Invesco

#04

4th issue 2020

Long live the digital revolution

NLP in portfolio management: an innovative approach

Searching for inner peace with value factors

Invesco's guide to alternative data

Investing through a macro factor lens

Risk & Reward

Research and investment strategies

This magazine is not intended for members of the public or retail investors. Full audience information is available inside the front cover.

Demographic and social change, the need for more sustainability and technological evolution – these three megatrends are shaping our world, and Covid-19 is likely to accelerate these trends further. In this issue of Risk & Reward, we discuss the impact of these forces and of the digital revolution on the world – and on asset management in particular.

At Invesco, we are not just observers of change. We work to drive change ourselves through state-of-the-art research. True to this commitment, four of my colleagues have developed a new method to identify promising stocks with the help of Natural Language Processing (NLP) – a truly digital approach to discovering innovative companies based on automated analysis of millions of text files.

Other articles in this issue deal with factor investing, our time-tested method of investment selection and an innovative concept in its own right.

We analyze the recent weakness of value stocks and examine what has happened to the value effect, one of the most persistent market anomalies of recent decades. How should investors react to this puzzling trend?

We also discuss the use of less structured alternative data in factor investing. What can investors look for when they decide to search beyond traditional financial data?

Finally, we consider the macro factors that factor investors can use to position their portfolios for likely macroeconomic developments as an effective complement to style preferences.

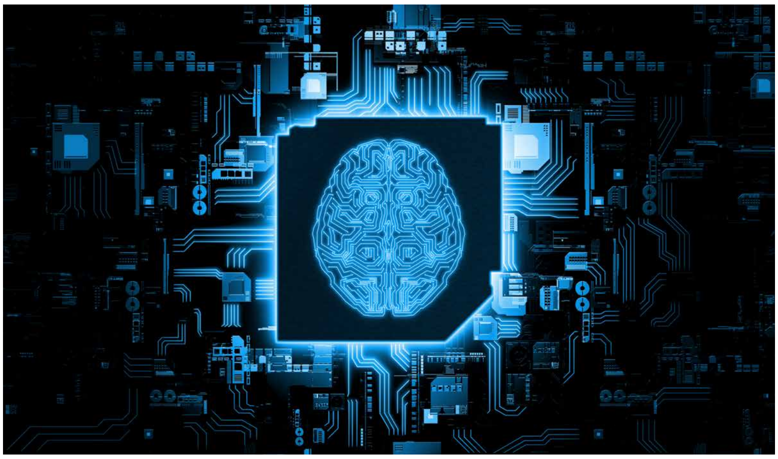
We hope you enjoy this edition of Risk and Reward.

Best regards,

Marty Flanagan
President and CEO of Invesco Ltd.

Contents

Feature



In focus

Long live the digital revolution

By Georg Elsaesser, Dr. Martin Kolrep, Alexandar Cherkezov and Michael Rosentritt

In brief

Digitalization and the accelerating pace of innovation are bringing about tremendous changes, not only in how we communicate, but also how people invest. We highlight some of the key developments, discuss the importance of innovations and show how new technologies on the near horizon are transforming the world – including the fund management industry.

With the arrival of the internet and smartphones, the vast processing power of personal computers and improvements in data storage, the past decades have seen significant innovations and life-changing progress. Together with, e.g. sustainability challenges and the aging population, this is likely to lead to even more and faster innovation in different fields. We discuss these developments and assess their significance for the fund management industry.

Anyone born around 1970 remembers when the first broadly available video game systems arrived around the end of the 1980s. Compared to the quality of computer games kids are playing today, it is easy to see just how far computing power has come – those early video games and gaming consoles might as well be from the stone age.

Over the last 30-40 years, we have lived through the progression from the Walkman to the mp3 player, and the arrival of the cell phone.

Over the last 30-40 years, we have lived through the progression from the Walkman to the mp3 player, and the arrival of the cell phone – all of which have now been replaced once more, by smartphones and streaming services. The speed of technological innovation has accelerated tremendously compared to 100 years ago. For example, social networks like Facebook, Twitter or WeChat did not even exist just 20 years ago. Today, these platforms have revolutionized how we access information and how we communicate with our friends or family. They are deeply integrated into our daily lives.

Of course, these platforms have only become possible due to the invention of the World Wide Web and the internet. Over the years, a significant increase in bandwidth made it possible to distribute information around the globe with almost no delay. Remember that around the beginning of the 1990s, there was only a handful of websites online. The main limitation was low bandwidth, and very few people had the technology available to access these websites. However, the underlying technology, the TCP/IP protocol, was actually invented around the year 1970 – roughly 20 years before the first website was available to the public.

Ultimately smartphones and tablets were a complete game changer, and significantly boosted the exchange of information through the internet on various platforms. Novel technologies and the improved speed of data transfers with the genesis of 5G mobile internet now promise to further expand the playing field. For the gaming industry alone, augmented reality and virtual reality will make video games ultra-realistic and accessible to the mainstream. And this is just one of a multitude of new horizons opening up as computers become even smaller and faster. The future will be shaped by the synergies between the myriad innovations that are now becoming market-ready and scalable so that almost every industry will be reimagined.

Innovation builds on invention

In this context, it is good to have a website like Wikipedia for a solid definition of the term innovation: *"An innovation is typically a new idea, creative thoughts, new imaginations in the form of a device or method. (…) Innovation is often also viewed as the application of better solutions that meet new requirements, unarticulated needs or existing market needs. (…) Such innovation takes place through the provision of more-effective products, processes, services, technologies or business models that are made available to markets, governments and society."*[1]

An invention, on the other hand, is something different. Another website defines "invention" as, "the creation of a product or introduction of a process for the first time."[2] The crucial point here: to be financially successful, one does not need to invent something. Quite often, innovation delivers more added value than inventions as it is the consequence of some invention that happened far earlier.

The TCP/IP example demonstrates this quite impressively. It was the development of smartphones and tablets, which according to the definition we would call innovations, that led to the breakthrough of the internet. This milestone resulted in companies developing sustainable business models and actually earning money. Thus, innovation is not the exclusive domain of private equity firms and venture capitalists – every company should be innovative and help propel the trends and technologies that will shape our future.

We believe that the most successful companies will be those working to evolve or reshape their business models compared to companies that are sticking to their traditional ways of doing business and might miss out on the benefits of innovation. For ages, technological innovations have been the driving force behind growth and productivity. Just recently, we have seen how the world changed with the outbreak of Covid-19 and the broad adoption of even more technology. In a very short time, this has had a massive impact on our society, with families staying closer together and people traveling much less to complete their work. It has been a challenge for many companies, but those that have been able to adapt and bring forward innovation have done very well.

The digital era

The key to all of these trends is digitalization. Over the past 100 years our economy and society have been living through the industrial era. But now, everything is going digital, meaning not only a transformation of the industrial era but a fundamental transition into what we call the "digital era".[3] The changes and the level of innovation we are about to experience are profound, and the potential arising from ever increasing interaction between technologies is exponential. 25 years ago, it was the internet that led to the innovations and developments we see today. We now have technologies like artificial intelligence (AI), 3D printing, Internet of Things, cloud computing, virtual reality, robotics… which were the stuff of science fiction back then.

The interplay and more widespread adoption of these technologies will help generate innovation in other areas, mainly demographics and society, as well as the protection of the environment and resource availability. AI, for example, has given rise to new methods in the development of drugs, drug target identification, drug screening and predictive modeling. Finding new drugs will likely become a much more rapid process than in the past. AI will also be useful in the area of sustainability: for example, when we talk about "smart cities", AI will play a significant role in helping to reduce emissions and manage waste.

And the fund management industry?

The fund management industry, too, has seen its share of changes over the years. When Invesco Quantitative Strategies started to manage money for clients in 1983, there was not a lot of information available in digital form. Computers had started to emerge and become available to more people and businesses. But compared to today's standards, they were very slow and limited. Like the world in general, the changes we have seen over the past 35+ years are best summarized by the following key trends:

1. Everything is becoming digital. Digitalization is unstoppable in the fund management industry as well as in many other industries. More and more information is available in digital form, for example annual reports, earnings call transcripts and all data related to company performance.

2. The amount of data is increasing at an exponential rate, meaning that more and more devices are producing data that is uploaded somewhere and is available for analysis. Everything is monitored, the movement of every vehicle, person, device, etc. If data can be analyzed and certain patterns can be extracted, then data is of course very valuable. Data is therefore named "the currency of the digitalized world".

3. Finally, computers have become so powerful and fast that they are able to analyze huge volumes of data in just seconds. To give an example, a human fund analyst can listen to one earnings call at a time. With the help of a computer and textual analysis methods like Natural Language Processing (NLP), thousands of earnings calls can be searched for certain patterns instantly.

Conclusion

It is obvious that, under these circumstances, the traditional fund manager will be forced to use the available tools to aid research in order to maintain a competitive advantage. The speed at which a computer can analyze and compare companies based on specific datasets is so far beyond what a human being can deliver, that the traditional fund manager is outpaced before even picking up a newspaper or a research report. In our companion article, "NLP: an innovative approach to finding innovative companies", we describe how Invesco makes use of these new possibilities in an innovative quantitative investment strategy.

About the authors

Georg Elsaesser
Senior Portfolio Manager
Invesco Quantitative Strategies
Georg Elsaesser is involved in the management of Invesco Quantitative Strategies equity portfolios and has long-standing experience in the implementation of factor-based investment solutions.

Dr. Martin Kolrep
Senior Portfolio Manager
Invesco Quantitative Strategies
Martin Kolrep is focused on the development and management of multi-asset portfolios and outcome-oriented strategies. As a physicist he has gained important experience in the field of artificial intelligence and machine learning.

Alexandar Cherkezov
Portfolio Manager
Invesco Quantitative Strategies
Alexandar Cherkezov manages multi-asset portfolios that include the elements factor-based investing, active asset allocation and downside risk management.

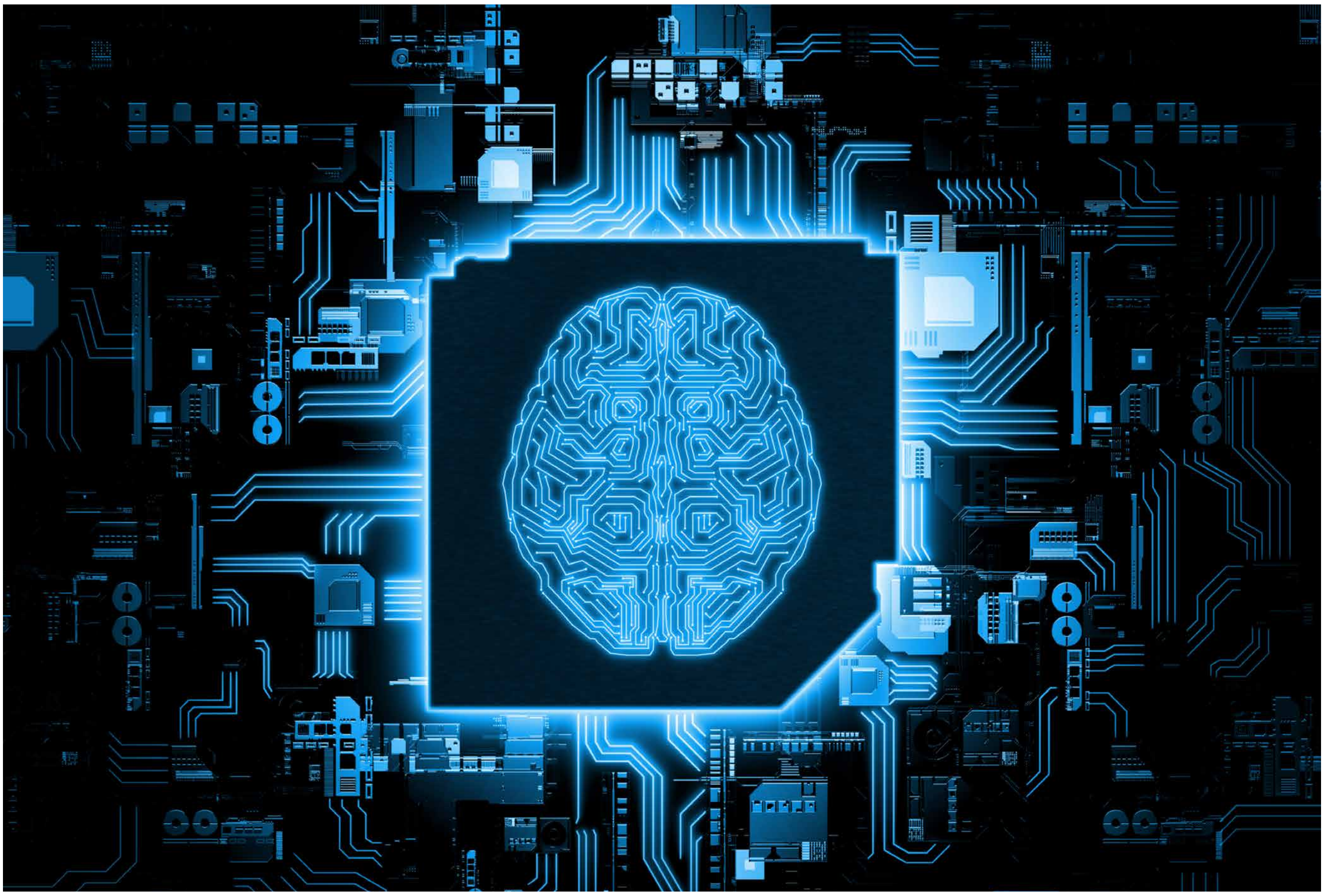
Michael Rosentritt
Portfolio Management Associate
Invesco Quantitative Strategies
Michael Rosentritt is responsible for model and performance analysis of quantitative equity strategies and investment communication.

Notes

1 Wikipedia, downloaded on 19 August 2020.
2 Institute for Economic Thinking, https://www.ineteconomics.org/perspectives/videos/invention-vs-innovation, downloaded on 19 August 2020.
3 O. Stengel et al. (2017): Digitalzeitalter – Digitalgesellschaft, das Ende des Industriezeitalters, Springer Verlag.

NLP in portfolio management: an innovative approach to finding innovative companies

By Georg Elsaesser, Dr. Martin Kolrep, Alexandar Cherkezov and Michael Rosentritt



In brief

Innovations are key to future growth, and they happen in many different areas. But with the abundance of available information, finding tomorrow's leading firms can be challenging. We develop an innovative two-step NLP approach that may help to find companies exposed to the megatrends that shape the future and weight them based on their exposure to these trends and the themes underlying them.

The world will continue to change at a rapid pace, driven by three megatrends: changes in demographics and society, the need for a more sustainable world and evolutions in technology. We show how a Natural Language Processing (NLP) algorithm applied to millions of news datasets can help identify innovative businesses that are likely to benefit from these changes.

We have developed a keyword extraction method to pinpoint innovation-related keywords from documents.

We live in times when the amount of data available, the speed in which the data is produced and the sources from which the data can be obtained have grown tremendously. It is increasingly difficult for humans to keep up with this pace as manual analysis and interpretation of the data is too time consuming. This is where Natural Language Processing (NLP) comes into play, a field of Artificial Intelligence in which computer algorithms analyze, understand and derive meaning from human language and text in an automated way.

In investment management, NLP techniques can be used to support investment decisions through:

- **automatic summarization:** generating concise summaries from vast amounts of text (e.g. research reports);
- **speech recognition:** identifying words and phrases in spoken language (e.g. earnings calls);
- **keyword extraction**;
- **sentiment analysis:** interpreting and classifying emotions within text data (e.g. management sentiment in earnings calls);
- **topic segmentation:** detecting whether different topics are discussed in a given text (e.g. in a longer conversation) and parsing text into relevant segments;
- **named entity recognition:** identifying "named entities" (e.g. people, places, organizations) as key information in a text and classifying them into categories like company, country, time, location etc.

A two-step method

To find innovative companies that may be of interest for a theme portfolio, we have developed a keyword extraction method to pinpoint innovation-related keywords from documents, like academic or broker research. Furthermore, we use NLP algorithms on an ongoing basis to scan millions of news reports and detect which companies are mentioned together with our keywords.

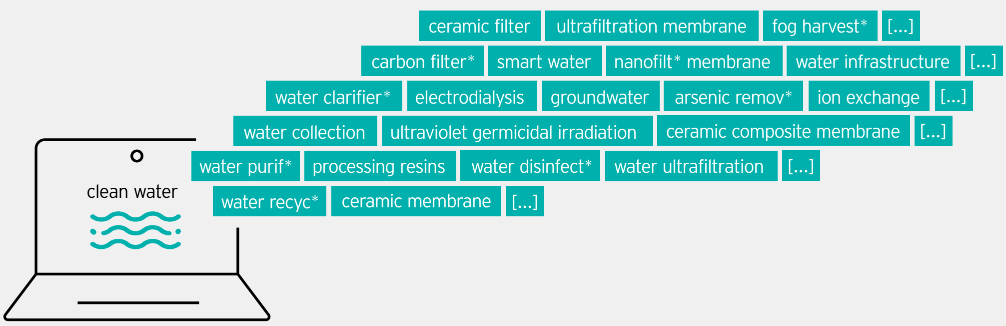
We start with the systematic analysis of innovation-related academic papers, news data, financial research, think tank and futurist publications, thematic websites etc. using NLP algorithms to identify investment themes underlying three key megatrends: changes in demographics and society, the need for a more sustainable world and evolutions in technology. We have chosen these trends because they are likely to be most relevant for the coming decades.

An investment theme can be any subject, topic or focus that breaks a megatrend down into investable topics. Basically, it is a dictionary of associated, innovation-related keywords extracted from the documents. For example, an extract from a theme dictionary for "clean water" could be: ceramic composite membrane, carbon filter*, arsenic remov*, ceramic filter, ceramic membrane, electrodialysis, groundwater, fog harvest*, ion exchange, processing resins, nanofilt* membrane, smart water, ultrafiltration membrane, ultraviolet germicidal irradiation, water clarifier*, water disinfect*, water purif*, water recyc*, water collection, water infrastructure, water ultrafiltration.

By continuously analyzing a broad set of innovation-related documents, a new investment theme is identified and captured as soon as it emerges. The outcome of this first step is a set of relevant investment themes (currently 17) that are shown in figure 2.

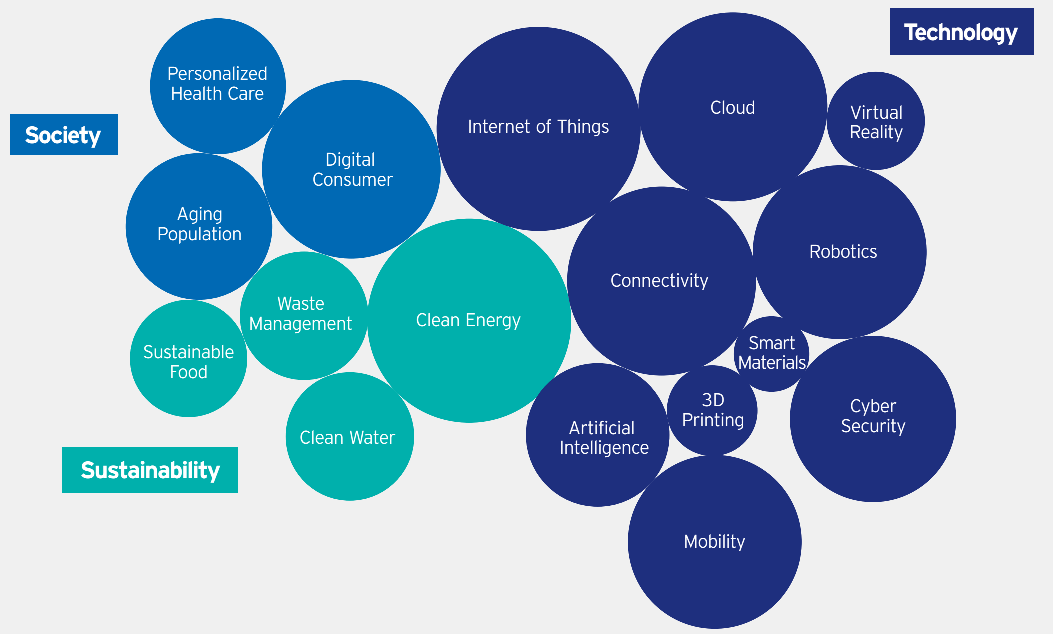
In a second step, we use NLP algorithms and the theme dictionaries to scan millions of news reports

Figure 1
Example of a dictionary for the theme "Clean Water" (extract)



Source: Invesco. For illustrative purposes only.

Figure 2
Investment themes and their grouping into the three megatrends

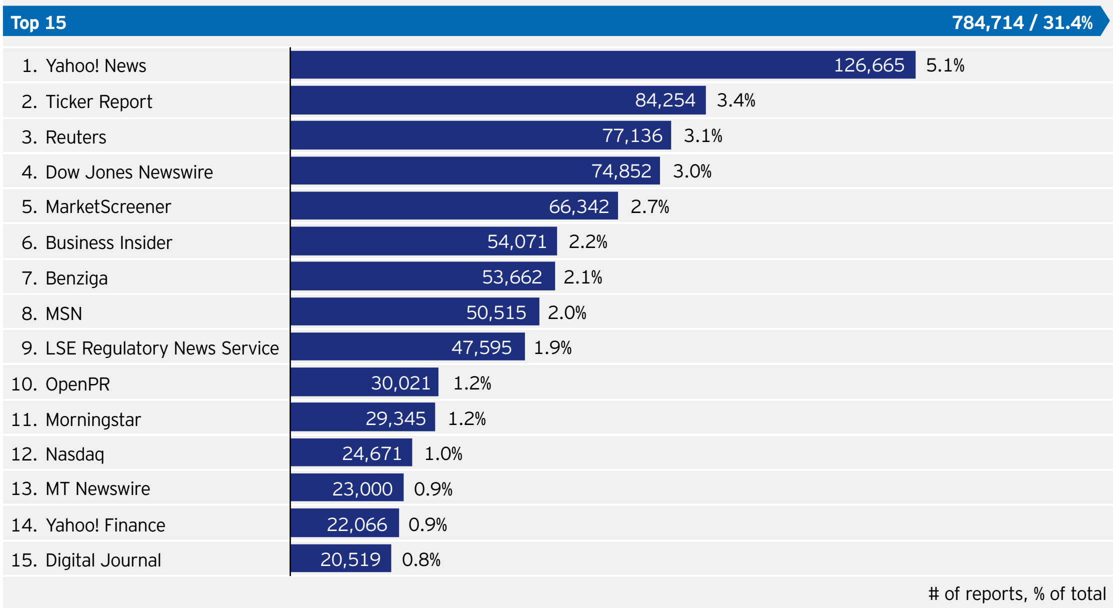


Source: Invesco. For illustrative purposes only.

with the goal of identifying the companies with the highest exposure to, as well as the highest relevance for, the investment themes. The news comes from up to 4,500 news channels and sources.[1] The reports are company-related, but do not typically come from the companies themselves. Figure 3 shows the number of news items and the percentage shares of different sources in April 2020.

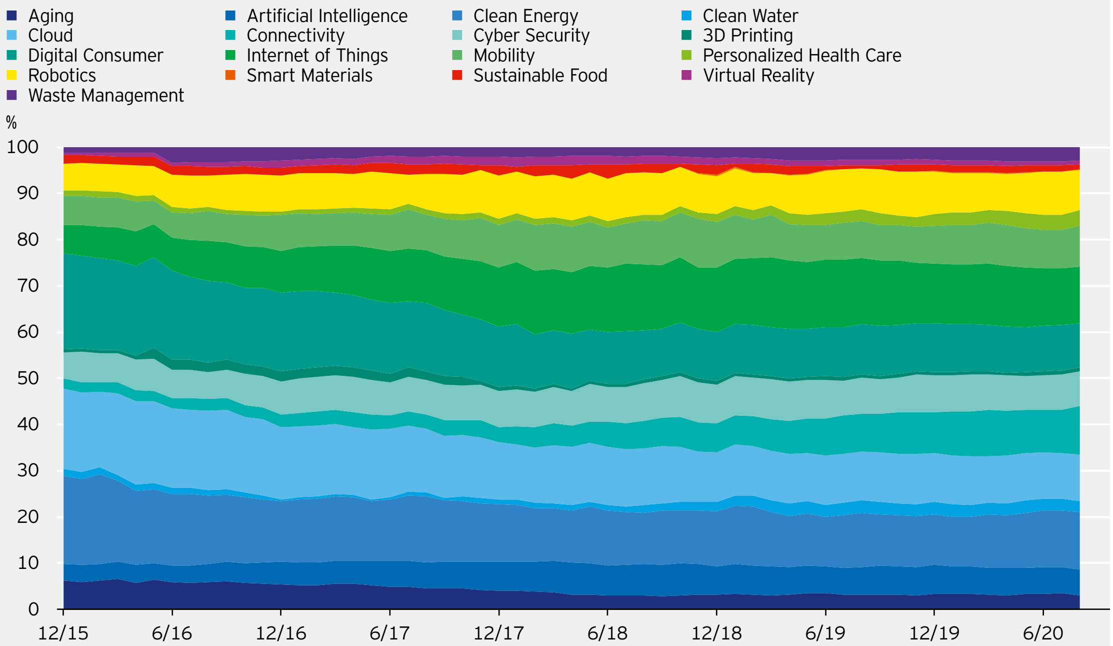
The rationale behind analyzing news data is the conviction that when a company is often mentioned in the context of a given theme and its underlying keywords, the theme is highly relevant for this company. Consequently, it should profit from the theme's rising importance. Relevance is determined through the occurrence of news hits and the strength of the narrative to the underlying themes.

Figure 3
News sources and number of news items per source for the month of April 2020
Total number of news reports: ~2,500,00



Source: Invesco. For illustrative purposes only.

Figure 4
Theme weights over time (simulated)



Source: Invesco. Data as at August 2020. Proprietary theme dictionaries as at 30 April 2020 were used to screen monthly news data for companies associated with the investment themes listed in the legend. Monthly portfolio rebalancing.

In the end, we have identified companies which represent several themes. Together, they would form a diversified multi-theme portfolio, where the theme weights and the company weights within the themes are driven by the number of occurrences in the news data.

Figure 4 shows how the theme weights would have developed over time: for example, personalized health care and connectivity would have risen in importance, whereas, surprisingly, the relative importance of clean energy and the digital consumer would have declined. However, this does not mean that these topics are no longer as relevant as they once were – it only proves that other, newer topics have emerged.

All in all, it turns out that the companies chosen by the NLP-based process are:

– pure play companies that have a high exposure to a single theme, but where the overall number of news stories may be low (e.g. a wind energy company);

– large companies that are highly relevant for multiple themes, but whose exposure to each single theme may be comparatively low (e.g. big technology companies or conglomerates in transition).

Furthermore, with the prominent selection of sustainability-related themes, the method explicitly targets companies with substantial exposure to "green" technologies that are managing their environmental resources and carbon footprint in a dedicated and sustainable manner. Additionally, an ESG screen could be implemented that seeks to filter out the companies scoring worst within the "green" themes.

We believe our new method has three main strengths:

1. The themes that form a megatrend are constantly changing. Themes with declining relevance will generate fewer news hits, and their weight will automatically decrease, while new themes with increasing relevance will receive higher weights over time so that more stocks that capture these themes are selected.

2. A huge dataset with respect to the themes underlying megatrends is analyzed, leading to a very broad view on the relevant topics. New information is constantly captured through the news data.

3. The main focus is on the selection of themes and the identification of relevant companies. The concept is less focused on identifying the one company that stands for a certain theme – it is more about identifying a group of companies associated with a certain theme and gaining exposure to these companies.

Summary
We have developed an NLP-based method to identify and weight companies exposed to certain megatrends and their underlying investment themes. By analyzing large datasets and alternative data sources, this may give asset managers an innovative toolkit to ensure they are invested in the right themes at the right times, and that their clients benefit from a diversified exposure to these themes and the companies driving them.

About the authors

Georg Elsaesser
Senior Portfolio Manager
Invesco Quantitative Strategies
Georg Elsaesser is involved in the management of Invesco Quantitative Strategies equity portfolios and has long-standing experience in the implementation of factor-based investment solutions.

Dr. Martin Kolrep
Senior Portfolio Manager
Invesco Quantitative Strategies
Martin Kolrep is focused on the development and management of multi-asset portfolios and outcome-oriented strategies. As a physicist he has gained important experience in the field of artificial intelligence and machine learning.

Alexandar Cherkezov
Portfolio Manager
Invesco Quantitative Strategies
Alexandar Cherkezov manages multi-asset portfolios that include the elements factor-based investing, active asset allocation and downside risk management.

Michael Rosentritt
Portfolio Management Associate
Invesco Quantitative Strategies
Michael Rosentritt is responsible for model and performance analysis of quantitative equity strategies and investment communication.

Note

1 News items are processed in the English language. News reports originally sourced in a foreign language are translated into English.

"A diversified exposure to a variety of themes makes sense."

Interview with Alexandar Cherkezov, Dr. Martin Kolrep and Michael Rosentritt

Alexandar Cherkezov

Dr. Martin Kolrep

Michael Rosentritt

Risk & Reward spoke to Invesco's Alexandar Cherkezov, Dr. Martin Kolrep and Michael Rosentritt about stock selection with Natural Language Processing (NLP) and the importance of innovation.

Risk & Reward
What was your intention when you began developing the NLP algorithm?

Martin Kolrep
We wanted to create something new and innovative while at the same time building upon our extensive knowledge of data analysis. Our main idea was to use novel and largely untapped data sources to view companies from a different perspective, from the angle of innovation. We believe innovation will be key for the success of businesses as we move into the future. This was already the case in the past, but we believe it will be even more important going forward. Companies that adopt and drive new technologies and develop solutions for challenges such as aging of the population and sustainability will likely shine. But companies holding fast to existing methods and ignoring these challenges may face difficulties – it's really very simple.

Risk & Reward
Can you give us an example?

Alexandar Cherkezov
There is a great deal of discussion about the adoption of artificial intelligence (AI). This is probably one of the primary technologies that will determine how businesses function in the future. It is anachronistic to ignore the potential. For example, when you are talking to your smartphone, there is AI working in the background. The same holds true for customer service chatbots, or the tools a doctor uses to screen for cancer. All of this is already quite commonplace, but we expect it will continue to advance rapidly. It is already the case that AI improves hit ratios in cancer screening. Would you go to a doctor who doesn't want to utilize this technology?

Risk & Reward
What does this mean for portfolio management?

Michael Rosentritt
Traditionally, theme managers talked to companies to find out what they have in the pipeline. Their portfolios were typically quite specialized in one area, and sometimes also quite concentrated. But even more, innovation drives productivity and productivity drives growth. In the long term, growth drives capital returns and market value. By focusing on innovative companies, portfolio managers can participate in this capital growth. This makes it important to identify them as early as possible.

Risk & Reward
We often hear that many innovative companies are not yet listed on the stock exchange.

Michael Rosentritt
There are enough innovative companies that are already in their scale-up phase, which have achieved a few milestones – and proven that their ideas work. Typically, this happens when companies need more capital and go public. Not many companies of meaningful size have chosen to stay private. Moreover, focusing on listed equities can make a portfolio much more liquid.

Risk & Reward
And what about the balance of pure plays versus conglomerates? Does your algorithm have any preference?

Alexandar Cherkezov
Our NPL-based process chooses both smaller pure play companies, which are highly active in one specific investment theme, and larger conglomerates with multiple business lines and activities in several themes. In fact, it takes both the relevance of an investment theme for a company and the relevance of a company for an investment theme into account. This is likely to lead to a truly diversified and balanced portfolio.

A smaller pure play company might not be highly important for its theme due to its smaller size, but the theme is often highly relevant for the company. A smaller wind energy company would be a good example.

For a larger conglomerate with activities in several investment themes, each individual theme might not be highly relevant, as it has many different business lines and generates its revenues in multiple themes. But, due to its size, the company will be highly relevant for each theme. An example might be a Big Tech company or an energy conglomerate transitioning into renewables.

Risk & Reward
From what you've said, I assume you prefer a multi-thematic approach.

Martin Kolrep
We are convinced that a very narrow focus and a limited number of investment themes is not the way to achieve the best outcomes. A diversified exposure to a variety of themes makes sense for several reasons.

First of all, investment themes are much more interconnected than they used to be. Innovations within one theme can have a significant impact on others. For example, developments within the "AI" theme not only have a significant impact on themes like "mobility" (e.g. self-driving cars) and "robotics", but also on "smart agriculture" – e.g. computer-based decisions on when to fertilize or harvest – and "personalized health care", including the development of new drugs. Exposure to a variety of themes can therefore be very beneficial.

Furthermore, many companies today are active in more than one theme, and the distinction between their different activities is fluid. For instance, we see companies from the industrial or automotive sector becoming active in artificial intelligence and cloud solutions to optimize their production, and former oil majors are pushing into renewable energies like wind or solar power. Thus, investors are likely to get a number of different theme exposures by investing in just one company.

Finally, the relevance of different investment themes is constantly changing. Since the outbreak of Covid-19, for example, we've seen a rising relevance of themes like "connectivity" or "personalized health care". Switching between single-theme investments can be tedious and expensive. Hence, a relevance-driven, multi-thematic approach can help investors remain exposed to the most relevant themes without holding on for too long to themes with declining relevance.

Risk & Reward
Thank you for your insight.

Searching for inner peace with value factors

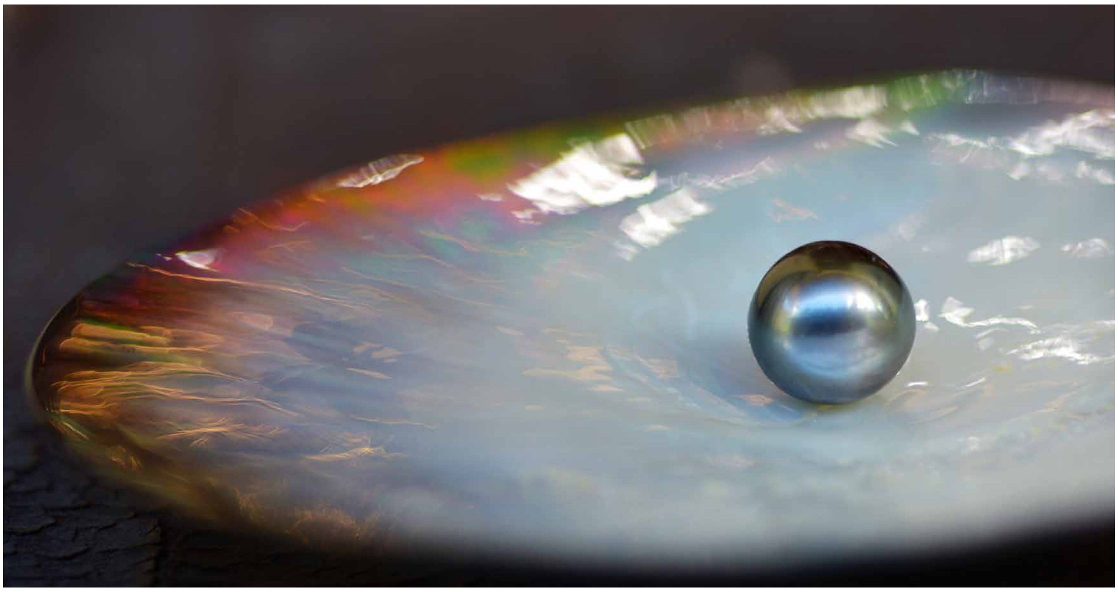
By Yifei Shea, Ph.D. and Erhard Radatz

In brief
Value factors have always been an essential part of quantitative investing processes. We show that the deterioration in performance of value stocks, as defined by high book-to-price (B/P) ratios, comes as no surprise given their relatively poor fundamentals as well as slower mean reversion of profitability. The opposite has been observed for low B/P "glamour" stocks. Value defined by high earnings yield (E/P) has also suffered from deteriorating performance despite being fundamentally different from value characterized by a high B/P.

The validity of value factors derived from various price multiples has been called into question amid their prolonged underperformance, which accelerated over the initial months of 2020. For instance, the Fama and French (1992) HML global value factor portfolio is down 40% from its last high in 2007.[1] To understand the value effect, i.e. the long-term excess return of value stocks over glamour stocks, we explore the links between value, profitability and EPS growth.

The validity of value factors derived from various price multiples has been called into question amid their prolonged underperformance.

The definition of value stocks as stocks with high B/P ratios was popularized by Fama and French (1992) and has been widely used among the quantitative equity community. We start by comparing the performance of such stocks and their opposites, low B/P "glamour" stocks, in the US large cap universe[2] from January 1991 to May 2020. Figure 1 shows



the prolonged drawdown of high B/P cheap stocks since mid-2007, while the most expensive glamour stocks delivered consistent outperformance.

To understand the deterioration in value's performance, we first need a better grasp of what B/P captures. While B/P can act as a value factor, it is not useful as a valuation metric to identify stocks that are cheap relative to their intrinsic value: book equity measures the equity capital input, while the actual output or profit is what a valuation model focuses on (e.g. Chapter 6 of Barker, 2001). Intuitively, given the same amount of capital input, higher future profits should command higher market valuation. Therefore, we expect to see a negative relationship between this input-output ratio (i.e. a profitability metric such as return-on-equity (ROE)) and book-to-price.[3] Indeed, figure 2 shows the median ROE of the most expensive B/P quintile being consistently higher than that of the universe, and the median ROE of the cheapest B/P quintile consistently lower.

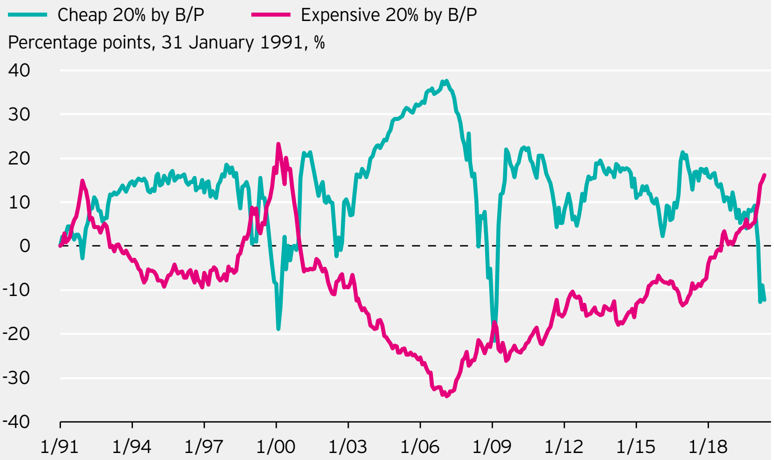
Figure 2 also indicates an increasing ROE gap between value and glamour; this is interesting as the widening of the value spread in recent years (shown in figure 3) is frequently highlighted and compared with the Technology, Media and Telecom (TMT) bubble period. However, comparing figure 2 and figure 3, we note an important distinction between the two periods: the glamour stocks of today are much more profitable than those during the tech bubble.

To further understand how value or glamour stocks may have differed over the last decade or so, we investigate how the profitability of a typical company within a value or glamour portfolio evolves through time. We introduce the concept of mean reversion, which is the tendency for individual companies' profitability to converge to the mean over time (e.g. Chapter 3 of Stigler, 1963). Mean reversion happens because any market segment with abnormally high profitability will attract competition and see profitability revert to average, while any segment with low profitability might be going through cyclical challenges and will strive to improve profitability or otherwise drop out. Given that value stocks tend to have low ROE and vice versa, we expect to see, on average, a trajectory of improving ROE for value stocks and deteriorating ROE for glamour stocks.

> Given that value stocks tend to have low ROE and vice versa, we expect to see, on average, a trajectory of improving ROE for value stocks and deteriorating ROE for glamour stocks.

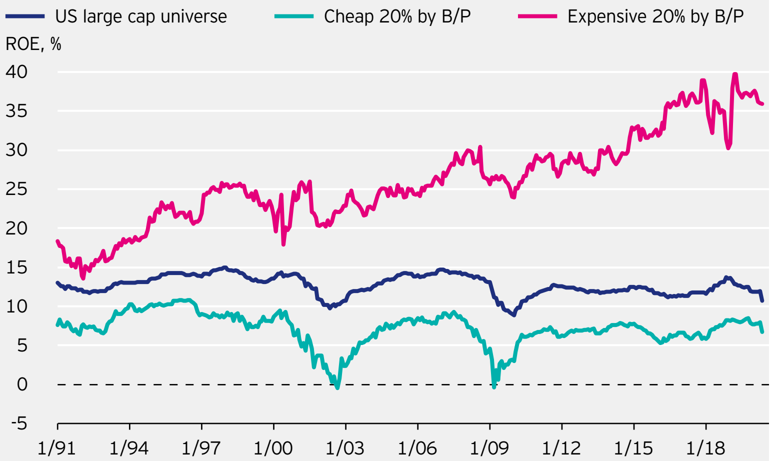
For our examination, we form buy-and-hold value, glamour and market portfolios for every month, and record their median ROE trajectories over the subsequent three-year period. We use median ROE

Figure 1
Cumulative return of value and glamour relative to the US large cap universe



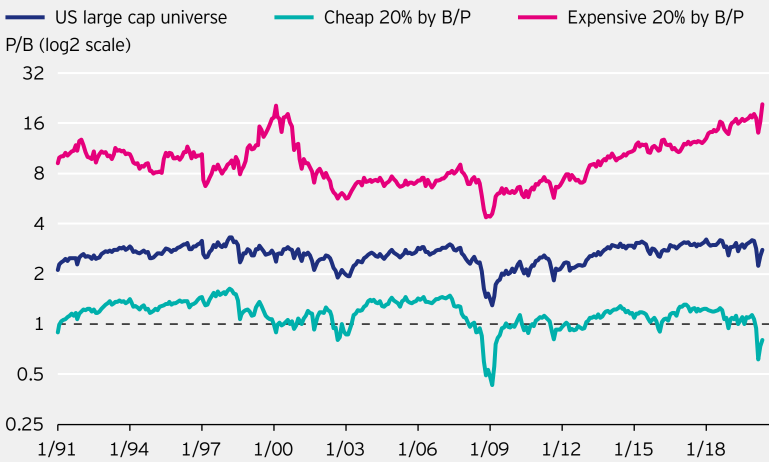
Source: Invesco. Data as at 31 May 2020. **Past performance is not a guide to future returns.**

Figure 2
Median ROE of value, glamour and the US large cap universe



Source: Invesco. Data as at 31 May 2020.

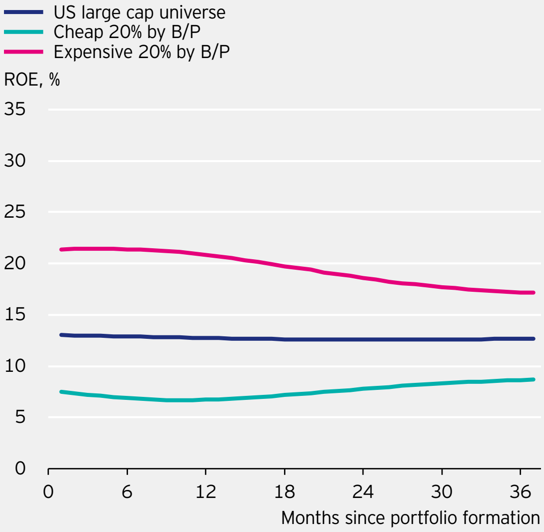
Figure 3
Median price-to-book of value, glamour and the US large cap universe



Source: Invesco. Data as at 31 May 2020.

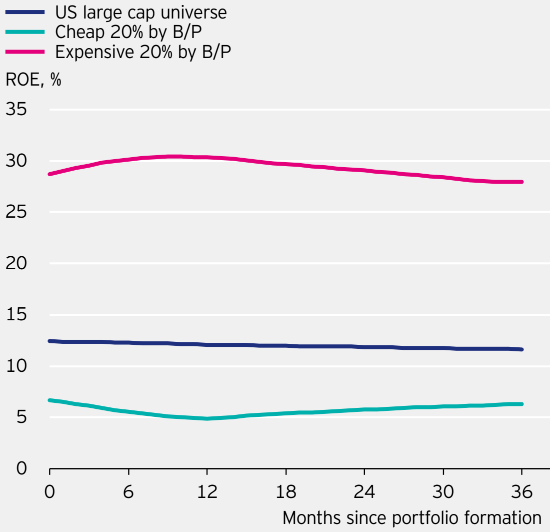
Figure 4
ROE of buy-and-hold value and glamour defined by B/P over two sub-periods

Portfolios formed between 31 January 1991 and 30 November 2004



Source: Invesco.

Portfolios formed between 31 December 2004 and 31 May 2017



Source: Invesco.

instead of aggregate ROE of the market cap weighted portfolio so that our results are not dominated by the ROE of large companies or driven by a few outliers. We then divide the entire sample period roughly evenly into two sub-periods to inspect how the ROE trajectories may have changed over time: one with portfolio formation between 31 January 1991 and 30 November 2004 and the other with portfolio formation between 31 December 2004 and 31 May 2017. The average of ROE trajectories across these two sub-periods are shown in figure 4.

First, consistent with figure 2, we note a wider gap in ROE between value and glamour stocks during the more recent period. Second, in both periods, the distance between median ROE in the cheap and expensive groups decreased over time, indicating mean reversion of profitability. Finally, comparing the ROE trajectories over the two sub-periods in figure 4, it should be noted that the rate of mean reversion slowed for value stocks relative to glamour stocks in the second period. Therein, we define the relative mean reversion rate of cheap versus expensive group of stocks as one minus the annualized ratio of their ROE differentials at the end of the three-year period and at the beginning of the period. We thus calculate that the relative mean reversion rate decreased from 15.1% p.a. in the first period to 0.6% p.a. in the second period.

To investigate whether these results could be driven by industry effects, as a next step we construct industry-neutral value factor portfolios. We note that industry neutralization is a common approach utilized by factor investing practitioners, in which a stock's B/P is compared with its industry peers. There are multiple ways one can construct industry-neutral factor portfolios. For illustration purposes, we simply rank each stock's B/P within its respective industry[4] at each point in time. We then form value and glamour portfolios based on the top or bottom 20% industry-adjusted B/P scores and analyze the 3-year median ROE trajectories for buy-and-hold portfolios. For the sake of robustness, we also investigate the same two groups' return-on-assets (ROA) and profit margin (ROS) trajectories. ROA is less impacted by stock repurchases, and ROS, despite not being a profitability metric, is meaningful for comparison within industries (additionally, sales is a cleaner accounting metric than book or assets). Consistent with earlier findings, we note a larger average profitability or profit margin gap over the more recent period in the industry-neutral setting. Furthermore, relative mean reversion rates are lower in the more recent time period across all three profitability or profit margin metrics. The results are summarized in table 1. For comparison, we also include ROE regression rates when B/P is not industry adjusted.

Table 1
Annualized 3-year mean reversion rates over two sub-periods

Value metrics		**B/P**	**Industry-neutralized B/P**		
		ROE	**ROE**	**ROA**	**ROS**
Portfolio formation periods	31 January 1991 - 30 November 2004	15.1%	10.4%	10.5%	6.9%
	31 December 2004 - 31 May 2017	0.6%	3.5%	2.4%	-1.1%

ROE: return-on-equity, ROA: return-on-assets, ROS: return-on-sales (profit margin). Note: Mean reversion rates in ROE, ROA and ROS are for value relative to glamour, annualized over a 3-year holding horizon.
Source: Invesco.

Given that the potential reasons for certain companies' ability to sustain abnormally high profitability or profit margin (e.g. technology advancement, low interest rate environment and globalization) would apply for almost all industries, we expected to see consistent results for industry-neutralized value and glamour, despite some industries benefiting more. At the same time, the slower profitability mean reversion rate of value relative to glamour, industry-neutralized or not, could simply indicate that the market has become better at deciding which stocks deserve to be priced cheaply.

Our analysis suggests that the ROE mean reversion rate for value relative to glamour slowed in the recent years, which potentially resulted in fewer EPS growth surprises for value.

We recognize that the mean reversion rate of profitability alone is not informative of whether the market has overestimated / underestimated a company's future profitability. However, inferring from the fact that value stocks defined by low B/P historically generated excess returns, it is reasonable to assume that the market has systematically underestimated the profitability mean reversion rate for both value and glamour. This may not be the case, however, in the more recent period. Some have shown that slower mean reversion in the B/P ratio itself is a major reason for deterioration in value's performance (e.g. Lev and Srivastava, 2020). We think it is likely that a slower mean reversion in profitability of value relative to glamour has driven the slower mean reversion in B/P.

The link between profitability and EPS growth

One may question why the profitability mean reversion rate, or more generally the trajectory of profitability, is relevant. Note that a company's profitability trajectory is directly linked to its future earnings-per-share (EPS) growth, which is an important determinant of intrinsic value. Based on clean surplus accounting, we can decompose EPS growth into "sustainable growth" (i.e. retention rate multiplied by ROE) and "efficiency growth" (i.e. percentage change in ROE), see Damodaran (2008):

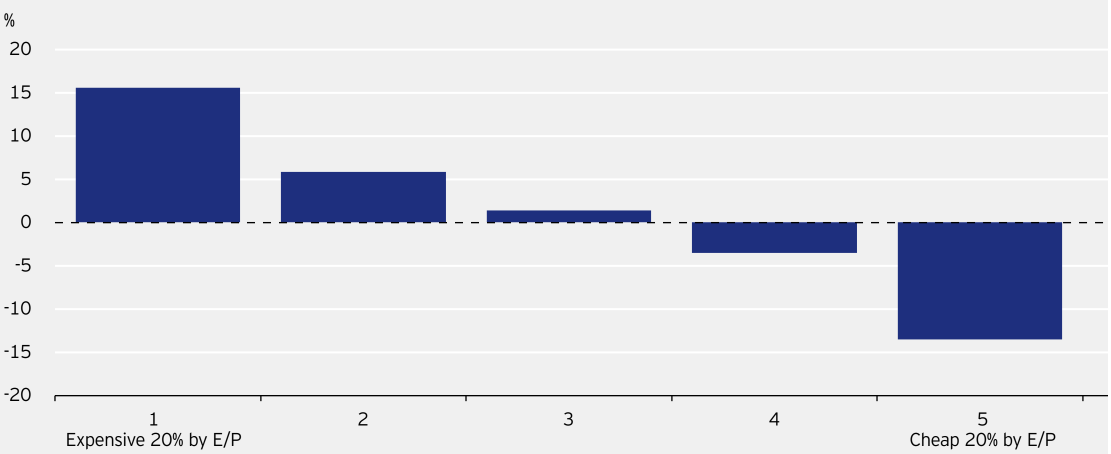
$$g_{EPS} = b \times ROE_{t+1} + (ROE_{t+1} - ROE_t)/ROE_t,$$

where b is the retention rate defined as the proportion of earnings retained or reinvested, and ROE_{t+1} is defined as next period earnings divided by current book equity.

The first term, retained earnings divided by current book equity, is equivalent to book-per-share (BPS) growth, which tends to be higher for companies with high and stable ROE. The second term is driving most of the volatility in EPS growth and, as a result, creates surprise. A simplified example: if a company with 50% retention rate has a stable ROE of 10%, then its EPS growth is equivalent to BPS growth at 5%. If its ROE increases to 11% in the next period, with the same retention rate, its EPS growth becomes 5.5% + (11%-10%)/10% = 15.5%. The 10.5% increase in EPS growth mainly comes from 10% efficiency growth.

In the previous section we argued that a value effect could have existed because the market under-estimates the profitability mean reversion rate of a subset of cheap B/P stocks. Based on the link between profitability trajectory and EPS growth, this is equivalent to saying that the value effect is driven by a subset of value stocks delivering better-than-expected EPS growth, which leads to their repricing. Our analysis suggests that the ROE mean reversion rate for value relative to glamour slowed in recent years, which potentially resulted in fewer EPS growth surprises for value.

Figure 5
Next-year EPS growth by E/P quintiles, relative to US large cap universe



Note: Non-positive earning companies are excluded from the universe in the current period to enable calculation of next-year EPS growth rates.
Source: Invesco. Data from 31 January 1991 to 31 May 2020.

Table 2
Annualized returns for monthly rebalanced portfolios relative to US large cap universe, over two sub-periods

	Value	Value trap	Value surprise
31 January 1991 – 31 December 2004	6.99%	-12.23%	13.07%
31 December 2004 – 30 June 2019	-0.15%	-15.22%	5.78%

Note: Non-positive earning companies in the current period are excluded from the universe (and subsequently the value portfolio), to enable calculation of next-year EPS growth rates for defining value trap and value surprise portfolios. The value portfolio consists of 20% highest E/P stocks in the US large cap universe. The value trap portfolio is defined as the intersection of value and 20% lowest next-year EPS growth stocks in the market; the value surprise portfolio includes the rest of stocks in value. All portfolio returns are market cap weighted; in each month, the return of the value portfolio is equivalent to the market cap weighted sum of returns of the value trap and value surprise portfolios. Within the value portfolio, the average percentage of value trap stocks by market cap is 24.1% in the first period and 26.1% in the second period.
Source: Invesco. **Past performance is not a guide to future returns.**

The link between value and EPS growth
The direct connection between profitability and EPS growth motivates additional insights into value, especially when we broaden value factors to include those defined by earnings yield (E/P), cash-flow yield and dividend yield etc. Indeed, such value metrics are more appropriate for valuation, since the numerators represent various forms of profit returned to shareholders.

In this section, we use value defined by E/P as an example to demonstrate how one can dissect the performance of value based on EPS growth. Unlike B/P, E/P does not have a clear relationship with ROE: we can rewrite E/P as the product of ROE and B/P, which indicates that a stock with high earnings yield could have either high ROE or high B/P (hence low ROE). On the other hand, there is a negative relationship between a stock's E/P ratio and its next year realized EPS growth rate (see figure 5) since earnings yield reflects EPS growth expectation by the market.

To test the hypothesis that the value effect arises because a subset of value stocks surprise the market with their EPS growth, we divide the stocks in the cheapest E/P quintile into two sub-groups: the "value trap" group (which includes stocks with next-year realized EPS growth in the lowest quintile) and the "value surprise" group (which includes the rest of value stocks in the cheapest E/P quintile). It is termed a "surprise" since a stock in the cheapest quintile E/P group is generally expected to have low EPS growth, as shown in figure 5. For diagnostic purposes, the groups are defined based on perfect foresight of next-year EPS growth, instead of EPS forecasts. We then define two sub-periods consistent with the earlier analysis.

First, we find in the more recent period (31 December 2004 – 31 May 2019) that the percentage of "value trap" stocks is 34.5%, compared to only 28.5% in the first period (31 January 1991 – 30 November 2004). Next we investigate the relative performance of value and its two sub-groups over each period (table 2). The results confirm that the historical outperformance of value stocks is driven by the subset with better-than-expected EPS growth. Table 2 also shows that the decrease in excess returns of value is mainly driven by worse performance of the "value surprise" sub-group, which experiences a greater than 7 percentage point reduction in excess returns in the more recent period. This might be another hint that the US market has become more efficient, but whether value will deliver excess returns still depends upon a subset of cheap stocks realizing better or worse-than-market's predicted EPS growth.

Summary and parting thoughts
The prolonged period of underperformance of value factors has brought the rationale of value into the spotlight. When measured by the B/P ratio, value stocks typically have low profitability, but their profitability reverts towards the average over time. The opposite happens for glamour stocks. We posit that the value effect exists when the market underestimates the mean reversion rate for value relative to glamour stocks, resulting in fewer EPS growth surprises and repricing for value. However, the mean reversion rate has slowed for value versus glamour in recent years, potentially leading to fewer such surprises and smaller price impacts of each surprise. These findings continue to hold in an industry-neutral setting, as well as when using earnings yield to measure value.

Still, value metrics such as those based on free cash flows or earnings provide a multi-factor model with a reasonable valuation anchor. Stocks with high multi-factor scores are those inexpensive relative to their quality and growth potential, mimicking an intrinsic value approach. Recent literature often highlights attempts to fix the performance of value as defined by B/P by capitalizing certain intangible assets. While this is a reasonable approach, the measurement of intangibles can be subjective, and we consider it preferable to incorporate this type of adjustments separately in the investment model. For instance, instead of capitalizing R&D to obtain an adjusted B/P factor, a more thoughtful approach would be to

Success in timing the outperformance/under-performance of value stocks relies upon one's ability to forecast when they will deliver better or worse future EPS growth relative to the market's expectations.

estimate the underappreciated value of patents held by a firm and utilize this as a quality or growth component.[5]

Given the headwinds involved with multiple-based value investing, one may wish to explore timing of value factors[6] via the much-watched value spread. Unfortunately, a wider value spread based on B/P does not necessarily lead to faster B/P reversion or higher profitability mean reversion rates in the future. Success in timing the outperformance/under-performance of value stocks relies upon one's ability to forecast when they will deliver better or worse future EPS growth relative to the market's expectations. Although value tilting may hold out reward for those confident in making a macro timing call on the economic recovery, a diversified multi-factor approach emerges as a prudent alternative to navigate the downside of value.

References

Barker, Richard (2001): Determining Value: Valuation Models and Financial Statements, Prentice Hall (1st edition).

Damodaran, Aswath (2008): The Origins of Growth: Past Growth, Predicted Growth and Fundamental Growth, Working Paper.

Dichtl, Hubert, Wolfgang Drobetz, Harald Lohre, Carsten Rother, and Patrick Vosskamp (2019): Optimal Timing and Tilting of Equity Factors, Financial Analyst Journal, 75, 84-102.

Fama, Eugene F., and Kenneth R. French (1992): The Cross-Section of Expected Stock Returns, Journal of Finance, 47, 427-465.

Fraikin, Michael, and Edward Leung (2019): Patent as a Driver of Equity Returns, Invesco Risk & Reward, #2/2019, 18-22.

Lev, Baruch I., and Anup Srivastava (2020): Explaining the Recent Failure of Value Investing, Working Paper.

Stigler, George J. (1963): Capital and Rates of Return in Manufacturing Industries, Princeton University Press.

Wilcox, Jarrod W. (1984): The P/B-ROE Valuation Model, Financial Analysts Journal, 40(1), 58-66.

Wilcox, Jarrod W., and Thomas K. Philips (2005): The P/B-ROE Valuation Model Revisited, Journal of Portfolio Management, 31(4), 56-66.

Acknowledgement: The authors would like to thank Harald Lohre, Tarun Gupta, and Bernhard Breloer for helpful discussions and feedback on the paper.

About the authors

Yifei Shea, Ph.D., CFA
Senior Quantitative Research Analyst
Invesco Quantitative Strategies

Yifei Shea conducts research into quantitative models that drive the investment decisions for multi-factor equity products. Her current focus includes fundamental equity factors, machine learning and natural language processing.

Erhard Radatz
Portfolio Manager
Invesco Quantitative Strategies

Erhard Radatz manages multi-asset portfolios that include the elements: factor-based investing, active asset allocation and downside risk management.

Notes

1 An HML portfolio (high minus low portfolio) goes long the cheapest stocks and shorts the most expensive stocks according to book-to-price ratio. The exact drawdown is -39.4% from end of July 2007 to end of May 2020 (Source: Invesco, Kenneth French's website).
2 Defined as highest or lowest 20% B/P stocks respectively. The US large cap universe is constructed based on Russell 1000 constituents. Value and glamour portfolio returns are calculated based on market cap weighting and relative-to-market returns.
3 For instance, Wilcox (1984) and Wilcox and Philips (2005) provide theory and examples illustrating this negative relationship.
4 We use 35 industries defined per the Axioma risk model.
5 A recent Risk & Reward article by our colleagues discusses how patent information can be utilized to predict subsequent equity returns (Fraikin and Leung, 2019).
6 In general, our colleagues show that there is little room to benefit from dynamic equity factor allocation after accounting for transaction costs (Dichtl et al., 2019).

Invesco's guide to alternative data

By Tarun Gupta, Ph.D. and Edward Leung, Ph.D.

In brief
We explore the use of alternative data in factor investing, showing which sources have become most popular. We discuss challenges and give a brief overview of what asset managers should consider when implementing a factor investing strategy using alternative data. Our two case studies, on sales transaction data and earnings transcripts, shed light on promising up and coming alternative data sources.

Over the years, factor investing[1] has grown ever more popular. But, recently, some researchers have suggested that traditional factor returns have been arbitraged away.[2] This is where alternative data comes into play: such data may contain less-well known information and often comes with higher barriers to access, because the necessary infrastructure and expertise are not easily built. The potential is therefore great for innovative asset managers to gain an information edge.

Loosely speaking, alternative data can be defined as data that analysts and portfolio managers typically ignore, unlike published financial statements and economic indicators. Moreover, alternative data is usually less structured, less organized and less defined. Consequently, it is not so easily discovered. In short, we are talking about unconventional information in an unstructured format, which is not easily accessible to investment managers.[3] Table 1 summarizes the differences between traditional and alternative data.

Alternative data is usually less structured, less organized and less defined.

Table 1
Traditional vs. alternative data

	Type	Sources	Examples
Traditional data	Structured	– Listed companies – Exchanges – Index providers	– Price – Volume – Reported financials
Alternative data	Structured/unstructured	– Individuals, i.e. internet usage, mobiles – Businesses – Sensors	– Consumer transactions – News – User reviews – Weather

Source: Forterre, Kalian, Van Hemert, Kilduff and Hurley (2020).

Interestingly, looking at traditional data in a new way is also counted as alternative. The best example is applying natural language processing (NLP) to 13F filings or earnings transcripts to build sentiment factors.[4] Earnings calls are traditional data, but extracting the sentiment or tone of the executives from them requires additional steps. By applying NLP, hidden sentiment signals of future market and industry trends are uncovered systematically.

Typical alternative data used for building factors are:

- ESG-related data to capture an ESG metric for companies
- Online reviews to reveal customers' perception of various brands
- Product sales estimates from transaction data (credit/debit card) and email receipts
- New product sentiment from social media and web traffic
- Export activity from analyzing bills of lading (shipping) data
- Foot traffic data for predicting economic activities and macro trends
- Job posting data for predicting company growth

According to the Alternative Investment Management Association (AIMA),[5] some of the most popular alternative data used by hedge funds are web crawled, social media and transaction data. Asset managers use alternative data as a research tool to source new investment opportunities, improve investment decisions and as an input for quantitative research. Monk, Prins and Rook (2019) argue that it is also important for risk management and can help improve a firm's operating processes. According to Neudata, an alternative data evaluator, the cumulative number of alternative data sets released commercially grew from 285 to 1037 between 2010 and July 2020.[6]

Similar to supply, the demand for alternative data has also increased in recent years. In 2016, buy side firms spent about USD 200 million on alternative data. For 2020, the estimate is USD 1.6 billion.

In 2016, buy side firms spent about USD 200 million on alternative data. For 2020, the estimate is USD 1.6 billion.

Coincidentally, the average number of daily big data job postings in the US went from 300 in 2010 to 900 in 2018.[7]

Challenges
This growth in alternative data poses challenges, and not all data is of equal quality. When applying alternative data to quantitative investment processes, we need to be mindful of a number of things:

- **Relevance:** Is the data useful for a given trading strategy or investment philosophy? How can we extract alpha?
- **Entity mapping:** Do the alternative data vendors have a robust method of entity mapping? Examples are (1) mapping merchant/company names to security identifiers (such as CUSIP or SEDOL) on a point-in-time basis and (2) mapping the names of corporate executives appearing in earnings call transcripts.
- **History:** Is the data history sufficient? As of 2019, more than 50% of the alternative dataset has a history of less than seven years. A short history makes meaningful backtests difficult.
- **Coverage:** Is there enough breadth? Ideally, the data should cover companies from all countries and industries. But alternative data tends to be industry specific. For example, transaction data tends to be mostly available for the US consumer discretionary sector and patent data tends to be mostly available for healthcare/biotech companies.
- **Granularity:** Is the data user level, company level etc.? For instance, the underlying consumer panel of transaction data is user level, but it can be aggregated to company level.

- **Frequency:** Data can be high frequency, daily or monthly. Does the frequency fit the purpose?

- **Data lag:** Is there a lag between collection and reporting? This may be due to the vendors' internal process or regulatory issues. Long data lags will lessen predictive ability.

- **No bias:** Is the data point-in-time and free of biases such as backfill, look-ahead or survivorship biases?

- **Representative:** Data such as transaction or mobile phone data is based on an underlying panel/sample of consumers. Is the underlying panel/sample stable enough and representative of the population?

- **Business risk:** What if the alternative data vendor goes out of business after subscription?[8] Small unknown providers are likely to have higher business risk.

- **Governance:** Does the data comply with data protection, copyright and other regulatory rules?

Whereas most of these challenges are self-explanatory, the final point appears more complex. Concerns surround obtaining material non-public information (MNPI), data owner consent, privacy issues and intellectual property rights infringement to name just a few. In the EU, the General Data Protection Regulation (GDPR)[9] requires anonymization to be irreversible so that it is impossible to re-engineer the data and back out the identity of individuals. In the US, personal identifiable information (PII) is limited to categories such as the name, address or telephone number. But under GDPR, PII can also include IP addresses, web cookies or photos. In short, all PII is personal data but not all personal data is PII.[10]

> Just because data is accessible does not mean it is public.

Web crawling/scraping can be a legal issue as well. Lots of web data appears behind paywalls, but many pages are publicly accessible. However, every website has its own terms of use, which may prohibit web scraping.[11] And the pages may contain material non-public information (MNPI); just because data is accessible does not mean it is public. In the absence of an industry-wide standard, we can refer to the US National Institute of Standards and Technology's guide to protecting the confidentiality of PII.[12]

Thorough due diligence[13] on the data source, as well as the alternative data provider, should be performed to ensure that no personal data is being used and that the set complies with data protection, copyright and other laws and regulations.

Last but not least is the issue of cost: 30% of alternative data series cost between USD 50,000 to USD 150,000 per year, and 15% between USD 150,000 and USD 300,000.[14] Transaction data is usually the most expensive. However, according on Neudata, prices are coming down because of growing competition,[15] with more and more vendors offering the same data. For the most ambitious quantitative asset managers, we estimate that an in-house data science team costs at least USD 1 to 2 million per year, depending on the technology, talent base and objective.

How to onboard
With these challenges in mind, integrating alternative data into a quantitative investment process should follow the steps enumerated below:

1. **Relevance:** Examine whether the data is useful for a given trading strategy or investment philosophy.

2. **Interview:** Interview the data provider to understand the nature and challenges of the data.

3. **Due diligence:** Perform due diligence on the vendor to assess business and legal risks.

4. **Trial testing:** Trial the data to see if it is indeed relevant, i.e. consistent with the factor rationale and expected performance. Conduct entity mapping and various data checks; usually a trial agreement is needed.

5. **Production license:** If the data passes the trial and users decide to subscribe, a production license is needed.

6. **Data onboarding/infrastructure:** Decide how to store the data; for example, If the data is linkage data, it is best stored in a graph database. If it is more structured, it is best stored in an SQL database.

7. **Implementation/production:** Decide on the best delivery method, e.g. via FTP (File Transfer Protocol).

8. **Maintenance:** Monitor the quality of the data and factor performance/discrepancies over time.

The following two case studies illustrate the use of alternative data in factor investing.

Case study 1: Transaction data
Factors constructed from transaction data (i.e. credit and debit card data) can predict earnings surprises and are effective in predicting subsequent cross-sectional stock returns. Even after deducting turnover and trading cost, factor performance has been concluded to be still substantial.[16]

> Credit/debit card sales are excellent proxies for consumer demand.

Credit/debit card sales are excellent proxies for consumer demand (which is the source of firms' cash flow) because consumers use these cards for larger value items and are more likely to spend more compared to cash. Importantly, transaction data is timelier than companies' quarterly filings because it captures firm-specific real-time economic activity that tracks

Table 2
Performance of a credit card factor for the US all cap universe
January 2003 to December 2019

Credit card data only	Long (gross)	Short (gross)	Long - short (gross)	Long - short (net)
Return p.a.	17.56%	-1.47%	19.03%	16.03%
Standard deviation p.a.	15.04%	19.41%	11.12%	11.12%
Information ratio	1.17	-0.08	1.71	1.44

Source: Invesco. The average returns of the top and bottom bins are equal weighted. The factor return spread and information ratio are calculated by dividing the average spread by the standard deviation of the spread over the same time span. **Past performance is not a guide to future returns.**

consumer activity. It is tightly linked to underlying sales fundamentals. For example, credit card data shows that sales of athletic retailer Finish Line were pacing down 15% in the 3rd quarter of 2016. The shares fell 8.7% after its 21 December 2016 earnings report. The same happened to Bed Bath & Beyond on 22 December 2016.[17]

Our analysis focuses on the consumer discretionary names in the Russell 3000, i.e. an all cap universe, in order to maximize breadth/coverage using monthly data from January 2013 to December 2019. For these companies, we construct a factor based on the year-to-year % change in sales. Our model portfolio consists of long positions in firms with a high factor score and short positions in companies with a low score.

Transaction data is ideally suited to illustrate some of the shortcomings of alternative data where the time span is short, coverage is low and the data is industry specific. The highest coverage of this transaction dataset is approximately 50% of the consumer discretionary sector. The time span – monthly – is shorter than most of the traditional data, which goes back to at least the 1970s. In extensive interviews with the alternative data vendor, we ensured that the underlying data is legal for us to use and free from biases.

We then divided the transaction factor scores into bins and calculated the average forward one-month return of the top and bottom bins. The return spread is computed as the difference between the top bin average forward one-month return and the bottom bin average forward one-month return.

Table 2 summarizes the performance of the transaction data factor in our universe: the first column shows long-side gross performance, the second shows short-side gross performance. The third column contains the overall gross performance (long minus short) and the fourth column the overall performance net of transaction cost.

Most of the performance is from the long side, i.e. the top tercile of the companies with the highest factor scores. Note that the overall performance net of trading cost has been still very significant.

Case study 2: Gleaning sentiment from earnings transcripts via NLP
Applying natural language processing (NLP) to 13F filings or earnings call transcripts to build sentiment factors is also considered to be a form of alternative data because a novel approach is applied to tease embedded information from traditional data.

In this vein, two of our colleagues investigated several indicators of managers' sentiment using conference calls of US companies and assessed their predictive power for future stock returns. They found that the average sentiment of managers over the last twelve months and their degree of emotional

Machine learning
Transaction data, sentiment from social media, web crawling and ESG are some of the most popular factor ideas in the realm of alternative data. Another strand of research that has gained a lot of traction is the application of machine learning in asset management. The verdict on applying machine learning to predict cross-sectional stock returns using traditional data is mixed; see Leung, Lohre, Mischlich, Shea and Stroh (2020). But machine learning combined with alternative data such as geolocation has shown promising performance; see Liew, Budavari, Kang, Li, Wang, Ma and Fremin (2020). Compared to traditional data, alternative data provides significantly more observations, which is essential for machine learning to succeed.

In the broadest sense, the term Internet of Things (IoT) encompasses everything connected to the internet. But it is increasingly used to define objects that "talk" to each other. Simply put, the IoT is made up of devices that connect together, from simple sensors to smartphones, and IP addresses. IoT can track customers' real-time location to better understand their behavior, generating micro-level information to better predict the future performance of corporations.[18] Liaukonyte and Zaldokas (2020) use TV advertisement and the internet protocol (IP) addresses from SEC EDGAR visitation logs at specific geographic locations to study retail investor searches for financial information and their subsequent trading activity. The ad-induced queries increase trading volume and contribute to a temporary increase in stock returns. More and more behavioral information on retail trading can be discovered from IoT data to provide insight to quantitative asset managers as we head into the future.

Table 3
Performance of combined sentiment factor for the US large cap universe
2004 to 2017

Annualized performance of sentiment factors	Long	Short	Long - short	Standard deviation	Information ratio
Net sentiment and self deprecation	1.56%	-1.44%	3.00%	5.06%	0.59

Source: Invesco. **Past performance is not a guide to future returns.**

levelness during these calls helps explain future returns. What makes their analysis unique is that they partnered with a team of linguists to devise a set of proprietary dictionaries to determine their factors. They focused their investigation on the US large cap universe (Russell 1000) from 2004 to 2017.[19]

For instance, self-deprecation is what managers do when describing the performance of their firms in unfavorable terms. The rationale for this indicator is that, when managers use an excessive amount of self-deprecation, they may unwillingly reveal their lack of certainty about their firm's prospects.

Sentiment strength requires that positive and negative sentiment words can have different levels of sentiment strength intensity, but the sentiment strength of a word can be further magnified through the addition of specific words. For instance, a low intensity "sad" can be turned into a high intensity "very sad". Similarly, a high-intensity emotion word like "great" can be lowered in intensity by an attenuating addition, "not so great".

Our colleagues sourced all available transcripts of quarterly earnings conference calls from FactSet Research Systems Inc. and focused on managers' answers to questions from analysts in the Q&A section. To mitigate the risk of analyzing text that does not relate to the managers' answers, they required answers to be flagged as such in the transcripts and that an ID be available for the person talking to ensure that a manager is indeed speaking. Secondly, these IDs had to be related to a participant identified as a corporate representative in the section on participant information.

This case study illustrates the difficulty of entity mapping when processing unstructured alternative data. Often, there are inconsistencies with respect to documenting the names of the corporate representatives, such as John Doe vs. John C. Doe, vs. Doe, J., etc. Hence, a systematic approach needs to be used to determine whether all these John Does are the same individual.

Table 3 summarizes the decent performance of self-deprecation and sentiment strength as a single factor.[20]

Conclusion
We have surveyed the application of alternative data in asset management, identifying issues that quantitative managers need to be mindful of when evaluating alternative data. Currently, there is no agreement on how to address some of the shortcomings of alternative data, such as low coverage, short history or industry specificity. One potential method to improve coverage is to combine the same type of alternative data from different providers. For instance, transaction data coverage may improve if we combine transaction data from more than one provider.

Adapting alternative data is costly, whether a new internal data science team is set up to process the raw unstructured data or more structured data is bought from external vendors. In any case, a thorough cost-benefit analysis is warranted. The associated legal and regulatory risks can be lessened by applying thorough due diligence of alternative data vendors – but, because of diverse government policies, no process is foolproof.

Asset managers need to contemplate wisely how to reap the potential benefits of alternative data.

In short, asset managers need to contemplate wisely how to reap the potential benefits of alternative data. Based on AIMA survey information, 61% of hedge fund managers expect "alt data" to become more widely adopted over the next one to five years, suggesting a bright future.

References

Chen, M., Lee, J., and Mussalli, G. (2020): Teaching Machines to understand Chinese Investment Slang, The Journal of Financial Data Science Winter 2020, 2 (1), 116-125.

Cong, L.W., Li, B. and Zhang, Q.T. (2019): Alternative Data in FinTech and Business Intelligence, SSRN Working Paper.

Denev, A. and Amen, S. (2020): The Book of Alternative Data, A Guide for Investors, Traders, and Risk Managers, Wiley, 2020.

Forterre, A., Kalian, H., Van Hemert, O., Kilduff, E. and Hurley, N. (2020): The Data on Data, Man Group/Eagle Alpha Man Institute Paper, May.

Gupta, T. and Leung, E. (2020): Get More out of Plastic: An Investigation on Transaction Data, Earning Surprise, and Cross-Sectional Stock Returns, Invesco Quantitative Strategies Working Paper.

Leung, E., Lohre, H., Mischlich, D., Shea, Y. and Stroh, M. (2020): The Promises and Pitfalls of Machine Learning for Predicting Stock Returns, SSRN Working Paper (2020).

Liaukonyte, J. and Zaldokas, A. (2020): Background Noise? TV Advertising Affects Real Time Investor Behavior, SSRN Working Paper.

Liew, J. K-S., Budavari, T., Kang, Z., Li, F., Wang, X., Ma, S. and Fremin, B. (2020): Pairs Trading Strategy with Geolocation Data – The Battle between Under Armour and Nike, The Journal of Financial Data Science Winter 2020, 2 (1), 126-143.

Mahdavi, M. and Kazemi, H. (2020): It's All About Data: How to Make Good Decisions in a World Awash with Information, The Journal of Financial Data Science Spring 2020, 2 (2), 8-16.

McLean, R.D. and Pontiff. J. (2016): Does academic research destroy stock return predictability?, The Journal of Finance 71.1, 5–32.

Monk, A., Prins, M. and Rook, D. (2019): Rethinking Alternative Data in Institutional Investment, The Journal of Financial Data Science Winter 2019, 1 (1), 14-31.

Noble, L. and Balint, A.F. (2020): Casting the Net, How Hedge Funds are Using Alternative Data, AIMA.

About the authors

Tarun Gupta, Ph.D.
Managing Director of Research and Global Head of Investment Technology
Invesco Quantitative Strategies
Tarun joined Invesco as part of the IQS team in 2019 to focus on innovative research to enhance quantitative research strategies and spearhead development of investment technology. He also serves as a member of IQS's management team responsible for strategic planning and direction.

Edward Leung, Ph.D.
Quantitative Research Analyst
Invesco Quantitative Strategies
Edward Leung conducts quantitative research into factors, portfolio construction and risk modelling, currently focusing on machine learning and alternative data.

Notes

1 Factor investing is an investment approach that attempts to capture systematic sources of return via transparent, rules-based portfolio construction. In each period, the expected return over the risk-free rate of a stock is modeled as a linear function of a set of systematic factors and an unsystematic return component that captures company-specific shocks. Such a multi-variate factor model reduces risk by reducing unintended factor bets and providing investors with a better understanding of the sources of returns. Some of the most common systematic factors are size, value, momentum and profitability.
2 E.g. MacLean and Pontiff (2016).
3 For a more extensive assessment of alternative data, see Denev and Amen (2020) and Mahdavi and Kazemi (2020).
4 NLP can also be applied to similar filings documents in other countries or even documents in non-English languages. See Chen, Lee, and Mussalli (2020). The Securities and Exchange Commission's (SEC) Form 13F is a quarterly report that is required to be filed by all institutional investment managers with at least $100 million in assets under management. It discloses their equity holdings and can provide some insights into what the smart money is doing in the market.
5 See https://www.aima.org/educate/aima-research/casting-the-net.html
6 See https://www.neudata.co/press-resources.
7 See Forterre, Kalian, Van Hemert, Kilduff and Hurley (2020).
8 Recent example is the shutdown of RobinTrack of Robinhood (See https://www.businessinsider.com/point72-contacts-other-investing-apps-after-robinhood-data-taken-down-2020-8) and Thasos, an alternative provider of Geolocational data: https://www.businessinsider.com/hard-times-at-thasos-might-be-sign-alternative-datas-future-2019-9
9 https://gdpr.eu/what-is-gdpr/.
10 https://www.wsj.com/articles/lawmakers-call-for-investigation-of-fintech-firm-yodlees-data-selling-11579269600
11 See https://www.zdnet.com/article/appeals-court-linkedin-cant-block-public-profile-data-scraping/
12 See https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf
13 https://www.aima.org/resource/aima-ddq-for-alternative-data-vendors.html
14 https://www.neudata.co/press-resources
15 This was verbally mentioned in a Neudata Seminar.
16 For a full description of this analysis see Gupta and Leung (2020). This is one of the few analyses of the investability of a hedge portfolio based on alternative data that takes turnover, trading cost and portfolio allocation into account.
17 See https://www.wsj.com/articles/how-credit-card-data-might-be-distorting-retail-stocks-1483468912
18 See Cong, Li and Zhang (2019) for more discussions of IoT as an alternative data source.
19 This section is based on the article "What do corporate managers' words reveal about their firms' value" by our colleagues Michael Fraikin and Xavier Gerard, which appeared in the #4/2018 edition of Risk & Reward.
20 Note that the performance of the credit card factor and the NLP factor is very different because of the difference in rationale, data, investment universe and time span.

Investing through a macro factor lens

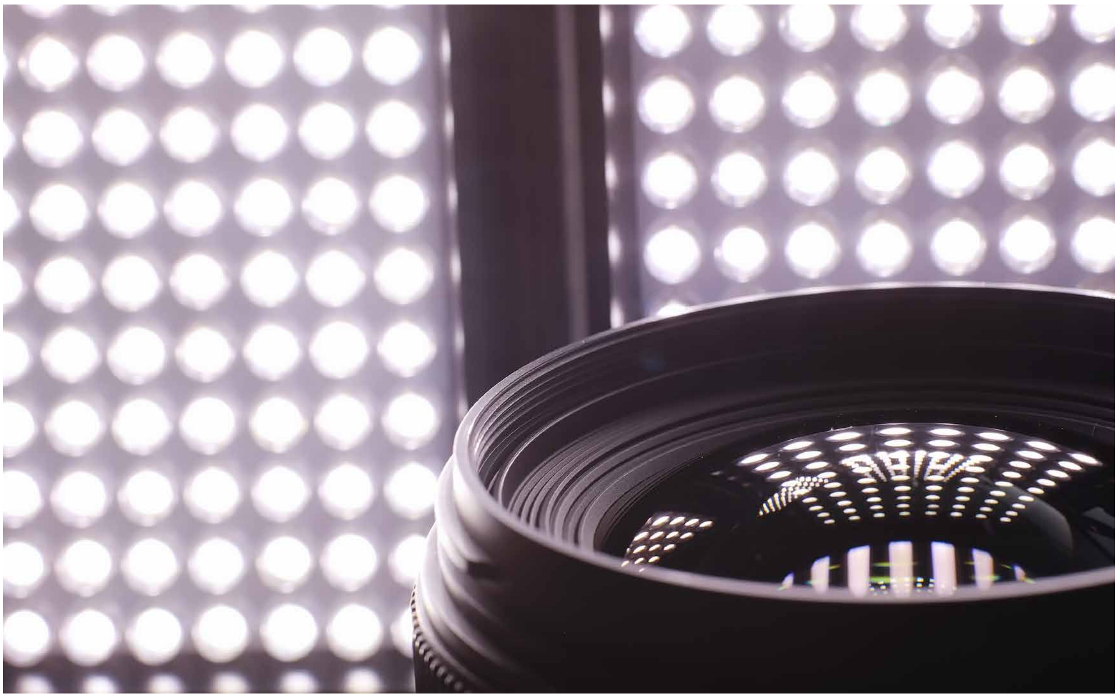
By Dr. Harald Lohre, Scott Hixon, Jay Raol, Ph.D., Alexander Swade, Hua Tao, Ph.D. and Scott Wolle

In brief

A macro factor perspective can help guide portfolio allocation by focusing on salient macroeconomic factors like growth or inflation. We study the link between such macro factors and common multi-asset multi-factor investment building blocks. Specifically, we investigate their macro factor sensitivities and propose a simple, yet effective, route to designing diversified macro factor-mimicking portfolios that prove beneficial in diversifying a given portfolio allocation with respect to its macro factor exposures.

The recent decade has seen a significant rise in factor-based investment propositions, most often focusing on style factor strategies, such as value or momentum. Style factors follow a clear investment rationale and are useful in diversifying a given traditional asset allocation. As many investors are concerned with shocks in macroeconomic variables like growth and inflation, they wish to understand and position multi-asset multi-factor allocations through a relevant macro factor lens.

To set the stage, we briefly recall three general types of factor models as juxtaposed in the seminal paper by Connor (1995). First, *macroeconomic* factor models use macroeconomic variables, e.g. inflation or interest rates, to explain asset returns. Second, *fundamental* factor models use factor portfolio returns related to certain asset characteristics, such as book-to-market or price momentum. Third, *statistical* factor models aim to create factors that naturally hold good explanatory power for the assets under consideration. Yet statistical factor models are often lacking when it comes to shaping the economic intuition of employed factors. Macroeconomic factors, on the other hand, are intuitive but generally provide the lowest explanatory power, leaving a sizable gap of unexplained specific risk.



To strike a balance between these three, we discuss and define macroeconomic factors and investigate the sensitivity of asset classes and style factors with respect to these macro factors. We show why economic regimes matter in constructing effective macro factor-mimicking portfolios and how these can help diversifying macroeconomic risk of a traditional 60/40 asset allocation.

Identifying macro factors

There are generally two distinct approaches for building out macroeconomic factors.[1] The first focuses on pure macroeconomic state variables that can be considered as ultimate drivers of co-movement in asset returns, as in Chen, Roll and Ross (1986). Common macroeconomic state variables are output (to measure growth), inflation, interest rates and risk aversion. However, the explanatory power regarding the returns of many asset classes proves to be modest, complicating the actual implementation of corresponding macro factor-based portfolio allocations.

The second approach focuses directly on the factors' ability to explain the cross section of different asset classes' returns. A common statistical methodology to achieve this objective is to run a principal component analysis (PCA) to derive the salient factors explaining most of the asset classes' return variation. In addition, this procedure often creates investable factors. For instance, a PCA typically identifies a portfolio of equities and other risk assets as the most important factor proxying for macroeconomic growth. Similarly, macro factor portfolios representing real rates or inflation risk emerge. Allowing for more granularity in the underlying asset class returns, one may also identify macro risk factors representing commodity, credit, emerging market or currency risk; see Greenberg, Babu and Ang (2016) among others.

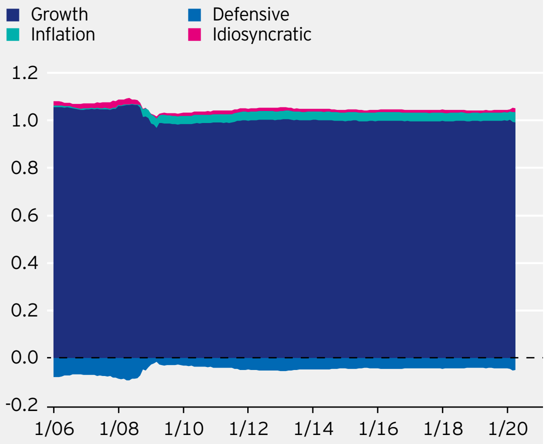
> The first factor is growth. The second factor is defensive. The third factor relates to inflation.

To examine the role of macroeconomic factors in portfolio management, we build on the above evidence and focus on three factors in particular: the first factor is growth, as measured by broad equity market exposure. The second factor is defensive, which we proxy by investing in US Treasuries. The third factor relates to inflation and is measured by the spread between inflation-linked bonds and US Treasuries.

Traditional asset allocation through the macro factor lens

To illustrate the relevance of macroeconomic factors, we X-ray a traditional asset allocation in terms of a risk model governed by these three macroeconomic factors. We particularly look into a 60/40 portfolio in global equities and bonds. The 60% equity allocation is represented by the MSCI ACWI index, and the 40% bond allocation splits into 30% in investment grade and 10% high yield bonds. Figure 1 decomposes its portfolio volatility from 2006 to 2020 into macro factor contributions and shows growth risk to be the biggest (if not sole) contributor to portfolio risk. In the following, we seek to reduce this obvious vulnerability through an allocation process that acknowledges macro factor sensitivities.

Figure 1
Macro factor risk decomposition of a 60/40 stock-bond portfolio



The chart decomposes the volatility of a 60/40 stock-bond allocation into macroeconomic factor contributions.
Sources: Invesco, Bloomberg, Goldman Sachs.
Sample period: 31 January 2006 to 31 May 2020.

Building out diversified macro factor-mimicking portfolios

Asset and style factor data

We wish to investigate the macroeconomic factor sensitivities of a broad set of asset classes and style factors. In each asset class, we aim to be as granular as possible in teasing out the differential element of a given investment. That is, next to broad world equity exposure, we are interested in the returns of certain regions (US, EAFE, EM) relative to the world equity market. Similarly, we look at long-short style factor returns for value, momentum, quality and low volatility investments, isolating the pure factor premia. For fixed income assets, we use US 10Y Treasuries to proxy for the market return and add TIPS, investment grade and high yield corporate bond spreads, as well as emerging market credit spreads. Similar to equities, the factor investing literature supports the notion of fixed income style factors (Kothe, Lohre and Rother, 2021), and we include the four rates factors: quality, value, momentum and carry.

Given the heterogeneity of commodities as an asset class, we abstain from utilizing a broad market index, as these commonly suffer from an extreme energy risk allocation; see Bernardi, Leippold and Lohre (2018) among others. Instead, we investigate the properties of four commodity sectors (precious metals, industrial metals, energy and agriculture) that show little correlation across sectors. We also consider long-short commodity factors along the dimensions carry, value, momentum and quality. Lastly, we include currency investments by allowing two currency baskets, representing the currency allocations implicit in the MSCI EAFE and the MSCI Emerging Markets indices, respectively. We also investigate the three salient currency investment styles carry, value and momentum.

A regime-based route to macro factor-mimicking portfolios

There are different techniques to determine the macroeconomic nature of assets and style factors. For instance, a simple statistical clustering of the multi-asset multi-factor data can help in assembling feasible sets to proxy for a given macroeconomic factor.[2] Another common alternative is to inspect macroeconomic factor sensitivities from multivariate factor regressions. Here, we instead pursue an innovative route that leverages insights from analysis of assets and factors in different economic regimes. As the two decades of multi-asset multi-factor data see a high correlation of growth and inflation assets, we believe such analysis to be vital in identification of genuine growth or inflation assets and factors.

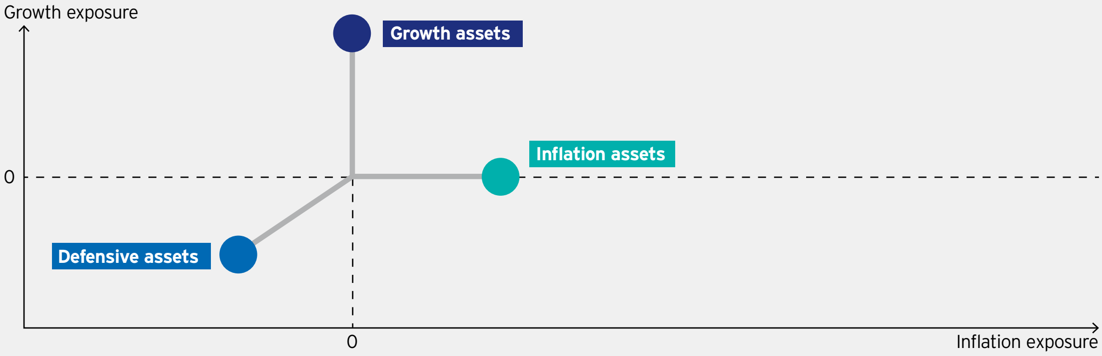
Therefore, to get a sense of how different assets and style factors perform under distinct growth-inflation

Table 1
Determining macro factor-mimicking portfolios

Assets and factors	Rising Growth + Rising Inflation	Rising Growth + Falling Inflation	Falling Growth + Rising Inflation	Falling Growth + Falling Inflation	Growth Exposure	Inflation Exposure	Growth MFMP	Inflation MFMP	Defensive MFMP
Equities									
ACWI	4.77	4.44	-3.91	-3.76	8.44	0.09	2.5%		
US-ACWI	0.77	0.38	-0.60	0.33	0.71	-0.27			16.8%
EAFE-ACWI	-0.74	-0.05	-0.18	-0.35	-0.13	-0.26			18.2%
EM-ACWI	0.60	-0.04	1.23	-0.08	-0.29	0.98		9.3%	
Cyclicals-Defensives	1.58	0.91	-2.03	-1.93	3.22	0.29	11.4%		
Quality	0.59	0.60	0.68	1.92	-0.70	-0.63			29.7%
Momentum	-0.02	0.50	1.30	1.54	-1.18	-0.38			14.9%
Value	-0.51	-0.27	-0.44	-0.86	0.26	0.09			20.4%
Low Volatility	-0.42	0.55	1.50	1.79	-1.57	-0.63			29.8%
Fixed Income									
US 10Y Tsy	-0.75	0.66	1.44	2.17	-1.85	-1.07			30.5%
TIPS	1.10	-0.05	3.42	-0.02	-1.18	2.29		20.9%	
IG Credit	2.25	1.54	-0.38	-2.24	3.21	1.28	16.4%		
HY Credit	2.97	2.18	-0.95	-2.50	4.30	1.16	11.8%		
EM Credit	2.22	1.24	-1.20	-2.64	3.65	1.21	8.5%		
Rates Value	0.44	1.11	-0.12	-0.80	1.24	0.00	12.5%		
Rates Momentum	0.13	0.33	2.18	1.53	-1.63	0.23			17.6%
Rates Quality	-0.48	-0.46	1.81	0.98	-1.86	0.41			26.2%
Rates Carry	0.80	0.93	1.02	0.10	0.31	0.40		13.5%	
Commodities									
Precious Metals	0.90	-0.42	1.04	-0.43	-0.06	1.40		5.5%	
Industrial Metals	1.53	0.45	0.11	-3.28	2.57	2.24		2.8%	
Energy	1.19	-1.08	0.68	-2.97	1.20	2.96		4.7%	
Agriculture	0.69	0.28	-0.44	-1.54	1.47	0.75		3.1%	
Carry	0.81	-0.29	1.09	0.10	-0.33	1.05		9.9%	
Quality	1.25	1.44	1.59	2.20	-0.55	-0.40			32.9%
Momentum	0.14	-0.02	0.97	-0.37	-0.24	0.75		6.5%	
Value	0.73	0.56	0.75	0.62	-0.04	0.15			8.5%
Currencies									
Developed Markets	0.99	-0.58	0.52	-1.35	0.62	1.72		16.8%	
Emerging Markets	2.13	1.05	0.01	-2.71	2.94	1.90	9.8%		
Carry	1.90	1.58	-0.04	-1.14	2.33	0.71	10.7%		
Value	-0.07	0.87	1.08	1.00	-0.64	-0.43			20.8%
Momentum	0.69	0.28	1.35	0.28	-0.33	0.74		15.1%	

The first section of the table gives risk-adjusted performance (Sharpe ratios) of the asset classes and style factors in four different growth-inflation regimes. Columns 5 and 6 synthesize this information into average growth and inflation exposures. The last section of the table shows macro factor-mimicking portfolio weights for the three macro factors: growth, inflation and defensive. For proxies used, please refer to the data appendix at the end of the article.
Sources: Invesco, Bloomberg, Goldman Sachs. Sample period from 31 January 2001 to 31 May 2020. **Past performance is not a guide to future returns.**

Figure 2
Classifying assets and factors based on growth and inflation exposures



Source: Invesco. For illustrative purposes only.

regimes, we divide the sample data in four regimes using the monthly returns of the simple growth and inflation factors as a delimiter. The rising growth/rising inflation regime comprises all months in which both growth and inflation assets rise. All remaining months fall into the other three regimes (table 1), formed by considering the alternative combinations of growth and inflation returns.[3]

Imagine we expect rising inflation and falling growth. Column 3 shows which assets and factors provided inflation protection in such a regime. Clearly, TIPS stand out with the highest Sharpe ratio. While we would expect good inflation hedge properties for all commodity sectors, this specific regime favors metals and energy. As for the three credit asset classes, we note that their positive correlation to inflation is mostly driven by their proximity to growth assets. However, this relation breaks down in negative growth environments, when credit markets failed to provide inflation protection. Hence, these asset classes clearly can be considered in the growth bucket alongside equity exposure.

Regarding style factors, we observe consistent inflation hedging returns for commodity carry and momentum, as well as FX momentum. To systematically pin down genuine growth, defensive or inflation assets, we use a straightforward procedure to determine the average growth and inflation exposure based on evaluating the differential performance of assets and factors across the various growth-inflation regimes. To illustrate, for an asset (or style factor) to be considered a growth asset, we would expect it to have higher risk-adjusted performance in positive versus negative growth regimes. Specifically, we would wish to observe such outperformance in inflationary and deflationary periods. Hence we define an asset's average growth exposure as the growth spread in risk-adjusted performance averaged across the two inflation regimes. Conversely, inflation assets are expected to do well in inflationary periods, independent of the prevailing growth regime. We thus define an asset's average inflation exposure as the inflation spread in risk-adjusted performance as a simple average across the two possible growth regimes. Table 1 gives these average growth and inflation exposures for all assets and style factors.

With this information, we can plot all assets and style factors on the growth and inflation dimensions. As shown in figure 2, defensive assets would ideally have negative loadings on both growth and inflation; inflationary assets would have zero exposure to growth and large positive loadings on inflation; and growth assets would have zero exposure to inflation and large positive loadings on growth. In practice, many assets will not fit cleanly into one of these three areas, but we at least have clear priors about what constitutes an ideal asset in each macro factor.

We operationalize this idea using simple parameters to define an area for each macro factor. More sophisticated approaches are certainly available, but in this case, we want to illustrate the usefulness of macro factors even with fairly simple definitions. For example, consistent with figure 2, an asset is labeled 'growth' if it is closest to the growth coordinates of the asset or factor with the highest growth exposure. Figure 3 illustrates this procedure using growth and inflation exposures computed over the full sample period. High yield credit has the highest growth exposure and the latter forms the center of the blue area, where growth assets and factors are located.[4] Similarly, energy assets have the highest inflation exposure, forming the center of the purple area, which contains inflation assets and factors. Lastly, the center of the turquoise defensive assets and factors area is determined by the asset with the smallest growth exposure, i.e. 10-year US Treasuries.

Such an approach gives rise to intuitively appealing classifications. For instance, the basket of inflation assets and factors features TIPS, all commodity assets but also a few style factors, such as commodity carry and momentum or FX momentum and rates carry. As for growth, the equity and credit assets are joined by cyclical versus defensive sectors, EM currencies and two style factors, rates value and FX carry, which resonates with the latter suffering in similar periods like equities.

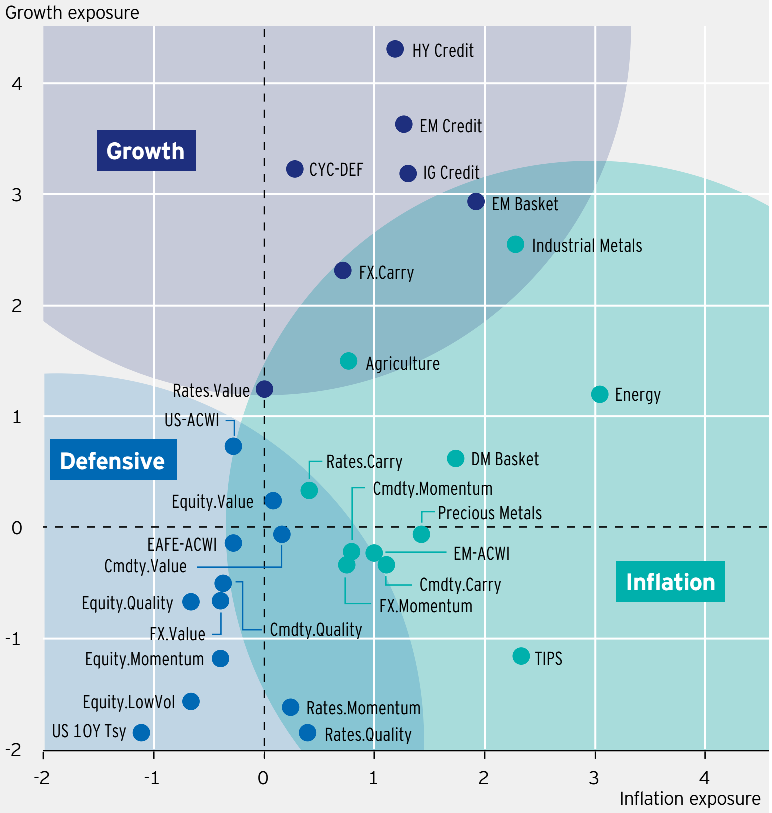
Interestingly, the defensive basket features a larger number of style factors, including almost all quality style factors as well as equity momentum and low volatility. Note that our procedure assigns every asset and factor to one of the three macroeconomic factors.

Based on this macroeconomic classification of asset and style factors, we build three macro factor-mimicking portfolios (MFMPs) that can help guide macro factor-based portfolio allocations. Obviously, the sensitivity of an asset to a certain macro factor decreases with the distance from the respective macro factor's center.

Applying this classification over time through an expanding window, we observe that, while some assets and factors can be clearly associated to one of the three macro factors, others might be reasonably close to more than just one factor. It seems natural to apply less weight to such distant assets and factors when constructing macro factor-mimicking portfolios. Also, we wish to diversify identified macro baskets in terms of risk and therefore apply a straightforward and robust weighting scheme. Specifically, we perform an inverse volatility allocation where the assets' and factors' volatilities are scaled according to their relevance for the macro factor concept. That is, a more distant asset or factor will experience a more severe volatility penalty than a very close one. As a result, the macro factor-mimicking portfolios focus on truly representative assets and factors rather than being unduly dominated by weaker contenders.

The specific constituents and weights for all three MFMPs are shown in the last three columns of table 1; the weights are scaled such that all MFMPs target a volatility of 5%. These portfolios each represent pure exposure to either growth or inflation or defensiveness and thus form meaningful instruments to navigate portfolios through a macro factor lens. Figure 4 shows the macro factor risk decomposition of the defensive and inflation portfolios, suggesting that both MFMPs live up to their respective objective.

Figure 3
Determining macro factor-mimicking portfolios

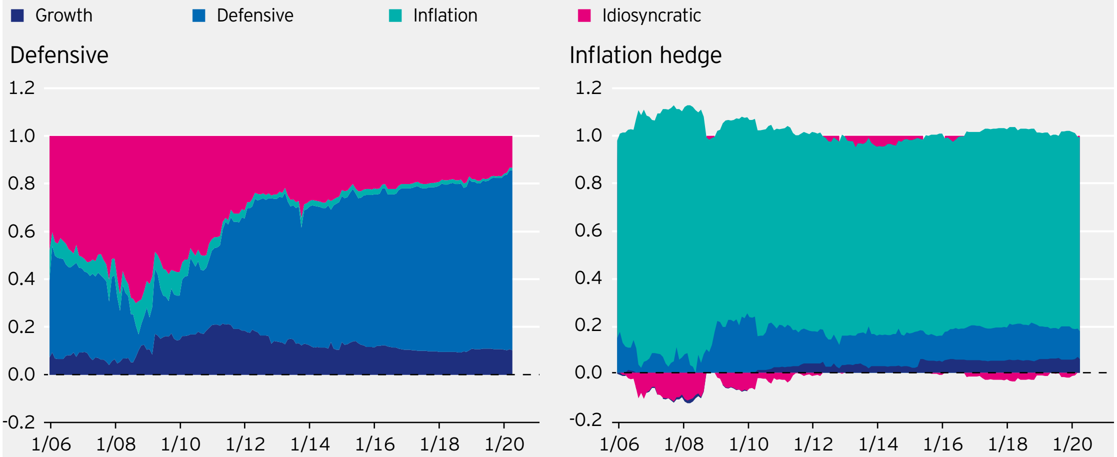


The chart depicts asset and style factors in a scatter plot according to their average inflation exposure (x-axis) and average growth exposure (y-axis).
Source: Invesco. Sample period from 31 January 2001 to 31 May 2020.

Macro factor-based portfolio overlays

We now make use of the macro factor-mimicking portfolios. Circling back to the concentrated growth risk allocation of the 60/40 stock-bond allocation, we explore ways of altering the risk profile. First, we add a defensive overlay to reduce growth risk in a 60/40 portfolio. Second, we add an inflation hedge to help protect against an increase in inflation. Third, we consider the effect of combining defensive and inflation hedge overlays with the 60/40 allocation.

Figure 4
Macro factor portfolios – defensive and inflation hedge



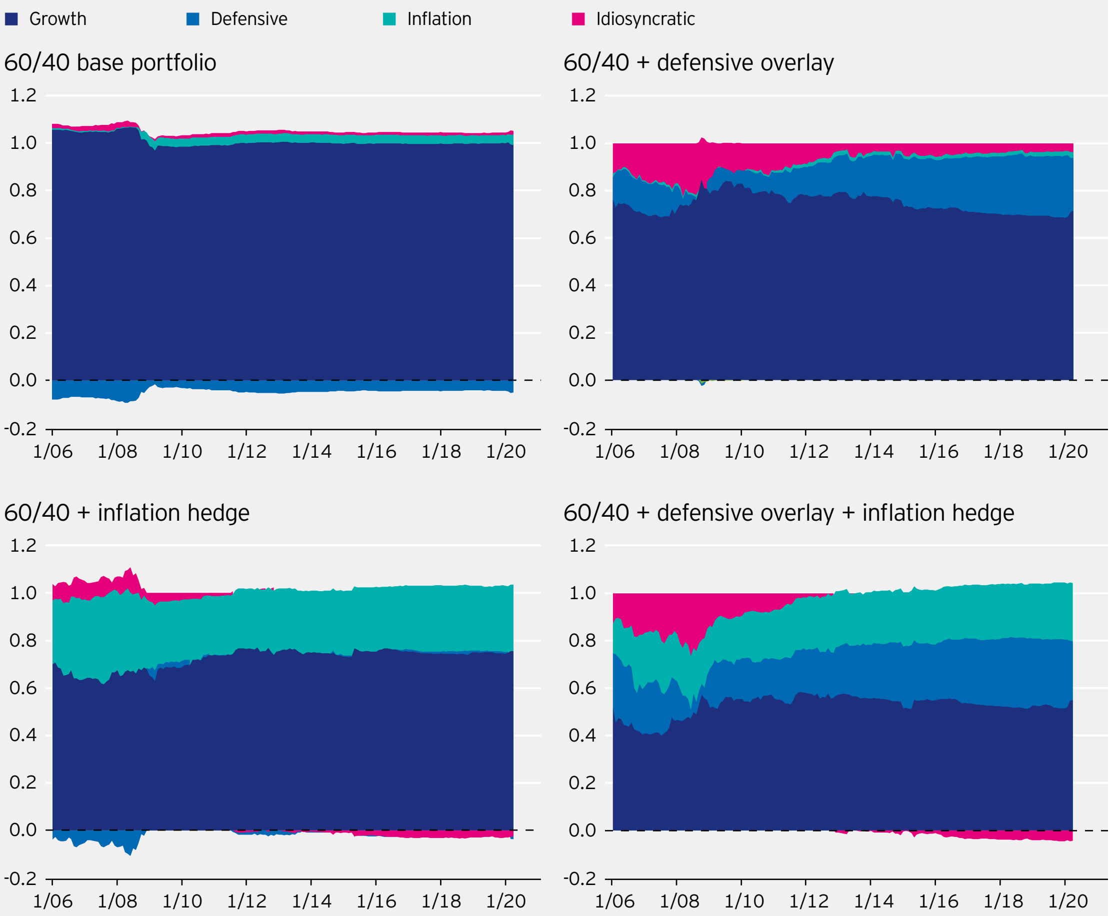
The chart decomposes the volatility of the macro factor-mimicking portfolio for defensive (left) and inflation (right) by macroeconomic factor contributions.
Sources: Invesco, Bloomberg, Goldman Sachs. Sample period 31 January 2006 to 31 May 2020.

Adding a defensive overlay
To keep the analysis simple, we add the defensive MFMP to the 60/40 stock-bond allocation using the exact defensive MFMP weights given in table 1. The addition of the defensive MFMP comes with a noticeable reduction in growth risk, as we can infer from the macro factor risk decomposition in figure 4 (upper right). Moreover, this addition has a considerable impact on the ensuing portfolio's risk-return profile. In the absence of the defensive overlay, the 60/40 portfolio operates at a 10% volatility level (see table 2). Given a Sharpe ratio of 0.54, it delivered some 6.2% annualized return over the sample period. Adding the defensive overlay barely affects the volatility level (which is slightly down to 9.44%) but crucially mitigates tail risk; maximum drawdown is considerably cut (-25.58%), which represents a reduction of more than 10 percentage points relative to that of the 60/40 portfolio (-36.54%). As a result, the annualized return is almost twice as high as that of the 60/40 (11.42% versus 6.22%).

Adding a diversified inflation hedge overlay
The effect of adding an inflation hedge clearly shows in the macro factor risk decomposition (figure 5, lower left). However, the inflation hedge portfolio is not a source of extra return in the backtest period. The combination with the 60/40 slightly raises volatility and tail risk due to the consideration of commodity assets. Therefore, we also provide performance statistics of a strategy variant that scales the full allocation such that the ensuing volatility is comparable to the one of the 60/40 base allocation. The annualized return of the scaled strategy is 3.97%, which is 225bp below the 60/40 portfolio's return. Still, the drawdown is likewise severe and comes in at -38.15%. Obviously, one would need to consult a longer history to better gauge the actual benefit of inflation hedging, as the considered sample period is lacking pronounced inflationary regimes.

Diversifying growth risk through combining defensive and inflation hedge overlays
Given the difficulties in predicting the economic environment, diversifying macroeconomic factor risk seems a natural path to follow. We thus investigate adding both overlays, defensive and inflation hedge, to the 60/40 allocation. First, we observe a fairly balanced risk profile through time, where growth risk ceases to be dominant most of the time. From a performance perspective, there is a slight increase in volatility risk but still a considerable decrease in tail

Figure 5
Diversifying a traditional asset allocation with macro factor overlays



The figure decomposes the volatility of a 60/40 stock-bond allocation by macroeconomic factor contributions (upper left). The remaining charts refer to 60/40 stock-bond allocations amended by a defensive overlay (upper right), an inflation hedge (lower left), as well as a combination of the two (lower right).
Sources: Invesco, Bloomberg, Goldman Sachs. Sample period 31 January 2006 to 31 May 2020.

Table 2
Performance statistics for macro factor-based allocations

	Defensive MFMP	Inflation hedge	60/40	60/40 + Defensive	60/40 + Inflation Hedge	60/40 + Defensive + Inflation Hedge	60/40 + Defensive (Scaled)	60/40 + Inflation Hedge (Scaled)	60/40 + Defensive + Inflation Hedge (Scaled)
Net return p.a.	5.92%	0.59%	6.22%	11.42%	5.48%	10.92%	11.91%	3.97%	8.96%
Volatility p.a.	4.41%	3.93%	10.04%	9.44%	12.18%	11.56%	10.33%	10.36%	10.14%
Sharpe Ratio	1.07	-0.12	0.54	1.08	0.41	0.86	1.04	0.32	0.79
Max Drawdown	-6.19%	-12.75%	-36.54%	-25.28%	-39.35%	-30.45%	-28.77%	-38.15%	-30.66%
Calmar Ratio	0.96	0.09	0.17	0.45	0.14	0.36	0.41	0.1	0.29

The table provides simulated performance figures for macro factor-based multi-asset multi-factor strategies from the perspective of a US-dollar investor. This model does not factor in all economic and market conditions that can impact results.
Source: Invesco, Bloomberg, Goldman Sachs. Period from January 2006 to May 2020. **The figures refer to simulated past performance and past performance is not a reliable indicator of future performance.**

Table 3
Macro factor allocations – performance by regime

	Rising Growth + Rising Inflation	Rising Growth + Falling Inflation	Falling Growth + Rising Inflation	Falling Growth + Falling Inflation
Growth	10.9%	4.8%	-2.8%	-17.8%
Defensive Overlay	1.2%	5.6%	8.8%	14.1%
Inflation Hedge	4.8%	-2.3%	4.7%	-8.0%
60/40	26.9%	17.5%	-18.9%	-29.5%
60/40 + Defensive Overlay	27.1%	21.4%	-11.5%	-16.3%
60/40 + Inflation Hedge	30.7%	13.6%	-15.5%	-38.3%
60/40 + Defensive + Inflation	31.2%	17.7%	-7.8%	-25.0%
60/40 + Defensive Overlay (scaled)	29.0%	23.5%	-13.2%	-18.4%
60/40 + Inflation Hedge (scaled)	24.7%	11.8%	-14.0%	-32.6%
60/40 + Defensive + Inflation (scaled)	26.5%	15.8%	-7.4%	-22.5%

The table displays annualized excess returns of several macro factor allocations performance under four different growth-inflation regimes.
Source: Invesco, Bloomberg, Goldman Sachs. Sample period from January 2006 to May 2020. **The figures refer to simulated past performance and past performance is not a reliable indicator of future performance.**

risk. The annualized return is almost 50% higher than in the 60/40 base case, resulting in a Sharpe ratio of 0.79. Also, the drawdown is reduced by some 6 percentage points to -30.66%.

Lastly, we investigate how the different macro factor strategies perform in the four growth-inflation regimes defined earlier; see table 3. Given its concentration in growth risk, we find the 60/40 portfolio outperforming in rising growth environments and underperforming when growth falls. As expected, the defensive macro factor-mimicking portfolio is particularly beneficial in both negative growth regimes. In fact, the regime-specific performance analysis highlights that the 60/40 with defensive overlay is on par with the pure 60/40 in rising growth environments, but considerably better in the two falling growth periods.

As for the efficacy of the inflation hedge, we observe that the corresponding macro factor-mimicking portfolio indeed earns positive returns in inflationary regimes and negative returns in deflationary regimes. Adding this inflation hedge to the 60/40 allocation, we note that the rising growth/falling inflation regime sees similar returns for this and the base portfolio. Yet, under the falling growth/rising inflation regime, we observe a return benefit for the inflation-hedged strategy. Judging by the scaled 60/40 + inflation hedge, the regime return is half that of the pure 60/40 portfolio (-14.0% vs. -18.9%). Obviously, if one is expecting a falling growth/rising inflation period, enhancing the 60/40 allocation through the addition of defensive and inflation hedge overlays is the method of choice, as demonstrated by a historic regime-specific return of -7.4% (based on the scaled version).

Building out macro factor-mimicking portfolios from a straightforward growth-inflation regime perspective enables effective diversification.

Conclusion

Style factors are often considered meaningful diversifiers and can help achieve various investor objectives; see Dichtl, Drobetz, Lohre and Rother (2021). In this article, we have looked at such investments through the overarching lens of macro factors that ultimately govern the dynamics of asset class and style factor returns. Building out macro factor-mimicking portfolios from a straightforward growth-inflation regime perspective enables effective diversification of traditional stock-bond allocations versus growth and inflation risks.

References

Amato, L., and H. Lohre (2020): Diversifying macroeconomic factors – for better or for worse, Working paper.

Bernardi, S., M. Leippold and H. Lohre (2018): Maximum diversification strategies along commodity risk factors, European Financial Management, 24(1), 53-78.

Chen, N., R. Roll and S.A. Ross (1986): Economic forces and the stock market, Journal of Business, 59(3), 383-403.

Dichtl, H., W. Drobetz, H. Lohre and C. Rother (2021): Active factor completion strategies, Journal of Portfolio Management, forthcoming.

Greenberg, D., A. Babu and A. Ang (2016): Factors to assets: Mapping factor exposures to asset allocations, Journal of Portfolio Management, Special Issue, 18-27.

Kothe, J., H. Lohre, and C. Rother (2021): Rates factors and global asset allocation, Journal of Fixed Income, forthcoming.

Notes

1 See Amato and Lohre (2020) for a comparison of associated diversified macro factor risk parity strategies.
2 Performing statistical clustering based on the multi-asset multi-factor dataset underlying this study, we find support for the relevance of the three factors growth, inflation and defensive; results available upon request.
3 Note that the rising growth / rising inflation regime prevails in about 40% of our sample period while the remaining three regimes each prevail in some 20% of the months. Hence, the derived statistics are not based on overly thin data.
4 By construction, the MSCI ACWI comes with inflated growth exposures and is thus not a suitable anchor.

About the authors

Dr. Harald Lohre
Director of Research
Invesco Quantitative Strategies
Harald Lohre and his team are responsible for maintaining and evolving the quantitative models that drive the investment decisions within multi-factor equity and balanced investment products.

Scott Hixon, CFA®
Portfolio Manager and Head of Research
Invesco Global Asset Allocation
Scott Hixon is a Portfolio Manager and the Head of Research for Invesco's Global Asset Allocation team, which invests in stock, bond and commodity markets worldwide. Mr. Hixon oversees and helps steer the team's research initiatives in the areas of model and strategy development, as well as portfolio construction.

Jay Raol, Ph.D.
Head of Fixed Income Factors
Systematic & Factor Investment Group
Jay Raol and his team researches and manages systematic and factor based strategies in global fixed income and currency markets.

Alexander Swade
Ph.D. Candidate, Lancaster University, and Invesco Quantitative Strategies
In a joint research initiative between Lancaster University and Invesco Quantitative Strategies, Alexander Swade investigates factor-based investment strategies across different asset classes and within equities.

Hua Tao, Ph.D., CFA®
Research Analyst
Invesco Global Asset Allocation
Hua Tao is a Research Analyst for the Invesco Global Asset Allocation team, responsible for model research and strategy development.

Scott Wolle, CFA®
Head of Systematic & Factor Investing
Invesco
Scott Wolle serves as the Head of Systematic & Factor Investing, which includes equity, fixed income, and macro strategies along with custom indices.

Data appendix

Name	Description	Ticker	Source	Construction details
Equities				
ACWI	MSCI ACWI Net TR Local index	NDLEACWF	Bloomberg	
US-ACWI	MSCI USA TR USD Index minus MSCI ACWI Net TR Local Index	NDDLUS, NDLEACWF	Bloomberg	NDDLUS - NDLEACWF
EAFE-ACWI	MSCI EAFE TR LCL Index minus MSCI ACWI Net TR Local Index	NDDLEAFE, NDLEACWF	Bloomberg	NDDLEAFE - NDLEACWF
EM-ACWI	MSCI EM TR LCL Index minus MSCI ACWI Net TR Local Index	NDLEEGF, NDLEACWF	Bloomberg	NDLEEGF - NDLEACWF
Cyclicals-Defensives	ACWI CYCLICAL SECTORS- ACWI DEFENSIVE SECTORS	MXCXDRN	Bloomberg	
Quality	MSCI ACWI Quality USD minus MSCI ACWI Net TR USD Index	M1WDQU, NDUEACWF	Bloomberg	M1WDQU - vol.adj. NDUEACWF
Momentum	MSCI ACWI Momentum USD minus MSCI ACWI Net TR USD Index	M1WD000$, NDUEACWF	Bloomberg	M1WD000$ - vol.adj. NDUEACWF
Value	MSCI ACWI Value USD minus MSCI ACWI Net TR USD Index	M1WD000V, NDUEACWF	Bloomberg	M1WD000V - vol.adj. NDUEACWF
Low Volatility	MSCI ACWI Minimum Volatility USD minus MSCI ACWI Net TR USD Index	M00IWD$O, NDUEACWF	Bloomberg	M00IWD$O - vol.adj. NDUEACWF
Fixed Income				
US 10Y Tsy	Bloomberg Barclays US Treasury Bellwethers 10 Year TR Index Value Unhedged USD	BW10TRUU	Bloomberg	
Cash	USD 3 Month T-Bill	USGG3M	Bloomberg	
TIPS	US TIPS TR	I01551US	Bloomberg	
IG Credit	Bloomberg Barclays US Agg Corp excess return	LUACER	Bloomberg	
HY Credit	Bloomberg Barclays US Corporat HY excess return	LF98ER	Bloomberg	
EM Credit	J.P. Morgan EMBI Global TR minus US Treasury	JPEIGLBL, LUATTRUU	Bloomberg	JPEIGLBL - vol.adj. LUATTRUU
Rates Value	Goldman Sachs Rates Value Strategy	GSIRVA03	GS	
Rates Momentum	Goldman Sachs Rates Momentum Strategy	GSIRTR03	GS	
Rates Quality	Goldman Sachs Rates Quality Strategy	GS Interest Rates Curve C0210	GS	
Rates Carry	Goldman Sachs Rates Carry Strategy	GSIRCA03	GS	
Commodities				
Precious Metals	Bloomberg Precious Metals Subindex	BCOMPR	Bloomberg	
Industrial Metals	Bloomberg Industrial Metals Subindex	BCOMIN	Bloomberg	
Energy	Bloomberg Energy Subindex	BCOMEN	Bloomberg	
Agriculture	Bloomberg Agriculture Subindex	BCOMAG	Bloomberg	
Carry	Goldman Sachs Commodity Carry Strategy	GS Macro Carry Index RP14	GS	
Quality	Bloomberg Roll Select Commodity Index minus Bloomberg Commodity Index	BCOMRST, BCOMTR	Bloomberg	BCOMRST - BCOMTR
Momentum	Goldman Sachs Commodity Momentum Strategy	GS Macro Momentum Index RP15	GS	
Value	Goldman Sachs Commodity Value Strategy	GS Commodity COT Strategy COT3	GS	
Currencies				
Developed Markets	MSCI EAFE Currency USD Index	MXEA0CX0	Bloomberg	
Emerging Markets	MSCI Emerging Markets Currency USD Index	MXEF0CX0	Bloomberg	
Carry	Goldman Sachs FX Carry Strategy	GS FX Carry C0115	GS	
Value	Goldman Sachs FX Value Strategy	GS FX Value C0114	GS	
Momentum	Goldman Sachs FX Momentum Strategy	GS FX Trend C0038	GS	
Macro factors				
Growth	MSCI ACWI Net TR Local index	NDLEACWF	Bloomberg	
Defensive	Bloomberg Barclays US Treasury Bellwethers 10 Year TR Index Value Unhedged USD	BW10TRUU	Bloomberg	
Inflation	USTreasuryTIP minus BBUSTreasury	SPBDUP3T, LT01TRUU	Bloomberg	SPBDUP3T - LT01TRUU

About risk

The value of investments and any income will fluctuate (this may partly be the result of exchange rate fluctuations) and investors may not get back the full amount invested.

Important information

This document has been prepared only for those persons to whom Invesco has provided it for informational purposes only. This document is not an offering of a financial product and is not intended for and should not be distributed to retail clients who are resident in jurisdiction where its distribution is not authorized or is unlawful.. Circulation, disclosure, or dissemination of all or any part of this document to any person without the consent of Invesco is prohibited.

This document may contain statements that are not purely historical in nature but are "forward-looking statements," which are based on certain assumptions of future events. Forward-looking statements are based on information available on the date hereof, and Invesco does not assume any duty to update any forward-looking statement. Actual events may differ from those assumed. There can be no assurance that forward-looking statements, including any projected returns, will materialize or that actual market conditions and/or performance results will not be materially different or worse than those presented.

The information in this document has been prepared without taking into account any investor's investment objectives, financial situation or particular needs. Before acting on the information the investor should consider its appropriateness having regard to their investment objectives, financial situation and needs.

You should note that this information:

- may contain references to amounts which are not in local currencies;
- may contain financial information which is not prepared in accordance with the laws or practices of your country of residence;
- may not address risks associated with investment in foreign currency denominated investments; and
- does not address local tax issues.

All material presented is compiled from sources believed to be reliable and current, but accuracy cannot be guaranteed. Investment involves risk. Please review all financial material carefully before investing. The opinions expressed are based on current market conditions and are subject to change without notice. These opinions may differ from those of other Invesco investment professionals.
The distribution and offering of this document in certain jurisdictions may be restricted by law.

Persons into whose possession this marketing material may come are required to inform themselves about and to comply with any relevant restrictions. This does not constitute an offer or solicitation by anyone in any jurisdiction in which such an offer is not authorised or to any person to whom it is unlawful to make such an offer or solicitation.