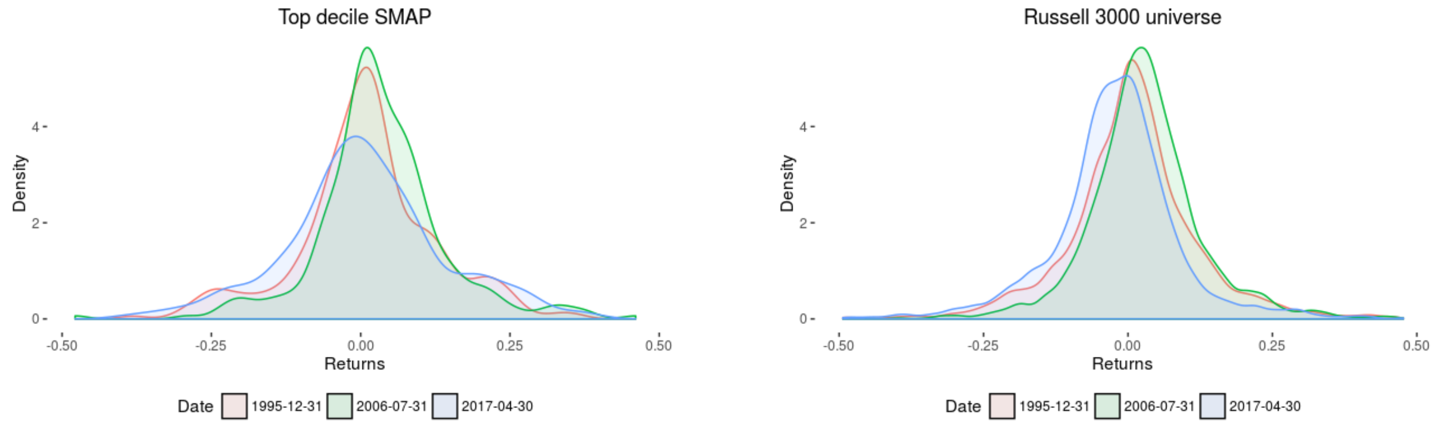
C) Median Skewness, by Decile

Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

The high positive skewness witnessed by the highest takeover probability decile can also be seen in Figure 62(A). The return distribution for the entire Russell 3000 universe appears to be more balanced (see Figure 62 B).

Figure 62 Return Distribution

A) Return Distribution for the Top Decile SMAP **B) Return Distribution for the Russell 3000 Universe**

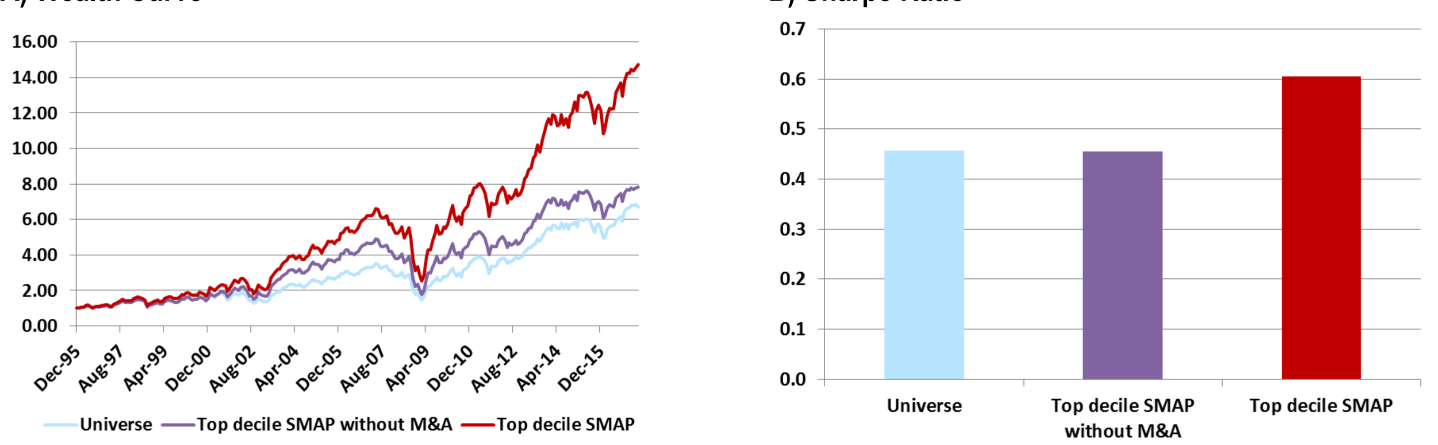


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

To further prove that the high skewness (and positive return) from the top SMAP decile portfolio is due to the returns from M&A targets, we compute the performance of the top decile portfolio with and without the actual M&A target names included. As shown in Figure 63 (A). if we exclude those actual takeover targets from the top decile, the actual return and Sharpe ratio of the portfolio is about the same as the overall market. On the other hand, once we put the targets back in the portfolio, the SMAP significantly outperforms the market.

Figure 63 Performance for the SMAP Top Decile with and without the M&A Target

A) Wealth Curve **B) Sharpe Ratio**



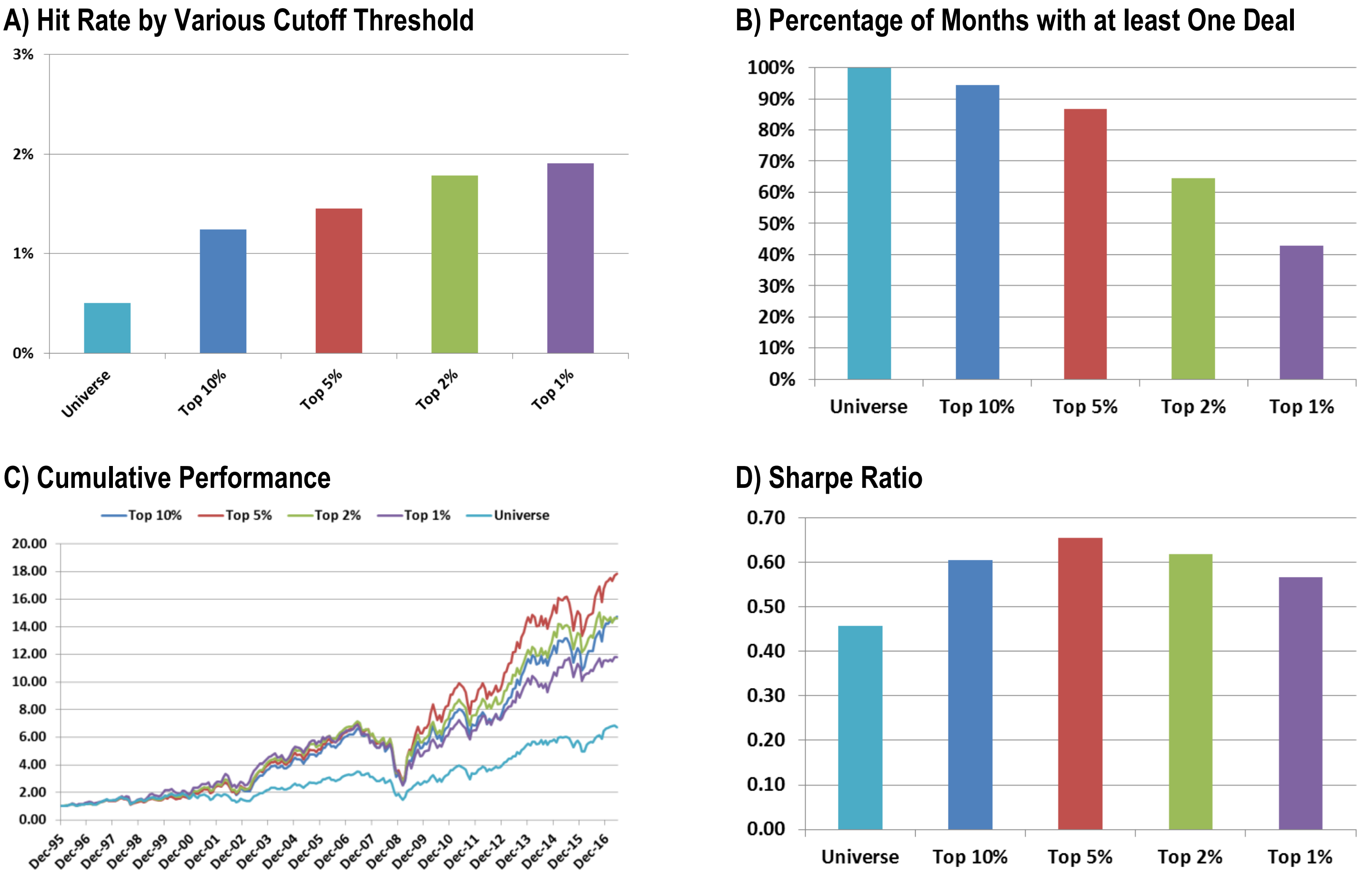
Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

The Optimal Cutoff Threshold

As shown in Figure 64 (A), the SMAP model's hit rate increases monotonically as we tighten out cutoff threshold. For example, if we limit our investment to the top 1% companies with the highest takeover probabilities, we can move up the hit rate to almost 2%, from the 1.3% hit rate for the top decile portfolio – an increase of almost 50%. However, as we lift up our threshold, our portfolio becomes more concentrated. The top 1% portfolio has 30 stocks, while the top decile portfolio includes 100 stocks. A more concentrated portfolio means that there might be no takeover target at all in certain months. As shown in Figure 64 (B), as the portfolio becomes more concentration, the percentage of months with at least one M&A transaction also decreases monotonically.

Using the SMAP as a direct investment tool, if we only buy the top $x\%$ stocks with the highest takeover probabilities, as shown in Figure 64 (B) and (C), 5% is the optimal cutoff threshold, which gives us 150 stocks per month.

Figure 64 Performance for the SMAP Top Decile with and without the M&A Target

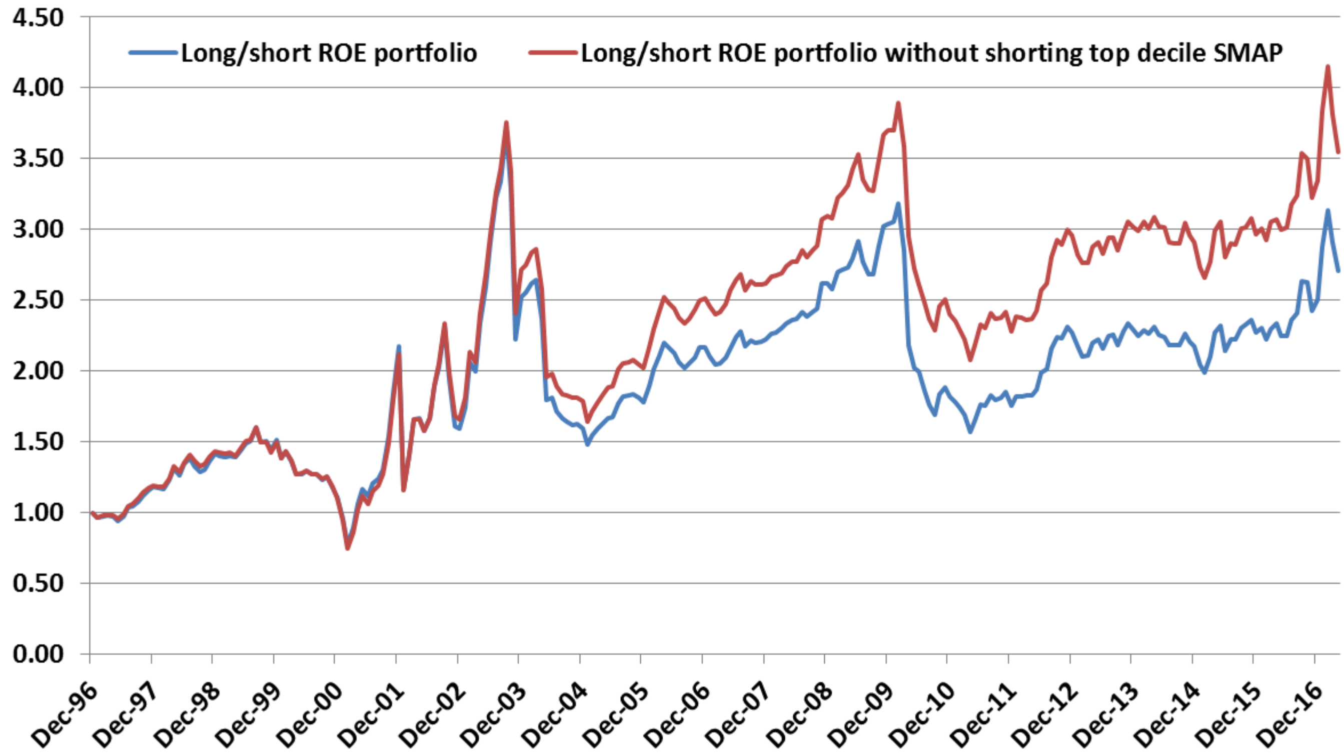


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

Avoiding Shorting Potential Takeover Targets

For long/short portfolio managers, one of the biggest worries is that some companies on the short portfolio become takeover targets. Therefore, another major application of the SMAP model is to help managers to avoid shorting potential M&A targets. As a simple example, if we remove those stocks with the highest 10% SMAP scores from the short portfolio based on the ROE factor, we can see noticeable performance boost (see Figure 65).

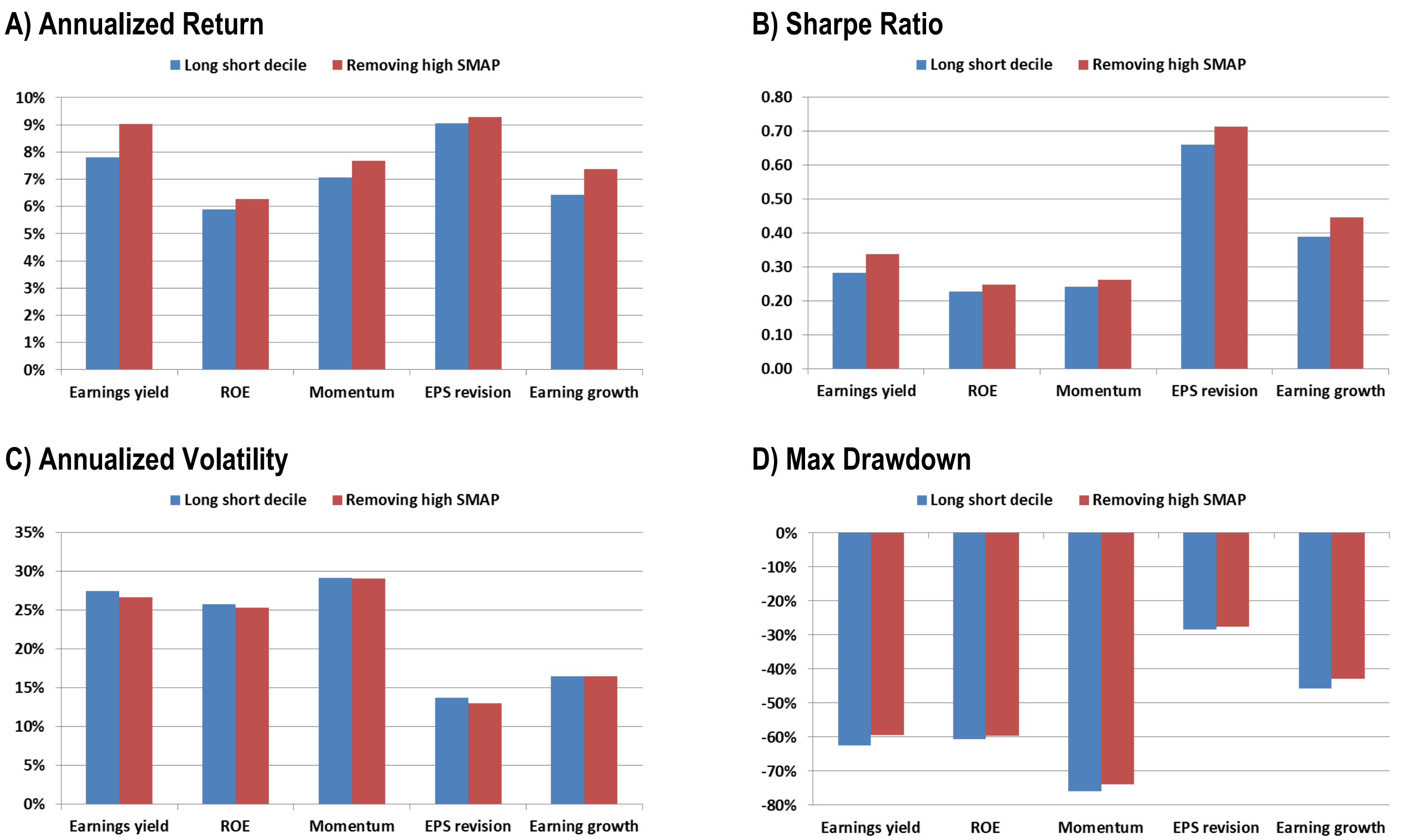
Figure 65 Cumulative Performance, ROE Factor Portfolio (Long/Short Decile Portfolio)



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

As shown in Figure 66 (A) and (B), we see across the board performance improvement by removing the highest takeover probability stocks from the short side, for value, growth, momentum, quality, and earnings revision factors. Furthermore, by not shorting potential M&A targets also reduces the risk (both volatility and maximum drawdown) for all common factors (see Figure 66 C and D).

Figure 66 Performance Increases after Removing High-Takeover-Probability Stocks (based on the SMAP) from the Short Side

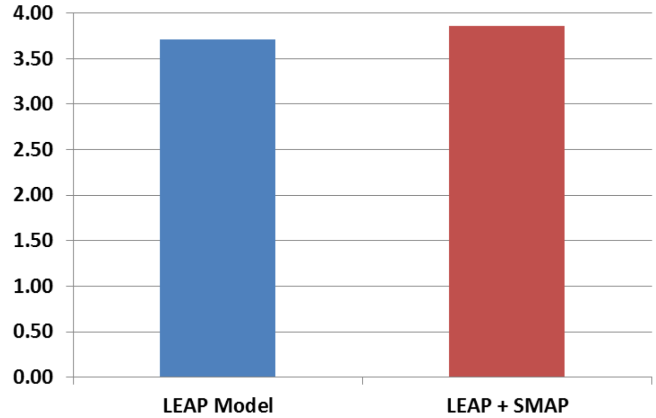


Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

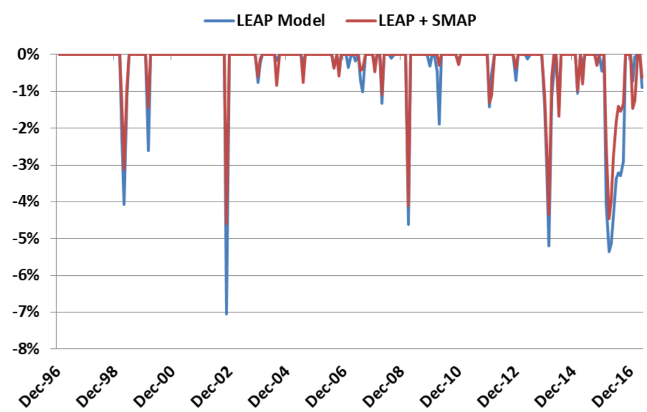
Finally, removing high takeover probability stocks from the short side also helps high efficacy alpha models such as the LEAP. As shown in Figure 67, the Sharpe ratio for the long/short decile portfolio based on the LEAP model increases by almost 5%, and the maximum drawdown dropped over one third, with the SMAP overlay.

Figure 67 Improving the Performance of LEAP using the SMAP Model

A) Sharpe Ratio



B) Max Drawdown



Sources: Bloomberg Finance LLP, FTSE Russell, S&P Capital IQ, Thomson Reuters, EDGAR, Wolfe Research Luo's QES

BIBLIOGRAPHY

Ambrose, B. W., and Megginson, W. L. [1992]. “The role of asset structure, ownership structure, and takeover defenses in determining acquisition likelihood”, *Journal of Financial and Quantitative Analysis*, 27(4), 575-589.

Bishop, C. [1995]. *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford

Breiman, L. [1996]. “Bagging Predictors”, *Machine Learning*, 24(2), 123-140

Chatterjee, S., John, K., & Yan, A. [2011]. “Takeovers and divergence of investor opinion”. *Review of Financial Studies*, 25(1), 227-277.

Clement, R.T. [1989]. “Combining Forecasts: A Review and Annotated Bibliography”, *International Journal of Forecasting*, 5, 559-583.

Cocco, J. F., & Volpin, P. F. [2013]. “Corporate pension plans as takeover deterrents”, *Journal of Financial and Quantitative Analysis*, 48(4), 1119-1144.

Comment, R., & Schwert, G. W. [1995]. “Poison or placebo? Evidence on the deterrence and wealth effects of modern antitakeover measures”, *Journal of Financial Economics*, 39(1), 3-43.

Creamer, G. G., & Freund, Y. [2005]. “Using AdaBoost for equity investment scorecards”, SSRN paper, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=940729

Creamer, G., & Freund, Y. [2010]. “Automated trading with boosting and expert weighting”, *Quantitative Finance*, 10(4), 401-420.

Edmans, A., Goldstein, I., & Jiang, W. [2012]. “The real effects of financial markets: The impact of prices on takeovers”, *Journal of Finance*, 67(3), 933-971.

Dessaint, O., Golubov, A., & Volpin, P. [2017]. “Employment protection and takeovers”, *Journal of Financial Economics*, May 10

Hasbrouck, J. [1985]. “The characteristics of takeover targets q and other measures”, *Journal of Banking & Finance*, 9(3), 351-362.

Hoberg, G., & Phillips, G. [2010]. “Product market synergies and competition in mergers and acquisitions: A text-based analysis”, *Review of Financial Studies*, 23(10), 3773-3811.

Jussa, J., Luo, Y., and Wang, S. [2017]. “Quant CSI”, Wolfe Research Luo’s QES, March 13, 2017

Luo, Y., Jussa, J., and Wang, S. [2017a]. “The Big and The Small Sides of Big Data”, Wolfe Research Luo’s QES, February 8, 2017

Luo, Y., Jussa, J., and Wang, S. [2017b]. “Signal Research and Multifactor Models”, Wolfe Research Luo’s QES, February 16, 2017

Luo, Y., Jussa, J., and Wang, S. [2017c]. “Style Rotation, Machine Learning, and The Quantum LEAP”, Wolfe Research Luo’s QES, February 24, 2017

Luo, Y., Alvarez, M., Jussa, J., Wang, S., and Rohal, G. [2017d]. “From Nowcasting to Forecasting”, Wolfe Research Luo’s QES, September 7, 2017

McCllullagh, P., and Nelder, J.A. [1989]. *Generalized Linear Models*, 2nd Edition, Chapman & Hall/CRC

Morck, R., Shleifer, A., and Vishny, R. W. [1988]. “Characteristics of targets of hostile and friendly takeovers. In Corporate takeovers: Causes and consequences”, (pp. 101-136). University of Chicago Press.

Palepu, K. G. [1986]. “Predicting takeover targets: A methodological and empirical analysis”. *Journal of Accounting and Economics*, 8(1), 3-35.

Ripley, B. [1996]. *Pattern Recognition and Neural Networks*, Cambridge University Press

Rohal, G., Luo, Y., Jussa, J., and Wang, S. [2017]. “Text Mining Unstructured Corporate Filing Data” Wolfe Research Luo’s QES, April 20, 2017

Routledge, B. R., Sacchetto, S., & Smith, N. A. [2013]. “Predicting Merger Targets and Acquirers from Text”, Working Paper, Carnegie Mellon University.

Schapire, R. E., Freund, Y., Bartlett, P., and Lee, W. S. [1998]. “Boosting the margin: A new explanation for the effectiveness of voting methods”. *Annals of Statistics*, 26(5), 1651-1686.

Shivdasani, A. [1993]. “Board composition, ownership structure, and hostile takeovers”, *Journal of Accounting and Economics*, 16(1), 167-198.

Vapnik, V. [2010]. *The Nature of Statistical Learning Theory*, Springer

Wang, S., Ai, H., Yamashita, T., & Lao, S. [2010]. “Combined top-down/bottom-up human articulated pose estimation using Adaboost learning”. In *Pattern Recognition* (ICPR), 2010 20th International Conference on (pp. 3670-3673). IEEE.

Wu, B., Ai, H., Huang, C., & Lao, S. [2004]. “Fast rotation invariant multi-view face detection based on real adaboost”. In Automatic Face and Gesture Recognition, 2004. *Proceedings. Sixth IEEE International Conference* on (pp. 79-84). IEEE.

DISCLOSURE SECTION

Analyst Certification:

The analyst of Wolfe Research primarily responsible for this research report whose name appears first on the front page of this research report hereby certifies that (i) the recommendations and opinions expressed in this research report accurately reflect the research analysts' personal views about the subject securities or issuers and (ii) no part of the research analysts' compensation was, is or will be directly or indirectly related to the specific recommendations or views contained in this report.

Other Disclosures:

Wolfe Research, LLC does not assign ratings of Buy, Hold or Sell to the stocks it covers. Outperform, Peer Perform and Underperform are not the respective equivalents of Buy, Hold and Sell but represent relative weightings as defined above. To satisfy regulatory requirements, Outperform has been designated to correspond with Buy, Peer Perform has been designated to correspond with Hold and Underperform has been designated to correspond with Sell.

Wolfe Research Securities and Wolfe Research, LLC have adopted the use of Wolfe Research as brand names. Wolfe Research Securities, a member of FINRA (www.finra.org) is the broker-dealer affiliate of Wolfe Research, LLC and is responsible for the contents of this material. Any analysts publishing these reports are dually employed by Wolfe Research, LLC and Wolfe Research Securities.

The content of this report is to be used solely for informational purposes and should not be regarded as an offer, or a solicitation of an offer, to buy or sell a security, financial instrument or service discussed herein. Opinions in this communication constitute the current judgment of the author as of the date and time of this report and are subject to change without notice. Information herein is believed to be reliable but Wolfe Research and its affiliates, including but not limited to Wolfe Research Securities, makes no representation that it is complete or accurate. The information provided in this communication is not designed to replace a recipient's own decision-making processes for assessing a proposed transaction or investment involving a financial instrument discussed herein. Recipients are encouraged to seek financial advice from their financial advisor regarding the appropriateness of investing in a security or financial instrument referred to in this report and should understand that statements regarding the future performance of the financial instruments or the securities referenced herein may not be realized. Past performance is not indicative of future results. This report is not intended for distribution to, or use by, any person or entity in any location where such distribution or use would be contrary to applicable law, or which would subject Wolfe Research, LLC or any affiliate to any registration requirement within such location. For additional important disclosures, please see www.wolferesearch.com/disclosures.

The views expressed in Wolfe Research, LLC research reports with regards to sectors and/or specific companies may from time to time be inconsistent with the views implied by inclusion of those sectors and companies in other Wolfe Research, LLC analysts' research reports and modeling screens. Wolfe Research communicates with clients across a variety of mediums of the clients' choosing including emails, voice blasts and electronic publication to our proprietary website.