

Big data & AI pose many challenges for quant investors

- Big data and AI could enable investors to take better decisions
- Data fitting and short histories raise robustness concerns
- We apply our cautious pioneering philosophy to this field

The advent of big data and artificial intelligence (AI) has emerged as a major game-changer for the financial industry. In the asset management sector, although the adoption of these innovations is still in its infancy, a growing number of players are examining how they can be used to design better investment strategies. This note explains **Robeco's** views on big data and AI, and how we already use large datasets in our investment strategies.

Defining big data and AI

Big data is a term for which there is no single consensual definition. The concept of big data usually describes the collection and analysis of large and complex information datasets, often from unstructured and changing sources, that does not fit on the hard disk of a conventional computer. Such data is increasingly being used by companies to analyze and predict the behavior of economic agents, mostly consumers.

Meanwhile, Artificial intelligence, or AI, can be defined as the use of computational tools to perform tasks that traditionally required human thinking. As a scientific field of research, AI is far from new. The term was actually coined in the mid-1950s by computer scientist John



David Blitz, Head of Quant Research, and Rob van Bommel, Portfolio Manager

'The adoption of these innovations is still in its infancy'

McCarthy, an assistant professor at Dartmouth College at the time. However, recent improvements in computational power and the dramatic surge in the amount of data available in the digital age have increased the scope of potential applications for these technologies considerably.

Big data and artificial intelligence are already being used in many industries, such as healthcare and the automotive industry. Although many asset managers have publicly embraced the use of big data and artificial intelligence, their focus has so far mostly remained on areas such as process automation, or sales and marketing. Other areas, in particular investments, still stand to benefit broadly from this kind of innovation. Only a limited number of asset managers,¹ most of them large hedge funds, have publicly announced they were implementing this kind of approach for their investment strategies, and the amount of assets managed solely based on these techniques remains relatively small at this stage.²

A critical look at big data and AI

While it is true that improvements in data analytics may bring valuable insights for investors, there is little hard evidence that big data and artificial intelligence actually work in practice, at least for now. Concrete cases of investment managers able to deliver consistent outperformance in real life using these techniques are still absent.

A solid investment strategy requires extensive empirical testing and falsification on broad samples of data and over long periods of time, but the evidence for big data and artificial intelligence is largely anecdotal. Strategies based on big data and artificial intelligence may also lack the necessary backing by a clear economic rationale. In fact, most investment ideas solely rely on paper back-test results, which should always be considered with caution.

These different elements are hard to reconcile with a prudent, evidence-based investment philosophy, such as Robeco's. We always strive to identify factors that are rewarded with superior risk-adjusted performance over the longer term. We also look beyond mere statistical patterns and aim to understand the economic drivers of returns.

As a result, managing money based on an untested AI algorithm that scrutinizes exotic datasets raises a lot of concerns, even though big data and artificial intelligence have become a popular discussion topic among investors. Moreover, these innovations pose a

¹ A. Satariano and N.Kumar, 'The Massive Hedge Fund Betting on AI', Bloomberg, September 2017.

² A recent report by the Financial Stability Board, indicated that 'pure' artificial intelligence and machine learning asset managers had only about USD 10 billion in assets under management, according to an investor focused on AI and machine learning funds. But the report also acknowledged that this figure is growing rapidly. See: 'Artificial intelligence and machine learning in financial services - Market developments and financial stability implications', Financial Stability Board, November 2017.

number of technical challenges for asset managers looking to incorporate them into their investment process.

The first of these technical challenges is that 'big data datasets' generally have a short or even very short history. The social network Twitter, for example, was created in March 2006 and it did not become widely used before 2010-2011. Moreover, the number of monthly active users reported for the last quarter of 2017 – 330 million people worldwide on average – is huge, but this number was only about six million a decade ago. As a result, this dataset can hardly be tested over a single full market cycle. For many other new datasets the available history is even shorter.

The second issue has to do with the lack of breadth, or at least the very fragmentary nature, of most big data signals. A good illustration of this is satellite or drone imagery that are used to track the number of car in the parking slots of supermarkets. Although many people would acknowledge the use of this kind of information as an appealing idea for an investment strategy, the lack of breadth of the data series makes it of limited use for global investors: for the vast majority of listed stocks this kind of data is meaningless or not available.

Another concern is that many signals provided by big data and AI tend to be very short-term focused. Useful trading signals often follow a decay function over time, as data are more widely used and hence become less valuable for gaining an edge over other investors. This has obviously limited value for institutional investors with a long-term focus.

Such lack of length and breadth and reliability in the data samples that can be analyzed raises serious concerns in terms of potential risk of data mining or 'p-hacking'. For instance, any quantitative investment strategy based on Twitter feed data would necessarily have to be back tested against a bull market backdrop only. Meanwhile rigorous back testing requires to look at how a strategy would have operated over multiple market cycles.

Finally, high quality datasets are not easy to obtain and can be very expensive. In fact, providing such data has turned into a lucrative business in recent years. A growing number of specialist firms are making available to asset managers, in particular hedge funds, novel data feeds as well as machine learning tools to gain insight from the vast volumes of news and market research available in financial markets. Meanwhile, some asset managers are also building themselves indicators using either significant in-house research resources, or using expensive tools supplied by third parties.

All these different issues pose serious challenges for financial research carried out on more conventional databases. In recent years, for example, many prominent academics, including Duke university professor and former president of the American Finance Association (AFA),

'Big data & AI pose technical challenges for managers looking to incorporate them into their investment process.'

Campbell Harvey,³ have been vocally warning against ‘p-hacking’ and the fact that a significant proportion of all empirical studies being published in academic journals are unlikely to hold up in the future. In the case of big data and AI, these threats to rigorous scientific analysis are even more acute, as there are many more degrees of freedom in the modeling process.

Given all the caveats mentioned above, we consider the current trend around big data and AI as a very interesting development, but one that should be treated with caution. Factors such as value, momentum, quality and low risk have been extensively studied for decades. Over the years, ample empirical evidence that these factors exist over time, over the cross-section and across asset classes has been accumulated. Researchers have also been able to find credible economic explanations of why the effect is there and why it should persist in the future.

What we are doing

As mentioned in the previous section of this note, many examples of big data information are not, or at least not yet, very useful to our investment strategies. This can be either because the data is of poor quality, has a very short history and very partial coverage, or because it has insufficiently been researched and tested.

Yet, there are also datasets around that satisfy our criteria. Recently, our efforts have centered around news flow data and company linkage data. For news flow, we are investigating whether our existing factor definitions can be further enhanced with news-related information. For example, we might be able to confirm positive momentum, based on positive sentiment detected in the news or overrule value, based on negative news sentiment. Similarly, for company linkages, we could extract sentiment on a business-to-business firm from its customers: if the customers are growing, this likely means more business for the supplier.

Moreover, some of the datasets we have looked at can be considered as “alternative data” in the sense that, although they may not be that ‘big’ because they still fit on the hard drive of a standard personal computer, they are clearly non-standard. Over the past few years, our research teams have investigated⁴ the use of various forms of such data, including option, short interest and detailed analysts data. Another example is the usage of corporate bond data in stock selection models. These alternative datasets have turned out to provide alpha while being less prone to overfitting.

³ See for example C. Harvey, ‘Presidential Address: The Scientific Outlook in Financial Economics’, 2017.

⁴ See for example: J. Blonk, B. van der Grient and W. de Groot, ‘Credit Momentum added to quant equity strategies’, June 2017.

Meanwhile, on the artificial intelligence front, we have experimented with advanced techniques, such as decision tree models, neural networks and genetic algorithms, for traditional quantitative selection data, in particular low-frequency stock exchange and financial statement data. However, our findings have been mostly disappointing. One major problem is that this data contains a huge amount of noise, which implies significant data mining risk.

As a result, we have focused our efforts on experimenting with techniques that improve upon the more conventional ones, that we have been using for many years. We consider the more advanced AI techniques particularly relevant as a tool for processing certain new kinds of datasets but not as something that is promising for every conceivable application.

Conclusion

Although big data and AI may look like promising developments for asset management, turning these techniques into efficient and reliable investment strategies remains a major challenge. For now, we see more potential in analyzing data signals available from credit, option and lending markets, using proven and transparent techniques, than in analyzing exotic big data variables with complex algorithms.

The reason is that these more traditional datasets show better quality, larger breadth and longer history. These datasets have also been thoroughly vetted, with ample empirical research published in the academic literature. In addition, these variables tend to be backed by a much more convincing economic rationale.

At the same time, we acknowledge the potential disruption this kind of innovation may lead to in the future. In the past, the datasets that are currently widely used by quantitative asset managers were subject to similar issues as big data today. Over the years, the quality, breadth and history of these datasets improved, and they became usable. Now, as time passes and more data becomes available,⁵ big data will probably also become increasingly usable.

Finally, the fundamental issue for investors may not necessarily be about choosing between one type of data or the other. There is a wide array of possibilities between sticking to traditional price and financial statement information, and solely relying on things like satellite imagery of parking slots. For example, big data and AI signals could be very useful to fundamental credit and equity analysts. This would feed through into our quantitative

'We see more potential in data available from credit, option and lending markets, using proven techniques'

⁵ In a 2017 white paper sponsored by Seagate, the International Data Corporation (IDC) forecasted that by 2025 the global datasphere will grow to 163 zettabytes (this is a trillion gigabytes). That's ten times the 16.1ZB of data generated in 2016. IDC also estimated that the amount of the global datasphere subject to data analysis will grow by a factor of 50 to 5.2ZB in 2025.

strategies, that take analyst revisions into account. In this case, we would be using big data and AI information in an indirect manner.

Important Information

Robeco Institutional Asset Management B.V. has a license as manager of Undertakings for Collective Investment in Transferable Securities (UCITS) and Alternative Investment Funds (AIFs) ("Fund(s)") from the Netherlands Authority for the Financial Markets in Amsterdam. This document is solely intended for professional investors, defined as investors which qualify as professional clients, have requested to be treated as professional clients or are authorized to receive such information under any applicable laws. Therefore, the information set forth herein is not addressed and must not be made available, in whole or in part, to other parties, such as retail clients. Robeco Institutional Asset Management B.V. and/or its related, affiliated and subsidiary companies, ("Robeco"), will not be liable for any direct, indirect, special, incidental or consequential damages arising out of the use of any opinion or information expressly or implicitly contained in this publication. The content of this document is based upon sources of information believed to be reliable. Without further explanation this document cannot be considered complete. It is intended to provide the professional investor with general information on Robeco's specific capabilities, but does not constitute a recommendation or an advice to buy or sell certain securities or investment products and/or to adopt any investment strategy. Nothing in this material constitutes investment, legal, accounting or tax advice, or a representation that any investment or strategy is suitable or appropriate to your individual circumstances, or otherwise constitutes a personal recommendation to you. Any opinions, estimates or forecasts may be changed at any time without prior warning. If in doubt, please seek independent advice. All rights relating to the information in this document are and will remain the property of Robeco. This material may not be copied or used with the public. No part of this document may be reproduced, saved in an automated data file or published in any form or by any means, either electronically, mechanically, by photocopy, recording or in any other way, without Robeco's prior written permission.

The material and information in this document are provided "as is" and without warranties of any kind, either expressed or implied. Robeco and its related, affiliated and subsidiary companies disclaim all warranties, expressed or implied, including, but not limited to, implied warranties of merchantability and fitness for a particular purpose.

Investment involves risks. Before investing, please note the initial capital is not guaranteed. The value of the investments may fluctuate. Past performance is no guarantee of future results. Investors should ensure that they fully understand the risk associated with the Fund. Investors should also consider their own investment objective and risk tolerance level. The information in this material may contain projections or other forward-looking statements regarding future events, targets, management discipline or other expectations which involve assumptions, risks, and uncertainties and is only as current as of the date indicated. Based on this, there is no assurance that such events will occur, and may be significantly different than that shown here, and Robeco cannot guarantee that these statistics and the assumptions derived from the statistics will reflect the market conditions that may be encountered or future performances. Historical returns are provided for illustrative purposes only. The price of units may go down as well as up and the past performance is not indicative of future performance. If the currency in which the past performance is displayed differs from the currency of the country in which you reside, then you should be aware that due to exchange rate fluctuations the performance shown may increase or decrease if converted into your local currency. The performance data do not take account of the commissions and costs incurred on the issue and redemption of units.

The prices used for the performance figures of the Luxembourg-based Funds are the end-of-month transaction prices net of fees up to 4 August 2010. From 4 August 2010, the transaction prices net of fees will be those of the first business day of the month. Return figures versus the benchmark show the investment management result before management and/or performance fees: the Fund returns are with dividends reinvested and based on net asset values with prices and exchange rates of the valuation moment of the benchmark. Please refer to the prospectus of the Funds for further details. Performance is quoted net of investment management fees. The ongoing charges mentioned in this document is the one stated in the Fund's latest annual report at closing date of the last calendar year.

This document is not directed to, or intended for distribution to or use by, any person or entity who is a citizen or resident of or located in any locality, state, country or other jurisdiction, such as US Persons, where such distribution, document, availability or use would be contrary to law or regulation or which would subject the Fund and its investment manager to any registration or licensing requirement within such jurisdiction.

Any decision to subscribe for interests in the Fund must be made solely on the basis of information contained in the prospectus which information may be different from the information contained in this document. The information contained in this document is qualified in its entirety by reference to the prospectus, and this document should, at all times, be read in conjunction with the prospectus. The prospectus and the Key Investor Information Document for the Robeco Funds can all be obtained free of charge at www.robeco.com.

Additional Information for US offshore investors

The Robeco Capital Growth Funds have not been registered under the United States Investment Company Act of 1940, as amended, nor the United States Securities Act of 1933, as amended. None of the shares may be offered or sold, directly or indirectly in the United States or to any US Person. A US Person is defined as (a) any individual who is a citizen or resident of the United States for federal income tax purposes; (b) a corporation, partnership or other entity created or organized under the laws of or existing in the United States; (c) an estate or trust the income of which is subject to United States federal income tax regardless of whether such income is effectively connected with a United States trade or business.

Additional Information for investors with residence or seat in Australia

This document is distributed in Australia by Robeco Hong Kong Limited (ARBN 156 512 659) ("Robeco") which is exempt from the requirement to hold an Australian financial services license under the Corporations Act 2001 (Cth) pursuant to ASIC Class Order 03/1103. Robeco is regulated by the Securities and Futures Commission under the laws of Hong Kong and those laws may differ from Australian laws. This document is distributed only to "wholesale clients" as that term is defined under the Corporations Act 2001 (Cth). This document is not for distribution or dissemination, directly or indirectly, to any other class of persons. In New Zealand, this document is only available to wholesale investors within the meaning of clause 3(2) of Schedule 1 of the Financial Markets Conduct Act 2013 ("FMCA"). This document is not for public distribution in Australia and New Zealand.

Additional Information for investors with residence or seat in Austria

This information is solely intended for professional investors or eligible counterparties in the meaning of the Austrian Securities Oversight Act.

Additional Information for investors with residence or seat in Brazil

The Fund may not be offered or sold to the public in Brazil. Accordingly, the Fund has not been nor will be registered with the Brazilian Securities Commission - CVM nor have they been submitted to the foregoing agency for approval. Documents relating to the Fund, as well as the information contained therein, may not be supplied to the public in Brazil, as the offering of the Fund is not a public offering of securities in Brazil, nor used in connection with any offer for subscription or sale of securities to the public in Brazil.

Additional Information for investors with residence or seat in Colombia

This document does not constitute a public offer in the Republic of Colombia. The offer of the Fund is addressed to less than one hundred specifically identified investors. The Fund may not be promoted or marketed in Colombia or to Colombian residents, unless such promotion and marketing is made in compliance with Decree 2555 of 2010 and other applicable rules and regulations related to the promotion of foreign Funds in Colombia. The distribution of this document and the offering of shares may be restricted in certain jurisdictions. Prospective applicants for the Fund should inform themselves of any applicable legal requirements, exchange control regulations and applicable taxes in the countries of their respective citizenship, residence or domicile.

Additional Information for investors with residence or seat in the Dubai International Financial Centre (DIFC), United Arab Emirates

This material is being distributed by Robeco Institutional Asset Management B.V. (Dubai Office) located at Office 209, Level 2, Gate Village Building 7, Dubai International Financial Centre, Dubai, PO Box 482060, UAE. Robeco Institutional Asset Management B.V. (Dubai office) is regulated by the Dubai Financial Services Authority ("DFSA") and only deals with Professional Clients or Market Counterparties and does not deal with Retail Clients as defined by the DFSA.

Additional Information for investors with residence or seat in France

Robeco is having the freedom to provide services in France. Robeco France (only authorized to offer investment advice service to professional investors) has been approved under registry number 10683 by the French prudential control and resolution authority (formerly ACP, now the ACPR) as an investment firm since 28 September 2012.

Additional Information for investors with residence or seat in Germany

This information is solely intended for professional investors or eligible counterparties in the meaning of the German Securities Trading Act.

Additional Information for investors with residence or seat in Hong Kong

The contents of this document have not been reviewed by the Securities and Futures Commission ("SFC") in Hong Kong. If you are in any doubt about any of the contents of this document, you should obtain independent professional advice. This document has been distributed by Robeco Hong Kong Limited ("Robeco"). Robeco is regulated by the SFC in Hong Kong.

Additional Information for investors with residence or seat in Italy

This document is considered for use solely by qualified investors and private professional clients (as defined in Article 26 (1) (b) and (d) of Consob Regulation No. 16190 dated 29 October 2007). If made available to Distributors and individuals authorized by Distributors to conduct promotion and marketing activity, it may only be used for the purpose for which it was conceived.

The data and information contained in this document may not be used for communications with Supervisory Authorities. This document does not include any information to determine, in concrete terms, the investment inclination and, therefore, this document can not and should not be the basis for making any investment decisions.

Additional Information for investors with residence or seat in Panama

The distribution of this Fund and the offering of shares may be restricted in certain jurisdictions. The above information is for general guidance only, and it is the responsibility of any person or persons in possession of the prospectus of the Fund and wishing to make application for shares to inform themselves of, and to observe, all applicable laws and regulations of any relevant jurisdiction. Prospective applicants for shares should inform themselves as to legal requirements also applying and any applicable exchange control regulations and applicable taxes in the countries of their respective citizenship, residence or domicile.

Additional Information for investors with residence or seat in Peru

The Fund has not been registered before the Superintendencia del Mercado de Valores (SMV) and are being placed by means of a private offer. SMV has not reviewed the information provided to the investor. This document is only for the exclusive use of institutional investors in Peru and is not for public distribution.

Additional Information for investors with residence or seat in Shanghai

This material is prepared by Robeco Investment Management Advisory (Shanghai) Limited Company ("Robeco Shanghai") and is only provided to the specific objects under the premise of confidentiality. This material must not be wholly or partially reproduced, distributed, circulated, disseminated, published or disclosed, in any form and for any purpose, to any third party without prior approval from Robeco Shanghai. The information contained herein may not reflect the latest information on account of the changes and Robeco Shanghai is not responsible for the updating of the material or the correction of inaccurate or missing information contained in the material. Robeco Shanghai has not yet been registered as the private fund manager with the Asset Management Association of China. Robeco Shanghai is a wholly foreign-owned enterprise established in accordance with the PRC laws, which enjoys independent civil rights and civil obligations. The statements of the shareholders or affiliates in the material shall not be deemed to a promise or guarantee of the shareholders or affiliates of Robeco Shanghai, or be deemed to any obligations or liabilities imposed to the shareholders or affiliates of Robeco Shanghai.

Additional Information for investors with residence or seat in Singapore

This document has not been registered with the Monetary Authority of Singapore ("MAS"). Accordingly, this document may not be circulated or distributed directly or indirectly to persons in Singapore other than (i) to an institutional investor under Section 304 of the SFA, (ii) to a relevant person pursuant to Section 305(1), or any person pursuant to Section 305(2), and in accordance with the conditions specified in Section 305, of the SFA, or (iii) otherwise pursuant to, and in accordance with the conditions of, any other applicable provision of the SFA. The contents of this document have not been reviewed by the MAS. This document is not intended as a recommendation or for the purpose of soliciting any action in relation to Robeco Capital Growth Funds or other Robeco Funds and should not be construed as an offer to sell shares of the Fund ("Shares") or solicitation by anyone in any jurisdiction in which such an offer or solicitation is not authorized or to any person to whom it is unlawful to make such an offer and solicitation. Any decision to subscribe for interests in the Fund must be made solely on the basis of information contained in the prospectus ("Prospectus"), which information may be different from the information contained in this document, and with independent analyses of your investment and financial situation and objectives. The information contained in this document is qualified in its entirety by reference to the Prospectus, and this document should, at all times, be read in conjunction with the Prospectus. Detailed information on the Fund and associated risks is contained in the Prospectus. Any decision to participate in the Fund should be made only after reviewing the sections regarding investment considerations, conflicts of interest, risk factors and the relevant Singapore selling restrictions (as described in the section entitled "Important Information for Singapore Investors") contained in the Prospectus. You should consult your professional adviser if you are in doubt about the stringent restrictions applicable to the use of this document, regulatory status of the Fund, applicable regulatory protection, associated risks and suitability of the Fund to your objectives. This document is not directed to, or intended for distribution to or use by, any person or entity who is a citizen or resident of or located in any locality, state, country or other jurisdiction where such distribution, publication, availability or use would be contrary to law or regulation or which would subject the Fund and its investment manager to any registration or licensing requirement within such jurisdiction. Investors should note that only the sub-funds listed in the appendix to the section entitled "Important Information for Singapore Investors" of the Prospectus ("Sub-Funds") are available to Singapore investors. The Sub-Funds are notified as restricted foreign schemes under the Securities and Futures Act, Chapter 289 of Singapore ("SFA") and are invoking the exemptions from compliance with prospectus registration requirements pursuant to the exemptions under Section 304 and Section 305 of the SFA. The Sub-Funds are not authorized or recognized by the Monetary Authority of Singapore and Shares in the Sub-Funds are not allowed to be offered to the retail public in Singapore. The Prospectus of the Fund is not a prospectus as defined in the SFA. Accordingly, statutory liability under the SFA in relation to the content of prospectuses would not apply. The Sub-Funds may only be promoted exclusively to persons who are sufficiently experienced and sophisticated to understand the risks involved in investing in such schemes, and who satisfy certain other criteria provided under Section 304, Section 305 or any other applicable provision of the SFA and the subsidiary legislation enacted thereunder. You should consider carefully whether the investment is suitable for you. Singapore Private Limited holds a capital markets services licence for fund management issued by the MAS and is subject to certain clientele restrictions under such licence. An investment will involve a high degree of risk, and you should consider carefully whether an investment is suitable for you. Additional Information for investors with residence or seat in Spain

The Spanish branch Robeco Institutional Asset Management BV, Sucursal en España, having its registered office at Paseo de la Castellana 42, 28046 Madrid, is registered with the Spanish Authority for the Financial Markets (CNMV) in Spain under registry number 24.

Additional Information for investors with residence or seat in Switzerland

This document is exclusively distributed in Switzerland to qualified investors as defined in the Swiss Collective Investment Schemes Act (CISA) by Robeco Switzerland AG which is authorized by the Swiss Financial Market Supervisory Authority FINMA as Swiss representative of foreign collective investment schemes, and UBS Switzerland AG, Bahnhofstrasse 45, 8001 Zurich, postal address: Europastrasse 2, P.O. Box, CH-8152 Opfikon, as Swiss paying agent. The prospectus, the key investor information documents (KIIDs), the articles of association, the annual and semi-annual reports of the Fund(s), as well as the list of the purchases and sales which the Fund(s) has undertaken during the financial year, may be obtained, on simple request and free of charge, at the office of the Swiss representative Robeco Switzerland AG, Josefstrasse 218, CH-8005 Zurich. The prospectuses are also available via the website www.robeco.ch.

Additional Information for investors with residence or seat in United Arab Emirates

Some Funds referred to in this marketing material have been registered with the UAE Securities and Commodities Authority (the Authority). Details of all Registered Funds can be found on the Authority's website. The Authority assumes no liability for the accuracy of the information set out in this material/document, nor for the failure of any persons engaged in the investment Fund in performing their duties and responsibilities.

Additional Information for investors with residence or seat in the United Kingdom

Robeco is subject to limited regulation in the UK by the Financial Conduct Authority. Details about the extent of our regulation by the Financial Conduct Authority are available from us on request.

Additional Information for investors with residence or seat in Uruguay

The sale of the Fund qualifies as a private placement pursuant to section 2 of Uruguayan law 18.627. The Fund must not be offered or sold to the public in Uruguay, except in circumstances which do not constitute a public offering or distribution under Uruguayan laws and regulations. The Fund is not and will not be registered with the Financial Services Superintendency of the Central Bank of Uruguay. The Fund corresponds to investment Funds that are not investment Funds regulated by Uruguayan law 16.774 dated September 27, 1996, as amended.