realindex INVESTMENTS

Part 2 – Conference call tone

Wang Chun Wei, Realindex Investments

For professional use only   December 2022

Introduction

Individuals use language to convey meaning. Beyond the basic definitions of words and sentences, we rely on expression for the full picture of what is being said. In other words, how we say something can be just as important as what we say.

If we are confident of our knowledge and the likely response of our audience, we speak with more confidence and clarity. If we are not, we can be hesitant, long winded, confusing and perhaps negative. That is - the way we form our speech in terms of the choice of words used and the tone of them can change markedly.

This maps very directly into the communication used by senior corporate figures in their regular conference calls on earnings results. If we can extract components of their speech, we can potentially infer good or bad news about the future prospects of the firm. These calls have the added benefit of being captured as recordings and come with accurate written transcripts. It is to these transcripts that we apply our technology.

The first paper in this series looked at the physical content of the transcripts – length of speeches, use of complicated words, and so on. Here, in part two of our series, we attempt to measure the *linguistic tone* or sentiment of the speech – that is, whether words used sound positive, neutral or negative.

Over time, conference calls have become an important channel for voluntary corporate disclosure. The interaction between analysts and company management, which is in part unscripted and spontaneous, provides an opportunity for investors to extract information that may not be fully incorporated into prices.

Linguistic tone

What is linguistic tone?

Linguistic tone is the sentiment of the speaker extracted from the transcript of the call. Firstly, the transcript text is extracted and stop words[1] are removed. To assess the importance of words, we also need to know how frequently they are used – more frequently used words should be downweighted in some way as they are less significant than rarer words. Infrequent words can often be more informative.

The simplest method of estimating the linguistic tone is by calculating the percentage of positive words versus negative words after term-frequency adjustment. This so-called "bag-of-words" approach is simple yet often effective in estimating sentiment. Future work will extend on this, but most of the initial benefit of measuring tone can be achieved through this simple approach.

There has been substantial empirical research on the role of linguistic tone in asset pricing. For instance:

- Price et al. (2012) find that conference call tone was a significant predictor of abnormal returns and trading volume. They also find that the information content of conference calls is also more significant for firms that do not pay dividends (i.e. firms with higher levels of cashflow uncertainty).

- Borochin et al. (2018) find that abnormally negative conference call tones impacts firm value uncertainty.

- Furthermore, Gillam et al. (2017) find that text-based sentiment measures from the Q&A section has positive predictive power for future returns.

As discussed in part 1 of this series (on linguistic complexity and obfuscation), we believe that analysing conference calls may be better than looking at 10-Q

[1] Stop words are words which are filtered out in the processing of text data as they are largely void of meaning, i.e. words such as determiners (eg. words like "the" and "a").

(annual) / 10-K (quarterly) reports as they are less constrained with formality and boilerplate text. This view is shared by:

- Brown and Tucker (2011) who show the year-on-year changes in the information content of the MD&A[2] section within the 10-K have been declining over time

- Bloomfield (2008) who highlights that the unscripted nature of the Q&A section allows for better examination of the information asymmetry between management and analysts.

'Off-the-cuff' tone by managers can be particularly informative, as they may reveal the underlying sentiment of the speaker. Indeed, our empirical analysis shows that for North American firms, Q&A tone is more effective than tone on management's prepared remarks in terms of predicting future returns.

Often, research on linguistic tone in conference call transcripts does not differentiate between manager and analyst remarks. We have already observed there are linguistic style differences between management-speaks versus analysts-speak. Therefore in our analysis, we differentiate between three sections: (a) manager prepared remarks, (b) analyst questions, and (c) manager responses.

Management behaviour

With regards to management tone, managers can either engage in truthful communication or at the other extreme, opportunistically manipulate language to make firm outlook sound better than it is. The reality is likely somewhere between, but nuances of language use in these different situations leave hints that we are able to extract quantitatively.

From our previous paper, we already know that managers in under-performing firms are more likely to employ linguistic obfuscation in their speech. This could be done in a bid to delay the release of bad news, otherwise known as bad-news hoarding. However, it is not clear whether they would actively manipulate tone.

To speak in a "round-about" way (i.e., obfuscate) when earnings outlook is poor is one thing, but to actively use positive tone when the situation is negative is another thing entirely. Reputational risk should be a major driver. Therefore, when there is bad news, we hypothesize that: *managers will make admissions, but also spend considerable time justifying the outcome*. Further, admissions like these are more likely in the *unscripted* sections than in the prepared sections of the call.

We believe managers are generally truthful when it comes to tone and sentiment. Therefore, positive (negative) tone may relate to future positive (negative) returns.

A further distinction can be made between management tones in their prepared remarks compared with their more "off the cuff" responses in the answers section. The latter is more spontaneous, and subsequently we expect the tone of the latter section to be more informative. With regards to analyst questions, we believe there is less incentive for analysts to engage in manipulative behaviour when it comes to sentiment or tone. Their tone is more likely to be transparent and to mirror their view of the firm, whether positive or negative.

The word dictionary

In order to determine whether a word is positive or negative, we employ the Loughran-MacDonald word dictionary[3] which is tailored for financial text. This is a standard methodology for looking at sentiment in conference calls, and at 10-K and 10-Q reports in the academic literature.

We test a battery of different sentiment measures based on the Loughran-MacDonald word dictionary on the three distinct parts of conference call transcripts (management discussion, questions by analysts and answers by managers). Positive and negative words from this dictionary can be counted.

Tone within a section is then calculated as:

---

[2] Management Discussion and Analysis

[3] See for example https://sraf.nd.edu/loughranmcdonald-master-dictionary/

$$TONE = \frac{num \ of \ positive \ words - num \ of \ negative \ words}{total \ words}$$

An extension is to incorporate what are known as *term frequency inverse document frequency (TF-IDF)* weights. This scales by word frequency across documents – less frequent words in more documents get a higher weight. We can do this by utilizing data from the Loughran-MacDonald master dictionary which provides word count and document frequency. Overall, we did not find substantial differences when using term frequencies for our long/short spread portfolios.

Describing conference calls

The Factset Conference Call Transcript data covers North America, Europe and Australia. The dataset begins around 2008, however, it is only in the last 5 years that we see MSCI ACWI coverage above 80%. Overall, we have transcripts for conference calls for more than 1,400 firms covered.

We then construct the tone measure above and in the charts below we summarise the results for the three sections (manager prepared remarks, analyst questions, and manager responses) by region. We see some noticeable differences. For example, in the Q&A sections, we note that North American tone to be more upbeat than other regions Tone is lower in Japan in all sections. Australian tone in the Manager Prepared section is about the highest. North American management responses are relentlessly upbeat.
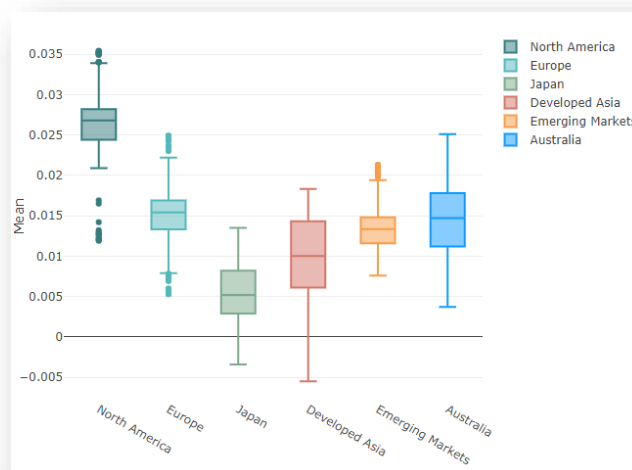
Chart 1. Global Sentiment

- *Panel A. Management Prepared Speech Tone*
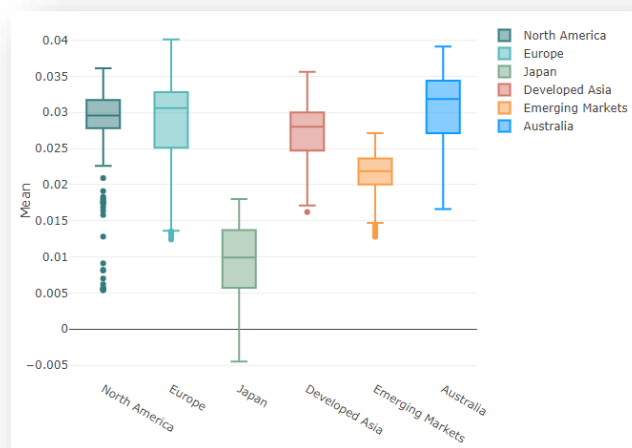
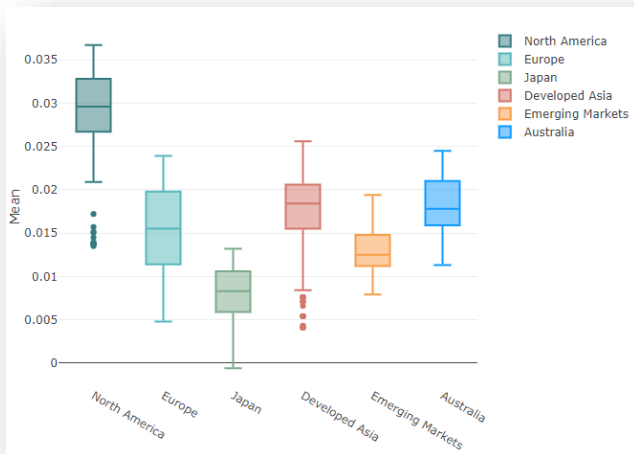*Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021*

- *Panel B. Analyst Question Tone*

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

- *Panel C. Management Answer Tone*

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

Why is this interesting? Tone as a leading indicator

If our supposition is correct, measuring the tone of conference calls in the way we have described will, we believe, give us some indication of the future of the firm, as seen by the management. If this insight is valuable, it should be reflected in firm outcomes and in future returns. That is – can conference call tone be used as an alpha signal?

North America

In Chart 2, we plot the equally-weighted long short spread performance (that is, top quintile return less bottom quintile return) of linguistic tone in North America. Linguistic tone is broken into three distinct sections:

- Question Tone: linguistic tone of analyst questions

- Answer Tone: linguistic tone of senior management answers to analyst questions

- Management Tone: linguistic tone of senior management's prepared remarks of company earnings prior to the Q&A section

We find that the performance of 'question tone' and 'answer tone' is highly correlated. We believe this makes sense, as the answers need to relate to the

topics raised by the questions. That is, in good news situations, analyst questions are likely to direct management towards discussing positive aspects of the results, and conversely when results are weak the analysts are more likely to direct their attention towards problems.

Of the three tones, management tone is based on a more scripted part of the transcript, as managers often would have prepared this beforehand. Question and answer tone should be more spontaneous, and potentially more informative. Our results suggest scripted speech gives away less information on sentiment for North American firms.

Chart 2. North American Long/Short Spread Performance

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

Europe ex UK

In Chart 3 below, we repeat the plot of equally-weighted long short spread performance (quintiles) of linguistic tone, but now in Europe (ex UK). We separate UK from EU here as often firms in the former hold semi-annual conference calls whilst the latter, like the US, is quarterly frequency. We find the results in Europe to be very promising. Consistent with North America, we find that the linguistic tone of analyst questions are a good predictor of future returns. We find the spread returns of the management and answer sections to be also correlated, as they are spoken by the same group of people.

Chart 3. European Long/Short Spread Performance

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

UK

In Chart 4, we broadly see a similar picture, however, given the lack of breadth in the UK, we find higher levels of volatility in spread portfolio returns.

Chart 4. UK Long/Short Spread Performance

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

Emerging Markets

As discussed in our first paper on this topic, the coverage of Factset transcripts in Emerging Markets is rather poor: around 30%, with about 350 names. Therefore, the results are not particularly meaningful. Regardless, we still see mildly positive results which is reassuring.

Chart 5. Emerging Markets Long/Short Spread Performance

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

Statistical summary

In Table 1 below, we document the same equally-weighted spread portfolios performance statistics, along with the Fama-MacBeth panel regression T-stats after controlling for a variety of other factors: momentum, quality, value, earnings revisions and sectors. The results vary by region, and emerging market results should be largely ignored due to data scarcity. We find results in Europe to be most compelling (across all sections of the conference call), along with mildly positive results in the Q&A section of North American and UK conference calls.

Table 1. Conference Call Tone Long/Short Spread Performance

Signal	Mean	Stdev	IR	Turnover	T-stat	Hit Rate	FM Mean	FM Stdev	FM T-stat
North American Tone									
Question	1.63%	4.75%	0.34	245%	1.19	54.50%	0.003	1.29%	0.9
Answer	1.85%	4.61%	0.4	220%	1.39	51.40%	0.005	1.23%	1.33
Management Prepared Remarks	0.62%	4.79%	0.13	198%	0.45	54.20%	0.001	1.33%	0.28
Europe ex UK Tone									
Question	4.33%	5.17%	0.84	250%	2.89	55.10%	0.004	1.40%	0.96
Answer	2.91%	5.92%	0.49	225%	1.7	55.10%	0.004	1.62%	0.97
Management Prepared Remarks	3.96%	5.94%	0.67	199%	2.3	54.50%	0.007	1.56%	1.53
UK Tone									
Question	2.65%	8.57%	0.31	218%	1.07	54.70%	0.006	2.67%	0.78
Answer	5.73%	9.37%	0.61	192%	2.12	56.90%	0.017	2.73%	2.19
Management Prepared Remarks	4.74%	9.71%	0.49	171%	1.69	54.20%	0.008	2.94%	0.95
Emerging Markets Tone									
Question	3.22%	7.36%	0.44	258%	1.51	51.90%	0.002	2.38%	0.3
Answer	2.52%	8.33%	0.3	249%	1.05	53.20%	0.001	2.59%	0.15
Management Prepared Remarks	2.12%	7.96%	0.27	230%	0.92	52.30%	-0.001	2.49%	-0.07

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

LM's own sentiment measures

The Loughran-MacDonald words lists can be examined by their own internal classifications, rather than by Tone as we have defined it. These are not limited to just good ("positive") and bad ("negative") words. There are many other classifications which are interesting. These include:

- Uncertainty - 297 words incl. words such as 'may', 'approximately', 'risks', 'pending', 'uncertainty' ...

- Litigious - 905 words incl. words such as 'attorney', 'appeal', 'indemnity', 'unlawful' ...

- Strong - 19 words incl. words such as 'always', 'best', 'definitely', 'never' ...

- Weak - 27 words incl. words such as 'may', 'could', 'sometimes', 'somewhat' ...

- Constraining 184 words incl. words such as 'required', 'comply', 'committed', 'abide' ...

We repeat our study above using these categories, and we also break down Tone, as analysed above, into its positive and negative components. These signals have been sector normalized.

Overall, we do not find anything particularly meaningful in the prepared remarks section. In other words,

prepared speeches in North America, do not give away much in terms of information. Negative words by analysts seems to be the most telling, whilst positive words, uncertain words and litigious words were all very informative in the management answers section of the Q&A.

Table 2. Other Sentiment Measures Long/Short Spread Performance (North America)

Signal	Mean	Stdev	IR	Turnover	T-stat	Hit Rate	FM Mean	FM Stdev	FM T-stat
Questions									
Positive	-0.15%	4.02%	-4.00%	243.90%	-0.13	53.00%	-0.001	1.26%	-0.32
Negative	-2.98%	5.28%	-56.00%	250.30%	-1.95	45.00%	-0.007	1.30%	-1.89
Uncertainty	-0.07%	3.58%	-2.00%	257.60%	-0.07	48.10%	0.003	1.10%	1.07
Litigious	-1.22%	4.56%	-27.00%	285.60%	-0.93	45.80%	-0.002	1.26%	-0.47
Strong	1.05%	4.76%	22.00%	382.90%	0.76	49.00%	-0.004	1.04%	-1.43
Weak	-0.26%	3.42%	-8.00%	250.10%	-0.27	51.90%	0.002	1.10%	0.58
Constraining	2.34%	5.04%	46.00%	368.90%	1.6	53.20%	0.003	1.26%	0.92
Answers									
Positive	1.72%	3.83%	45.00%	221.10%	1.56	50.80%	0.005	1.12%	1.71
Negative	-1.54%	5.17%	-30.00%	229.00%	-1.03	46.90%	-0.001	1.35%	-0.33
Uncertainty	-1.37%	4.15%	-33.00%	230.40%	-1.14	46.50%	-0.005	1.42%	-1.26
Litigious	-1.41%	4.17%	-34.00%	283.40%	-1.17	47.60%	-0.006	1.36%	-1.47
Strong	-1.27%	3.81%	-33.00%	255.50%	-1.15	50.60%	0	1.02%	-0.06
Weak	-0.17%	3.49%	-5.00%	240.30%	-0.16	49.50%	-0.003	1.13%	-0.93
Constraining	0.64%	3.92%	16.00%	253.00%	0.56	49.80%	0.002	1.24%	0.58
Management Prepared Remarks									
Positive	0.09%	4.19%	0.02	188.10%	0.07	52.60%	0	1.21%	0.06
Negative	-1.63%	6.07%	-0.27	201.00%	-0.93	43.70%	-0.003	1.51%	-0.73
Uncertainty	-0.44%	3.37%	-0.13	184.40%	-0.46	50.00%	-0.003	1.12%	-1.02
Litigious	-1.10%	3.77%	-0.29	215.80%	-1.01	50.60%	-0.001	1.31%	-0.24
Strong	-2.07%	5.28%	-0.39	365.70%	-1.35	46.30%	-0.004	0.98%	-1.25
Weak	0.44%	3.78%	0.12	280.60%	0.4	51.90%	0.002	0.89%	0.73
Constraining	-0.95%	4.15%	-0.23	219.10%	-0.79	47.90%	-0.002	1.22%	-0.61

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

In Europe the prepared management discussion section is more straightforward. For instance, positive tone is associated with strong positive T-stat. Similarly, litigious and constraining words is associated with future negative returns. Consistent with North America, negative sentiment by analyst is the strong indicator, whilst positive and litigious words by management responses were also significant.

Table 3. Other Sentiment Measures Long/Short Spread Performance (Europe)

Signal	Mean	Stdev	IR	Turnover	T-stat	Hit Rate	FM Mean	FM Stdev	FM T-stat
Questions									
Positive	2.85%	4.44%	0.64	243.50%	2.22	56.10%	0.002	1.39%	0.45
Negative	-3.69%	5.59%	-0.66	247.30%	-2.28	46.60%	-0.005	1.35%	-1.32
Uncertainty	-0.84%	5.05%	-0.17	248.00%	-0.57	49.80%	0	1.50%	-0.06
Litigious	-1.19%	5.31%	-0.22	312.20%	-0.77	47.80%	-0.006	1.70%	-1.17
Strong	0.55%	5.10%	0.11	374.30%	0.37	51.40%	-0.002	1.38%	-0.37
Weak	0.32%	4.52%	0.07	239.50%	0.25	50.80%	-0.003	1.38%	-0.63
Constraining	1.63%	4.95%	0.33	362.30%	1.14	50.20%	0.006	1.48%	1.49
Answers									
Positive	2.14%	5.42%	0.4	222.60%	1.37	54.30%	0.004	1.63%	0.87
Negative	-1.99%	5.92%	-0.34	225.90%	-1.16	48.20%	-0.002	1.50%	-0.39
Uncertainty	0.83%	4.41%	0.19	235.70%	0.65	50.60%	0.004	1.52%	0.81
Litigious	-3.71%	5.07%	-0.73	238.70%	-2.53	44.20%	-0.012	1.67%	-2.46
Strong	-0.62%	4.29%	-0.14	238.10%	-0.5	50.00%	-0.002	1.33%	-0.6
Weak	0.36%	4.77%	0.08	234.40%	0.26	50.80%	0.004	1.45%	0.88
Constraining	-1.35%	4.58%	-0.29	249.30%	-1.02	45.50%	-0.002	1.48%	-0.38
Management Prepared Remarks									
Positive	5.04%	5.56%	0.91	192.30%	3.14	55.80%	0.009	1.56%	1.99
Negative	-2.96%	6.39%	-0.46	199.90%	-1.6	43.40%	0	1.56%	-0.05
Uncertainty	-0.87%	5.07%	-0.17	214.40%	-0.59	47.60%	-0.002	1.69%	-0.46
Litigious	-2.77%	5.47%	-0.51	231.50%	-1.75	47.90%	-0.011	1.75%	-2.11
Strong	-0.33%	4.05%	-0.08	269.80%	-0.28	48.90%	0	1.26%	-0.06
Weak	0.22%	4.71%	0.05	264.40%	0.16	49.00%	0	1.53%	0.02
Constraining	-4.18%	5.81%	-0.72	233.00%	-2.49	43.60%	-0.01	1.56%	-2.14

Source: Realindex, Factset   Date range: Jan 2010 to Dec 2021

Overall, we find these other sentiment measures tagged by the Loughran-MacDonald master dictionary to be too noisy to be implemented as straightforward signals. Instead our preference is for the original tone metric described earlier.

We do however find some interesting asymmetric characteristics. For example, negative words by analysts are more meaningful than positive words, i.e., we should be more aware of analyst concerns than analyst praises. Conversely, positive tone by managers in response to analysts is a more potent signal than negative tone.

Conclusions

This paper describes how conference call tone can be used as an alpha signal. These signals are robust, easy to calculate and have good intuitive sense. Capturing the speech via textual analysis – both in content and in tone – gives us insight into the strength of conviction or uncertainty of management and

analyst community regarding the future prospects of the firm.

In particular we find that:

- Whilst managers may obfuscate (as discussed in our previous paper), they are generally truthful when it comes to sentiment. This means positive (negative) tone relates to future positive (negative) returns.
- In North America, tone works best in the unscripted Q&A section. And in particular, negative tone by analysts is a strong signal.
- Tone is particularly strong in Europe, and also works in the management discussion section

One last point: application of these ideas – and extensions on them – to Realindex' investment process is both additive and diversifying.

Reference

Bloomfield (2008) Discussion of annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics, 45, 248-252.

Borochin, P. Cicon, J, DeLisle, J. and S. Price (2016) The Effects of Conference Call Tones on Market Perceptions of Value Uncertainty, Journal of Financial Markets

Brown and Tucker (2011) Large-Sample Evidence on Firms' Year-over-Year MD&A Modifications, Journal of Accounting Research, Vol 49 (2), 309 - 346

Gillam, R., Guerard, J. and R. Cahan (2017) Getting Sentimental: Conference Call Sentiment and Stock Returns, McKinley Capital

Li, F. (2006) Do Stock Market Investors Understand the Risk Sentiment of Corporate Annual Reports?, SSRN, https://ssrn.com/abstract=898181

Price, S. Peterson D. and B. Bliss (2012) Earnings conference calls and stock returns: The incremental informativeness of textual tone, Journal of Banking and Finance, Vol 36 (4), 992-1011

Important Information

This material is for general information purposes only. It does not constitute investment or financial advice and does not take into account any specific investment objectives, financial situation or needs. This is not an offer to provide asset management services, is not a recommendation or an offer or solicitation to buy, hold or sell any security or to execute any agreement for portfolio management or investment advisory services and this material has not been prepared in connection with any such offer. Before making any investment decision you should consider, with the assistance of a financial advisor, your individual investment needs, objectives and financial situation. We have taken reasonable care to ensure that this material is accurate, current, and complete and fit for its intended purpose and audience as at the date of publication. No assurance is given or liability accepted regarding the accuracy, validity or completeness of this material and we do not undertake to update it in future if circumstances change. To the extent this material contains any expression of opinion or forward-looking statements, such opinions and statements are based on assumptions, matters and sources believed to be true and reliable at the time of publication only. This material reflects the views of the individual writers only. Those views may change, may not prove to be valid and may not reflect the views of everyone at First Sentier Investors.

About First Sentier Investors

References to 'we', 'us' or 'our' are references to First Sentier Investors, a global asset management business which is ultimately owned by Mitsubishi UFJ Financial Group. Certain of our investment teams operate under the trading names FSSA Investment Managers, Stewart Investors and Realindex Investments, all of which are part of the First Sentier Investors group. We communicate and conduct business through different legal entities in different locations. This material is communicated in:

- **European Economic Area** by First Sentier Investors (Ireland) Limited, authorised and regulated in Ireland by the Central Bank of Ireland (CBI reg no. C182306; reg office 70 Sir John Rogerson's Quay, Dublin 2, Ireland; reg company no. 629188)
- **United Kingdom** by First Sentier Investors (UK) Funds Limited, authorised and regulated by the Financial Conduct Authority (reg. no. 2294743; reg office Finsbury Circus House, 15 Finsbury Circus, London EC2M 7EB)
- **Other jurisdictions,** where this document may lawfully be issued, by First Sentier Investors International IM Limited, authorised and regulated in the UK by the Financial Conduct Authority (FCA ref no. 122512; Registered office: 23 St. Andrew Square, Edinburgh, EH2 1BB; Company no. SC079063).

To the extent permitted by law, MUFG and its subsidiaries are not liable for any loss or damage as a result of reliance on any statement or information contained in this document. Neither MUFG nor any of its subsidiaries guarantee the performance of any investment products referred to in this document or the repayment of capital. Any investments referred to are not deposits or other liabilities of MUFG or its subsidiaries, and are subject to investment risk, including loss of income and capital invested. © First Sentier Investors Group