Expanding our Short Squeeze Model to global markets

Research Signals

In May 2015 we introduced the Research Signals **Short Squeeze Model** to systematically score US stocks based on their potential for a short squeeze event. We now expand coverage of the model globally to Developed Europe and Developed Pacific markets and expand our US universe size. Using IHS Markit Securities Finance unique short loan transaction data, our model incorporates capital constraint indicators, which identify names where short sellers have increased potential to cover positions, and events, identifying catalysts for short squeezes. The model can be used to control risk of short squeezes and enhance alpha forecasts based on short interest measures, which we demonstrate by measuring the improvement of our short sentiment factors in a Short Squeeze Model overlay strategy.

- Short squeeze candidates identified by the model within our highly shorted Developed Europe universe had a 59% greater likelihood of squeezing compared with the base universe

- In Developed Pacific, squeezes occurred 1.23% of the time among stocks that the model predicted to squeeze, compared with 0.77% for the universe

- The Short Squeeze Model had the highest hit rate in the US (2.70%) in isolating stocks most likely to squeeze, handily exceeding the universe rate of 1.29%

- US stocks with the highest probability to squeeze outperformed the universe for open-to-close returns, with an additional 4 bps of return on average and 13 bps versus names least likely to squeeze, persisting out to 1-month (6-month) holding periods with 22 bps (1.52%) and 71 bps (4.72%) of additional alpha, respectively, with similarly robust results to our expanded global coverage

Contacts

Research Signals · MK-ResearchSignals@ihsmarkit.com

IHS Markit™

Introduction

We have extensively researched the phenomenon of short squeezes and the complications surrounding their identification. Given the loose use of the term 'short squeeze' in the media and the debatable prevalence of their occurrences, we have developed a systematic identification process for squeezes.

The conditions that we require for a squeeze include a sudden spike in price (3 standard deviation move versus prior 60 trading days over 1-3 days) followed by a decrease in shares on loan (over 5 consecutive days), for names with insufficient supply of shares or high borrowing costs in the securities lending market (bottom 40% of Demand Supply Ratio or Implied Loan Rate). We are also careful to filter out activity related to corporate actions such as dividends which may affect the demand for borrowing shares, with further adjustments in developed Europe and Pacific regions to normalize the securities lending transactions over 10 business days before and after ex-dividend dates to avoid false squeeze prediction (see the Appendix for a detailed definition).

We combine the underlying short loan transaction signals uniquely sourced from IHS Markit Securities Finance (MSF)with our proprietary factor data and news events to measure short squeeze risk. Our MSF daily feed provides an advantage in estimating short squeeze expectations when compared to the users of traditional short interest data provided by the US exchanges on a bi-monthly basis, and adds transparency globally with measures of both lendable supply and demand for borrowing.

We establish that underlying every short sale is a securities lending transaction captured by our comprehensive database. With this transaction-level detail, we can approximate the money-ness of each short sale, i.e., in- or out-of-the-money, an issue confronting short sellers in deciding to maintain or unwind their positions and the urgency of that decision.

We have examined MSF's transaction-level shorting flow data to produce unique indicators that describe the composition of short positions - Profit and Loss Impact, Out-of-the-money Percent, Transact Duration, Max Quantity Bins and Out-of-the-money Days-to-cover. Our research finds a number of these metrics can further distinguish short squeeze risk among highly shorted stocks.

In the remainder of this report we review the Research Signals Short Squeeze Model which incorporates unique insights from the transaction-level capital constraint factors and event indicators that demonstrate its use in predicting squeezes as well as alpha generation. We start with an overview of the datasets used and the model construction. Next, we discuss backtest results of the model, including short squeeze prediction frequencies and model returns. We round out the report with results of strategies using the model as an overlay to short sentiment factors.

Data and methodology

We begin with a description of background data and underlying methodology. First, we review MSF data which provides the underpinnings for the transaction-level indicators. Next, we introduce RavenPack data which is used to classify news events. Lastly, we describe the construction of the model.

IHS Markit Securities Finance data

MSF data provides a timely, detailed look at the short interest market. Names in high demand, a proxy for highly shorted and those with a high cost to borrow tend to underperform the market. At the same time, highly utilized names are at risk of short squeeze. Within the highly utilized set of stocks, we aim to identify those at risk of a squeeze to improve accuracy of short interest signals and provide deeper insight into short positions, the principles behind our detailed definition of a short squeeze.

We hypothesize that short squeezes are more likely to occur for stocks in which short sellers are experiencing capital constraints; in other words, they are losing or are at risk of losing money on their positions. Thus, we turn to MSF transaction data from both lenders and borrowers (net of double counting) to provide insight into the underlying stock lending transactions. The details we look to uncover include the duration of the open position, the price at which the position was entered, the quantity of shares on loan that are losing money and the average break-even price.

Profit and loss (PnL) is a key parameter in the construction of several pertinent factors. Briefly, to compute PnL, we begin by determining the start date of the short sale to set the initial price based on the date that the initial short was placed with the broker. The aggregate PnL for a stock is the weighted sum of all PnLs for each short position, using the number of shares on loan.

RavenPack data

RavenPack produces a structured sentiment scoring system based on unstructured news articles from major media providers and newswires. Over 40,000 listed stocks are covered spanning the Americas, Europe and Asia providing real time statistical summaries of the amount and content of text news. In this model, we use the RavenPack News Analytics, which sources news articles from Dow Jones Newswires, online publications and blogs. The news data provides information about companies cited in each news article, descriptions of the type of news event and sentiment and relevance scores for each event. We leverage both individual articles as well as aggregated sentiment values in this model.

While Ravenpack produces sentiment analysis on a wide range of news stories, our research has narrowed the focus to four specific event types:

1. M&A - we capture confirmed and expected M&A events as well as rumors about M&A activity that can affect short sellers

2. Earnings Sentiment - these articles identify positive sentiment related to earnings releases and forecasts

3. Trading Activity - news articles describing order imbalances with a positive sentiment, index rebalances and stock buybacks are captured in this category

4. Other positive news events such as patent approvals and completed debt restructuring

We look for events that fall into these four categories and which also have high relevance and sentiment scores. In the next section, we describe how this data is implemented into the model.

Model description

The model is constructed with two subcomposites composed of underlying subfactors. The first, Capital Constraints, uses unique MSF transaction-level data to assist in identifying potential short squeezes based on our estimates of short seller profitability and crowding:

- Out-of-the-money Percent (OTM%) - the sum of shares for short positions that are experiencing losses based on their PnL divided by the total shorted quantity. We expect names with a high percent of short sellers out-of-the-money to be at risk of a short squeeze.
- Out-of-the-money Percent - 20-day maximum - the maximum OTM% over the prior 20 trading days. The 20-day maximum value removes the effect of short-term price movement and identifies the 'worst case' scenario for short sellers.

- Short Position Profit Concentration - the distribution of a stock's short loan position profit/loss based on a predefined set of bins. We expect names with a high concentration of short sellers near the break-even point to be at higher risk of a short squeeze.

We also find certain events increase the probability of a short squeeze, which we combine in the second subcomposite called Market Sentiment:

- Earnings announcement events - our research finds that short squeezes happen more frequently around earnings announcement dates. We use this as an indicator to increase the probability of a squeeze five weekdays prior to an earnings announcement and the three weekdays following the announcement.
- Positive news events - we use RavenPack news events to identify potential positive news events that can trigger a short squeeze. Event types include merger and acquisition, earnings, trading and other positive events, as described in the methodology section.
- Abnormal trading volume - we find cases where abnormal trading volume levels paired with positive price movement are indicative of a positive event known to market participants which can trigger a short squeeze.

Finally, our Short Squeeze Model (Figure 1) incorporates the capital constraint and event indicators into a final score. The capital constraint indicators identify names with potential for a short squeeze and are ranked from 1 to 100 and then averaged on an equal-weight basis into a composite rank. The event indicators identify the catalysts for the short squeeze and improve the composite rank based on event type. Positive news events are rewarded with an increase of 20 ranks since we have found the highest connection between these events and future short squeezes. Earnings announcements and abnormal trading volume events increase the composite by 10 ranks. When multiple events occur at the same time, the maximum score increase is 30 ranks.

Figure 1
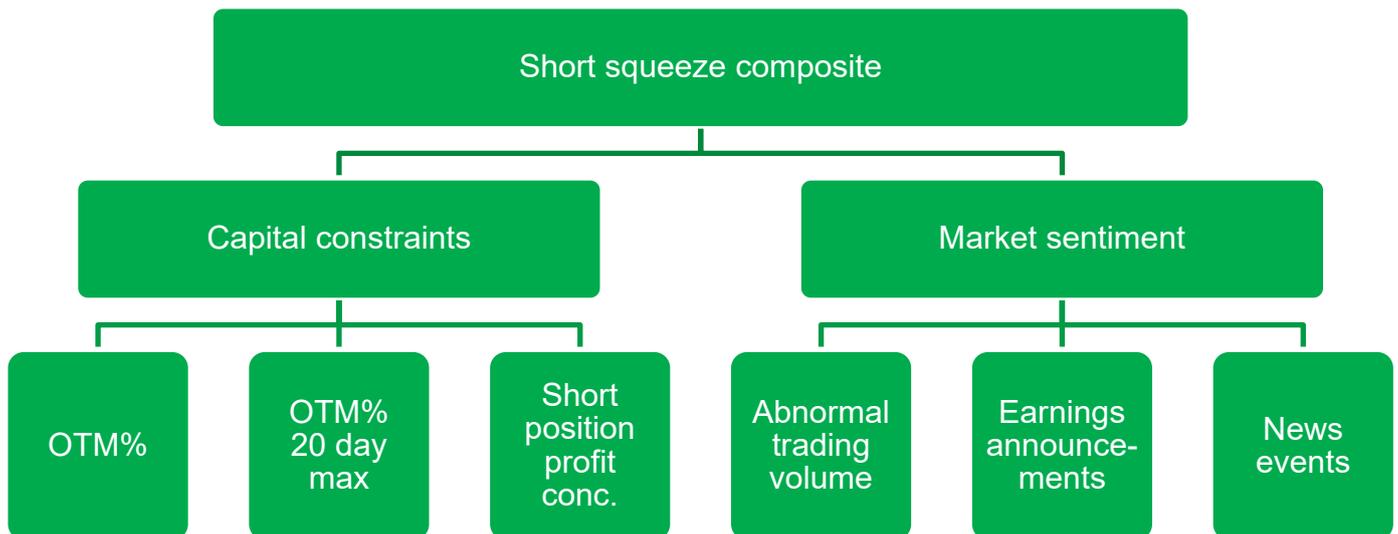
In the following section, we review model performance of highly shorted companies across Developed Europe and Developed Pacific universes (95% of cumulative market cap for each member country), in addition to the newly expanded coverage for the US Total Cap universe (98% of cumulative market cap, or 3,000+ stocks). While our

previous research defined the highly shorted universe as stocks in the bottom 20% of both Demand Supply Ratio and Implied Loan Rate, we now expand the universe to include stocks in the bottom 40% of either of the two factors, enabling us to capture a higher proportion of stocks of interest to short sellers.

Our analysis spans January 2011 to March 2020. We remark that short squeezes do not occur as frequently as commonly cited and results in a minimal set of outcomes, with an annual average of 1,487 (daily average rate of 0.99%) in Developed Europe, 1,782 (0.77%) in Developed Pacific and 3,886 (1.41%) in the US.

Results

Our model performance review covers two aspects of results - short squeeze prediction and alpha generation. We begin with analysis of the likelihood that model scores predict short squeezes. First, we analyze decile groups with decile 1 (D1) representing the names most likely to squeeze and decile 10 (D10) those least likely. Figure 2 displays the percent of names that experienced a short squeeze on average in each decile for each region.

Starting in Developed Europe, based on our model scores we find that squeezes occurred 1.57% of the time in D1, in other words, a 59% greater likelihood versus the universe overall. D1 results for Developed Pacific were slightly lower at 1.23% but represented a 61% greater likelihood compared with the full universe, while the US had the highest hit rate of 2.70%, a 92% improvement over the full universe frequency. Furthermore, the occurrences decrease across deciles for each universe with D10, representing names least likely to squeeze, exhibiting the lowest occurrence.

Figure 2

Percent of names that short squeeze, Jan 2011 - Mar 2020

Source: IHS Markit

© 2020 IHS Markit
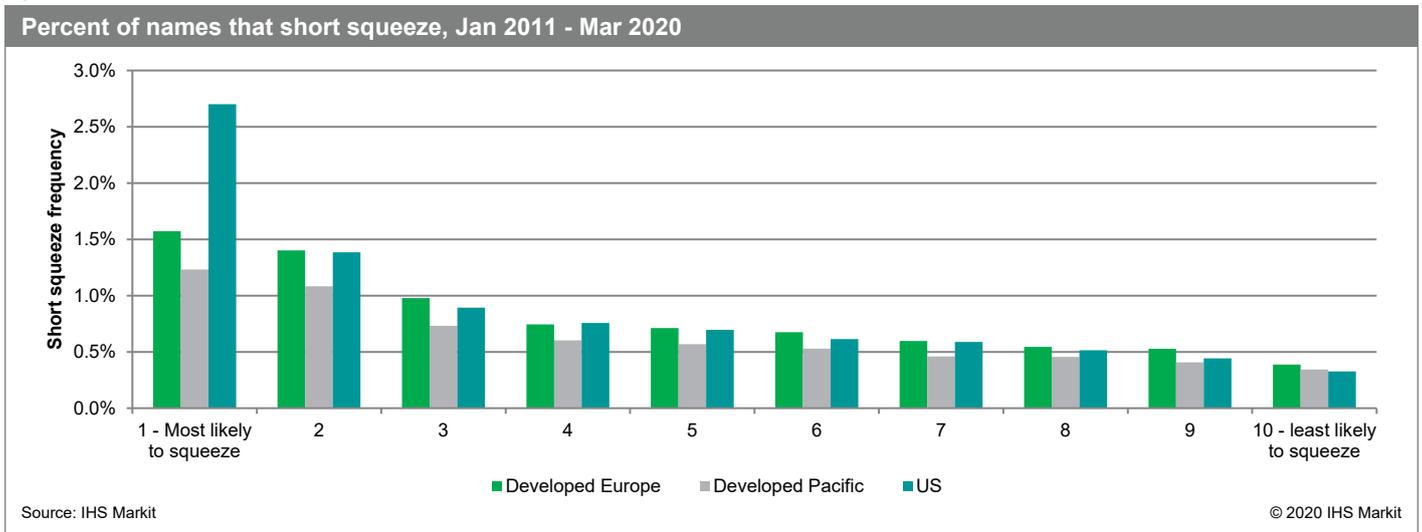
With squeeze prediction established, we consider application of our model in terms of its alpha generating capabilities. Our premise is that stocks which are identified as the most likely to squeeze are expected to outperform given the higher propensity for their prices to increase as short sellers cover their positions.

Thus, we analyze open-to-close, 1-week, 1-month, 3-month and 6-month subsequent spread returns based on model scores for the coverage universes over the analysis period (Table 1). We report the excess return of D1 and D10 short squeeze stocks versus the highly shorted universe along with the spread returns of D1 (highest probability) versus D10 (lowest probability) stocks.

In Developed Europe, our results show that the highest probability stocks (D1) outperform the universe over multiple holding periods. For open-to-close returns, D1 provides an additional 4 bps of return on average versus the universe

and 14 bps versus D10. Positive returns extended out to longer holding periods where D1 outperforms the universe (D10) by 13 bps (64 bps) for 1-month returns and 2.29% (5.33%) out to 6-month returns.

Returns were similarly delivered by the model across Developed Pacific and the US. Developed Pacific stocks with the highest probability to squeeze outperformed for open-to-close returns, with an additional 3 bps on average versus the universe and 7 bps versus the least likely to squeeze. D1 excess returns persisted out to 1-month (40 bps) and 6-month (1.52%) holding periods. Likewise, spread returns remained robust at the 1-month (75 bps) and 6-month (4.25%) time horizons. For the US, we report an open-to-close D1 excess return 4 bps and spread of 13 bps and a 1-month (6-month) D1 excess return of 22 bps (1.52%) and associated D1-D10 spread of 71 bps (4.72%).

Table 1

Short Squeeze Model performance (%), Jan 2011 – Mar 2020				
Universe	Holding period	D1 excess return	D10 excess return	D1 - D10 return spread
Developed Europe	Open-to-close	0.04	-0.09	0.14
	1-week	0.06	-0.12	0.17
	1-month	0.13	-0.52	0.64
	3-month	0.95	-1.59	2.54
	6-month	2.29	-3.04	5.33
Developed Pacific	Open-to-close	0.03	-0.04	0.07
	1-week	0.10	-0.07	0.17
	1-month	0.40	-0.35	0.75
	3-month	0.85	-1.26	2.11
	6-month	1.52	-2.73	4.25
US	Open-to-close	0.04	-0.09	0.13
	1-week	0.06	-0.11	0.17
	1-month	0.22	-0.50	0.71
	3-month	0.84	-1.63	2.47
	6-month	1.52	-3.20	4.72

Source: IHS Markit © 2020 IHS Markit

We further illustrate the favorable model performance with accompanying time series graphs. Figures 3 through 5 display the D1 open-to-close returns relative to each respective universe over the analysis period. We also include the cumulative growth of $1 to demonstrate the persistence in outperformance cumulating to 116% growth in Developed Europe, 75% in Developed Pacific and 115% in the US. Lastly, we take a closer look at D1 monthly returns versus the respective universes (see Figure A1 through A3 in the Appendix). The images show the consistency in outperformance with positive returns in 52% of observations in Developed Europe, 59% in Developed Pacific and 59% in the US.
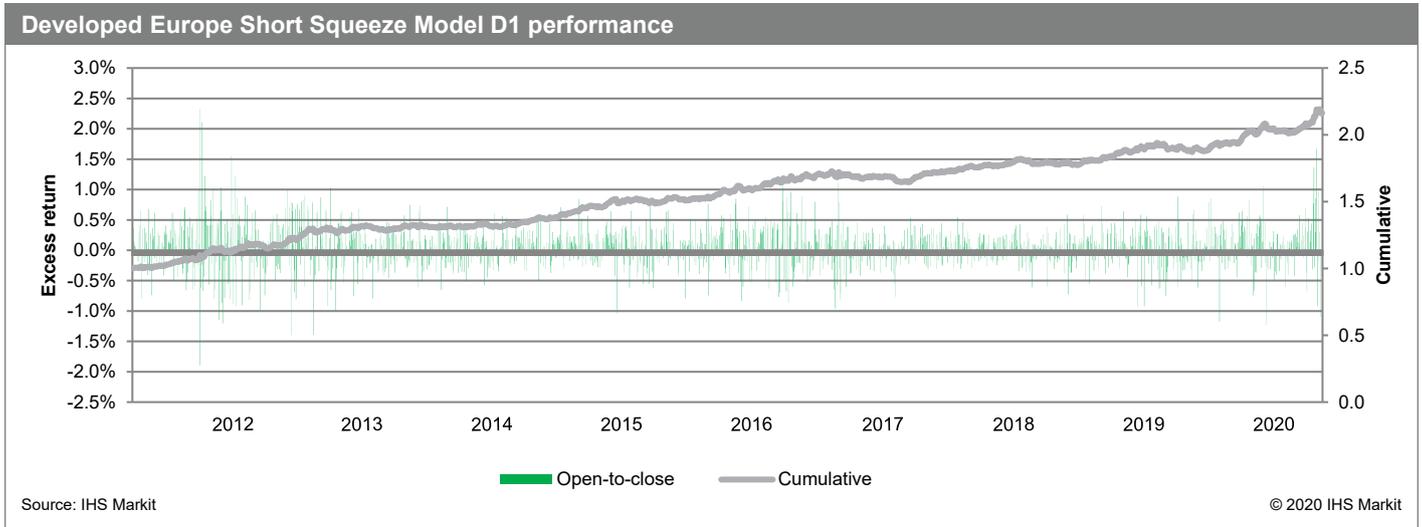
Figure 3

Developed Europe Short Squeeze Model D1 performance

Source: IHS Markit

© 2020 IHS Markit

Figure 4

Developed Pacific Short Squeeze Model D1 performance

Source: IHS Markit

© 2020 IHS Markit

Figure 5

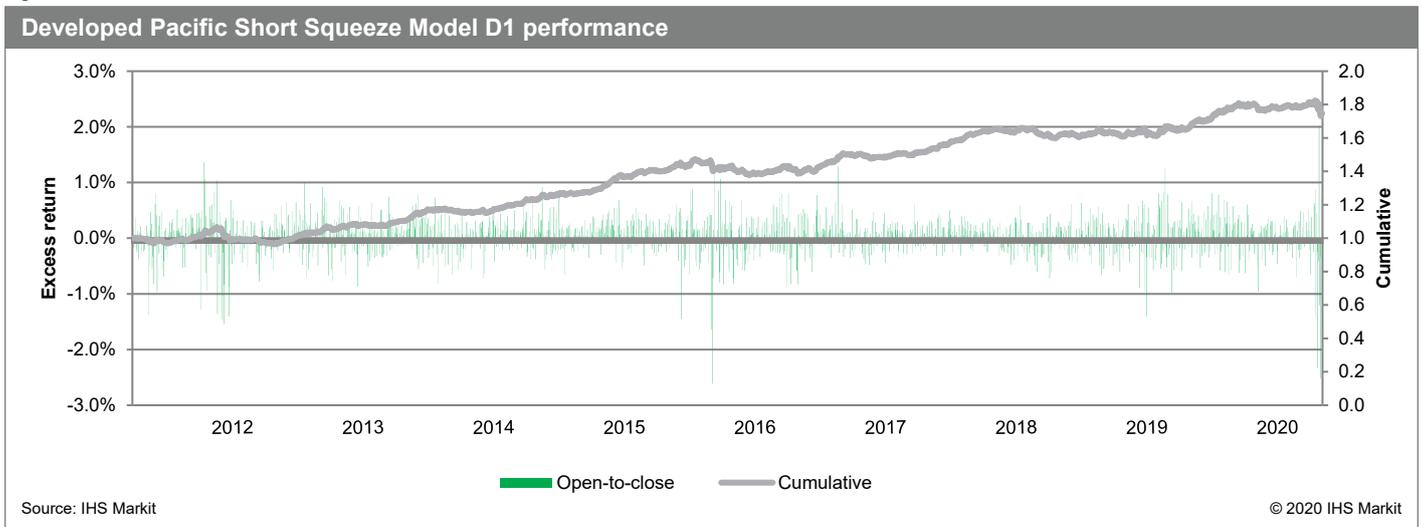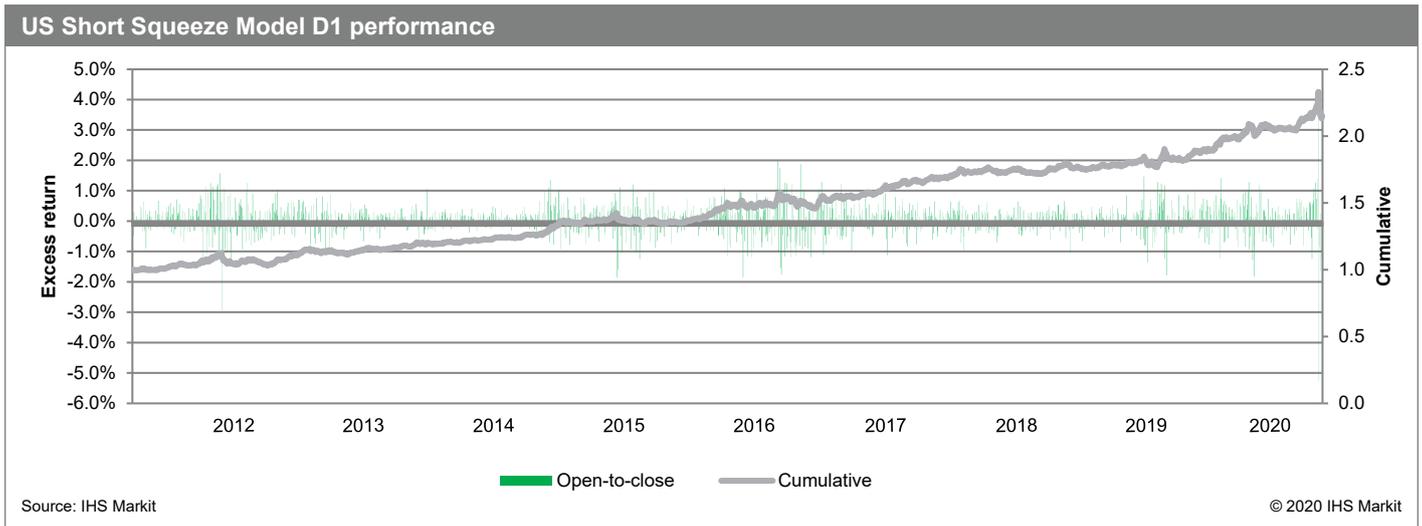US Short Squeeze Model D1 performance

Source: IHS Markit

© 2020 IHS Markit

We round out the results summary by detailing daily turnover statistics of the model ranks, focusing specifically on the top decile (Figure 6). We report the percent of stocks that remain in D1 on a daily basis. As may be expected, when incorporating event measures that are short term in nature, the turnover of the model can by higher than our typical models. However, the results show that the model ranks are relatively stable for D1, with 74% of stocks in Developed Europe and 81% in Developed Pacific remaining in the same decile the next day, though with a lower frequency in the US (62%).

Figure 6

Decile 1 daily turnover, Jan 2011 - Mar 2020

Source: IHS Markit © 2020 IHS Markit
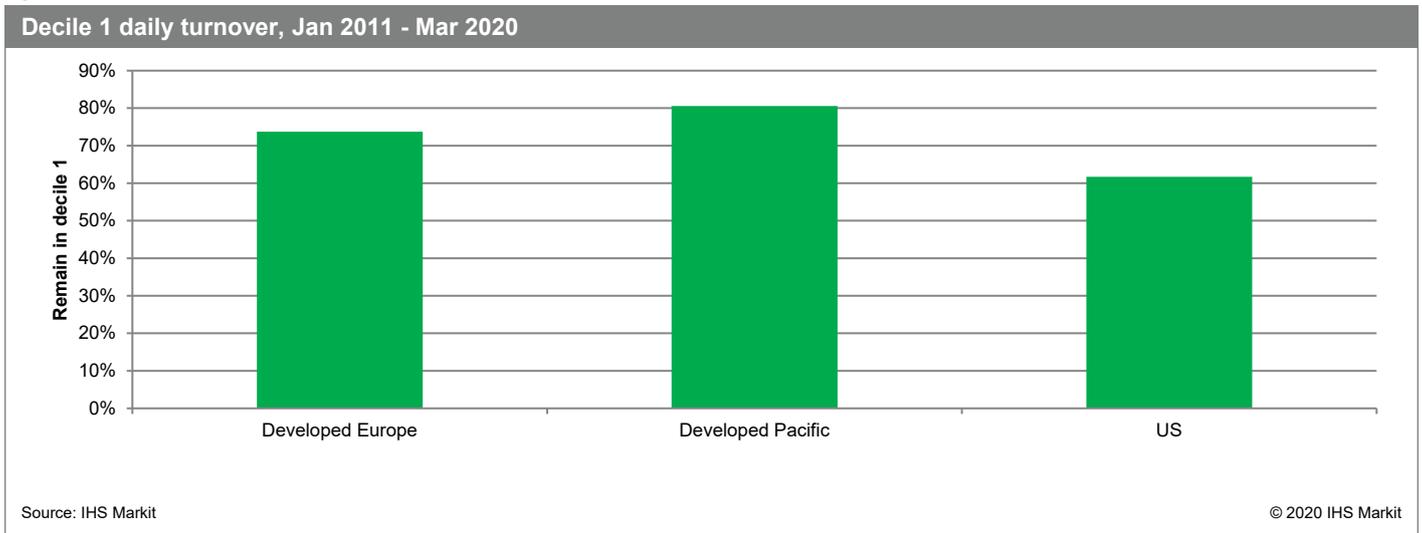
Short Squeeze Model application

Lastly, we turn to using the Short Squeeze Model in combination with Short Sentiment factors, using data from the MSF dataset measuring short interest levels and lending supply. (Table 2). To implement the Short Squeeze Model as an overlay to an existing strategy, we form our long/short portfolios based on factor ranks at the end of each month. The strategy is to go long the names in decile 1 and short the names in decile 10 with a one-month time horizon. We then use the Short Squeeze Model to close out positions in our short portfolio which are at risk of a short squeeze. More specifically, if a stock in decile 10 is ranked in the top decile of the Short Squeeze Model on the date of portfolio formation, we will not include the stock in our short portfolio. In addition, as we move forward throughout the month, if a stock moves into the top decile of the Short Squeeze Model we will close that short position at the close the day following the signal. We compare factor results for the model both with and without adjustment for short squeeze using both mean return and information ratio.

We find that applying the Short Squeeze Model as an overlay to single factors based on short interest data or securities lending data from MSF improves the performance of the factors. In Developed Europe, Short Interest, measuring shares on loan to shares outstanding, saw an improvement on average return spread from 0.34% to 0.94% monthly. For Days to Cover, also known as the short interest ratio measuring the ratio of shares shorted to trading volume, we again see a healthy improvement from 0.32% to 1.00%. Lending Supply, a proxy for institutional ownership calculated as the total quantity of stock made available by custodians relative to shares outstanding, posted a more modest 0.24% improvement on average per month.

In Developed Pacific, Short Interest's average monthly spread improved from 0.28% to 1.13% with the Short Squeeze Model overlay. Increases of 0.36% and 0.71% were also posted by Days to Cover and Lending Supply, respectively. In the US, we observe less impactful, though positive, results on the securities lending factors from the model overlay strategy.

Table 2

Factor and Short Squeeze Model overlay average monthly performance (%), Jan 2011 – Mar 2020				
Universe	Factor	D1 – D10	D1 – D10 with Short Squeeze Model overlay	Improvement
Developed Europe	Short Interest	0.34	0.94	0.60
	Days to Cover	0.32	1.00	0.68
	Lending Supply	0.24	0.48	0.24
Developed Pacific	Short Interest	0.28	1.13	0.85
	Days to Cover	0.06	0.43	0.36
	Lending Supply	0.62	1.33	0.71
US	Short Interest	0.51	0.52	0.02
	Days to Cover	0.35	0.38	0.03
	Lending Supply	0.78	0.80	0.03

Source: IHS Markit © 2020 IHS Markit

Conclusion

We expand our coverage of the Research Signals Short Squeeze Model globally to Developed Europe and Developed Pacific markets. The model was first introduced in US markets in May 2015 to systematically score stocks based on their potential for a short squeeze event. Our model incorporates capital constraint indicators - Out-of-the-money Percent, Out-of-the-money Percent 20-day Maximum and Short Position Concentration - constructed using MSF's unique short loan transaction data to identify names where short sellers have increased potential to cover positions. We also find that certain events including earnings announcements, positive news events identified from RavenPack and abnormal trading volume increase the probability of a short squeeze acting as catalysts for the squeeze.

We tested model efficacy in terms of its short squeeze predictability and found that, for Developed Europe, squeezes occurred 1.57% of the time in decile 1 which isolates names most likely to squeeze, in other words, a 59% greater likelihood than the base universe at 0.99%. Results for Developed Pacific were marginally lower at 1.23%, while US markets posted stronger results at 2.70%, nearly double the rate of the full universe. Furthermore, the occurrences decreased across deciles for each universe with decile 10, representing names least likely to squeeze, exhibiting the lowest occurrence.

In terms of the model's alpha generating capabilities, stocks with the highest probability to squeeze on average outperform the universe and those stocks with the lowest probability to squeeze for open-to-close returns, and persist out to 1-week as well as 1-, 3- and 6-month holding periods. Developed Europe posted an average open-to-close return of 4 bps versus the universe and 14 bps versus names least likely to squeeze, with comparable results of 3 bps and 7 bps, respectively, in Developed Pacific and 4 bps and 13 bps, respectively in the US. Positive returns extended out to 1-month (6-month) holding periods with decile 1 returns in excess of the universe of 13 bps (2.29%) in Developed Europe, 40 bps (1.52%) in Developed Pacific and 22 bps (1.52%) in the US, with similarly robust spreads over the least likely to squeeze.

The model can also be used to improve alpha forecasts based on several well-followed short interest measures. In Developed Europe, Short Interest decile spreads increased by an average of 0.60% monthly, while Days to Cover (0.68%) and Lending Supply (0.24%) also improved. Likewise, improvements of 0.85%, 0.36% and 0.71%, respectively, were posted in Developed Pacific, while US results were more neutral.

Appendix

Background

Short selling refers to sale of a security that the seller does not own, where the delivered security is borrowed by the short seller. The intention is to buy the security at a lower price in the future. In order to lock-in profit, or to avoid further losses (where the price of the security has gone up), short sellers need to cover a short position which involves buying securities in the market and returning the borrowed stock to the lender. The short seller may also be forced to cover positions due to failure to meet a margin call or when the security lender recalls the stock.

Short sellers need to deliver the stock on settlement day, in the same way as any other trade. Since they do not own the stock, they have to borrow it or face penalties for a failed trade. Naked short sales (where the security has not been located and/or borrowed in advance) are now banned in most jurisdictions across the globe. As a result, short sellers almost always need to borrow stock, and as such the resulting lending data provides a close proxy for short selling volumes.

Securities lending is a market practice whereby securities are temporarily transferred by the lender to the borrower. The borrower is obliged to return the securities either on demand or at the end of any pre-agreed term. Securities lending operates as an "over the counter" market. MSF's data provides benchmarking and transparency for all participants in the securities lending market by capturing the daily supply, demand and borrowing costs of individual securities. Information is sourced directly from leading industry participants including prime brokers, custodians, asset managers and hedge funds.

MSF covers more than 3 million intraday transactions, spanning $25 trillion of securities in the lending programs of over 20,000 institutional funds globally. This comprehensive dataset includes a wide range of securities lending metrics collected on a daily frequency. It provides content on the securities lending market including daily shares borrowed, inventory of available shares on loan, level of utilization, loan concentration and stock borrowing costs. It captures around 90% of the securities lending market in developed markets. The coverage can be lower for emerging and frontier areas where the securities lending market is not yet fully developed.

Short squeeze definition

Short selling refers to the sale of a security that the seller does not own, where the delivered security is borrowed by the short seller. The intention is to buy the security at a lower price than that at which the security was sold short. When the price of the security rises, the short seller can incur significant losses as the downside potential due to a price rise is unlimited.

In order to lock-in a profit or avoid further losses (where the price of the security has gone up), short sellers need to cover a short position. This involves buying securities in the market and returning the borrowed stock to the lender. The short seller may also be forced to cover positions due to failure to meet a margin call or when the security lender recalls the stock. The resulting buying pressure can drive prices higher in a phenomenon known as a short squeeze.

While fears of a short squeeze may act as a constraint on short sale activity, particularly in the event of manipulative short squeezes by original buyers who would benefit from inflated prices, the role of short sellers is considered a vital market practice to keep stock prices in-line with fair value.

Yet, the actual occurrence of a squeeze is a debatable subject. One such issue arises from general informed market trading activity which can easily be misconstrued as a short squeeze. As such, there is a clear need to identify specific characteristics to isolate their existence. However, many differing definitions are utilized in practice, forcing the need for a systematic identification process.

Thus, we outline the following steps in our definition of a short squeeze to systematically isolate their occurrence:

Pre-squeeze - Filter out securities ranked in the bottom 40% of Demand Supply Ratio or Implied Loan Rate. These factors are primarily used to identify securities that are heavily shorted. Demand Supply Ratio categorizes stocks that are heavily borrowed in the market relative to the lendable inventory of that stock and Implied Loan Rate measures the cost of borrowing which is indicative of the shorting flow. Stocks are ranked in a percentile form (1-100) on a relative basis by universe. Those securities having the best (worst) factor scores are assigned a 1 (100). Therefore, the closer a rank is to 1 (100), the more (less) prominent is the investment style for that stock.

For robustness, we also apply IHS Markit Securities Finance's proprietary algorithm to filter out positions associated with corporate actions. One well-documented bias in securities lending data is related to dividend activity. Raw securities lending information is affected by this phenomenon and we take special care to remove any bias. The execution of such a transaction ultimately results in a gradual increase in the demand (and cost) to borrow a stock around the dividend record date as firms hedge the associated market risk. This clouds the ability to detect negative sentiment around company prospects. For example, it is prevalent in European stocks as taxation policy there is highly fragmented. For developed Europe and Pacific regions, we apply a further adjustment to normalize the securities lending transactions over 10 business days before and after ex-dividend dates by using the aggregated securities lending count from the 10th day prior to the ex-dividend date over the next 20 days.

Short Squeeze - Identify situations where a stock's price increases significantly (i.e., a 3-standard-deviation move relative to the prior 60 trading days) over a 3-day period as we know that a surge in price could be staggered and last for a few days depending on the news announcement and the degree of positive sentiment. Also, stock recalls are settled in the same way as stock purchases (i.e., borrowers have 3 business days to return the stock).

Post-squeeze - Include only securities that are heavily shorted and the potential squeeze event is followed by a decrease in the shares on loan. In our definition of a short squeeze, we identify securities that have had a recent price surge and are heavily shorted. This should be accompanied by a reduction in stock loan quantity over five consecutive days. One could argue that returns of securities could reflect price manipulation that is characterized by large positive abnormal returns in the absence of any news announcements; however, we reduce the possibility of including any price manipulations in this step.
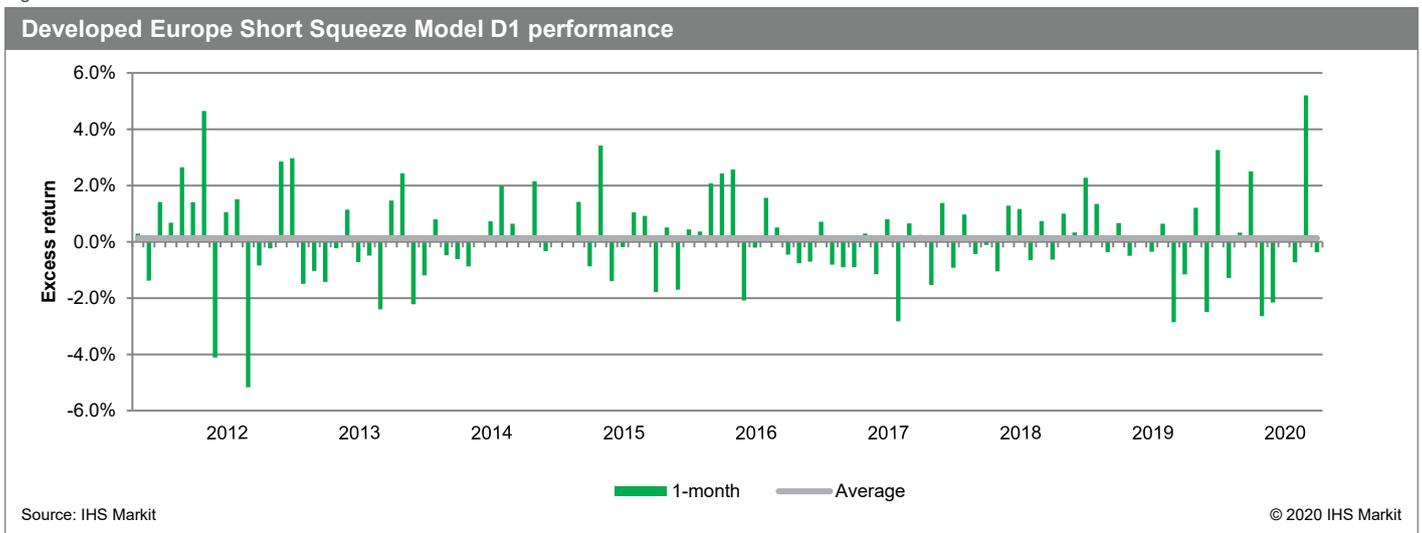
Figures

Figure A1

Developed Europe Short Squeeze Model D1 performance

Source: IHS Markit

© 2020 IHS Markit

Figure A2

Developed Pacific Short Squeeze Model D1 performance

Source: IHS Markit

© 2020 IHS Markit

Figure A3

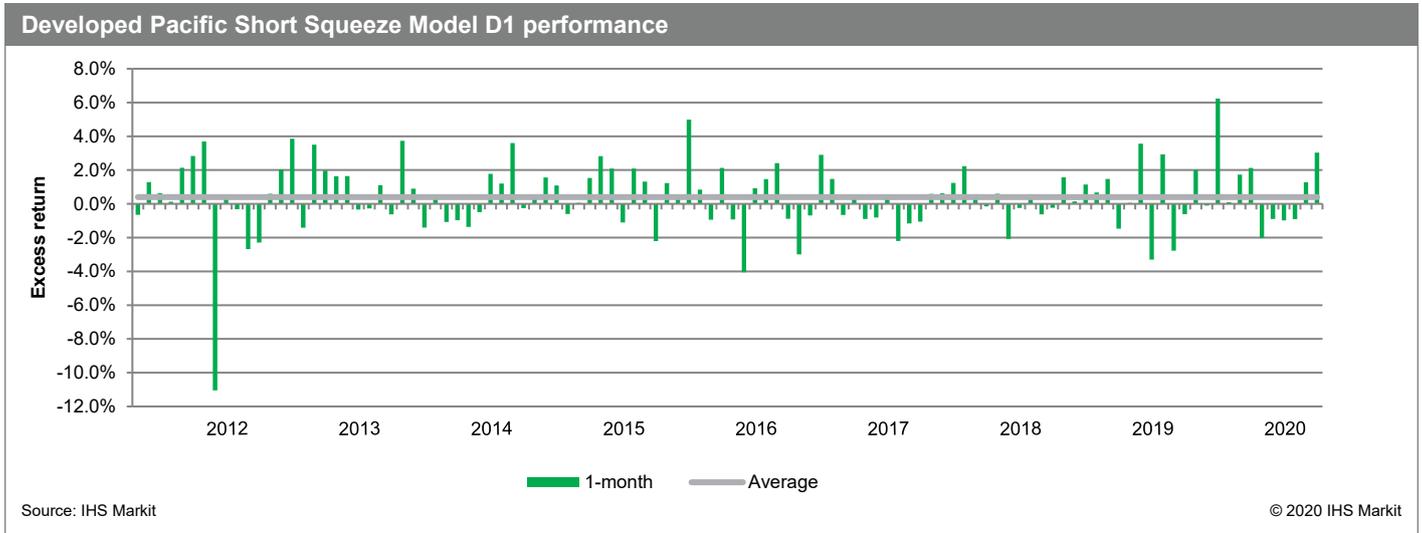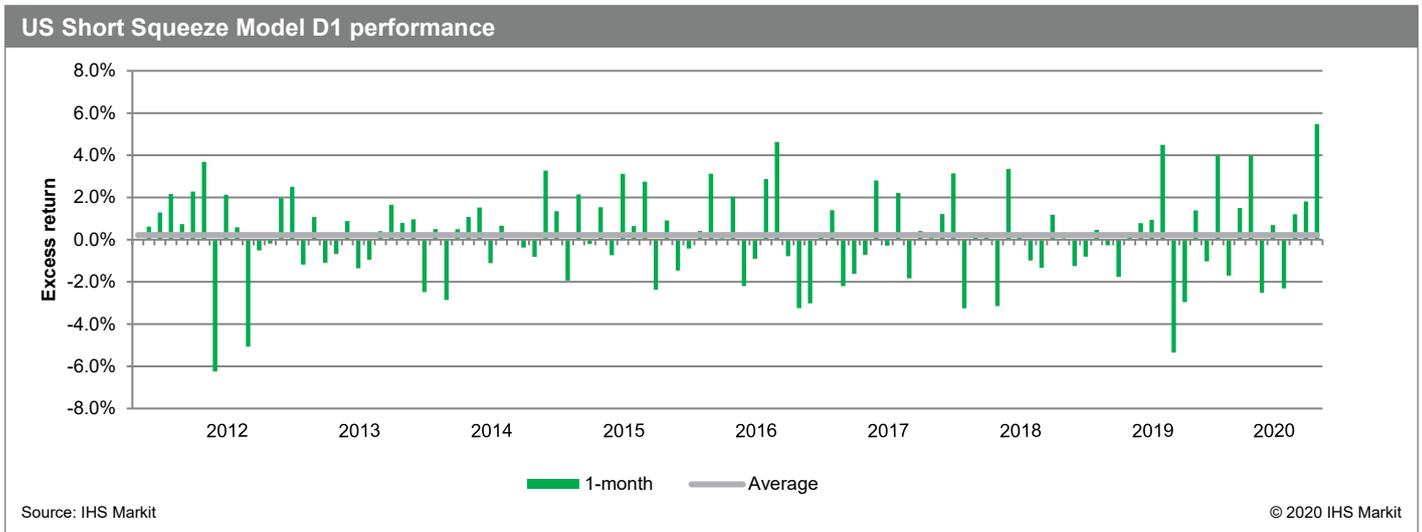## US Short Squeeze Model D1 performance

Source: IHS Markit

© 2020 IHS Markit

IHS Markit Customer Support:

Support@ihsmarkit.com
Americas: +1 877 762 7548
Europe, Middle East, and Africa: 00800 6275 4800
Asia and the Pacific Rim: +65 6922 4210