Financial data science for responsible investors

By: Andreas Hoepner, Associate Professor of Finance,

ICMA Centre, Henley Business School

Introduction by Faith Ward

Chief Responsible Investment and Risk Officer, Environment Agency Pension Fund

A core component of our responsible investment policy is to "apply evidenced based decision making in the implementation of responsible investment." In doing that we draw upon and contribute to academic research to inform our decision making. We have been working with the ICMA Henley Business School on a long term project to analyse the outcomes of our decision making and the integration of the consideration of environmental, social and governance risks from over a decade's worth of data. We will share the findings of this research in due course. Before that we wanted to share how financial data science (FDS) can help investors consider the perennial and thorny question as to whether ESG integration in investment processes can be beneficial to returns and risk of their portfolio. In summary Andreas makes the following observations;

- Financial data science uses real world data and focuses on the predictability of outcomes, allowing practical solutions to be developed more quickly than work based on economic theory. He parallels the work to medical science where you would start treating the symptoms of disease long before you know the exact cause.
- Financial data science also bases its analysis on what actually happens, including the impacts of human behaviour, more rationally acknowledging that the world rarely black and white.
- Apply FDS techniques to the analysis of so called' sin stocks' e.g. tobacco, demonstrates they do not outperform, contrary to other studies. He argues that the quality of the data model using FDS is greater as it can explain more data variables and dig deeper in to what is actually happening.

Responsible investment and paradigms of financial research

Probably the first academic question in responsible investment has been, does integrating environmental, social or governance (ESG) criteria into the investment process have a significant impact on return or risk?

Financial economists such as Andrew Rudd quickly concluded as early as 1981 that theoretically, in an efficient market, any deviation from a shareholder value creation objective would be seen as detrimental to portfolio diversification and hence negative for return/risk ratios. This conclusion poses four questions to the critical mind though: first, can complex financial markets ever be fully explained from a theoretical perspective? Second, are financial markets really fully efficient? Third, does the integration of ESG criteria really represent a deviation from a long term shareholder value creation perspective? Fourth, would investments in firms underappreciated by responsible investors perform any better than those investments integrating ESG criteria?

As a financial data scientist, I aim to minimize my use of assumptions. Whenever I am inevitably exposed to assumptions, I make every effort to empirically test these. In other words, while financial economists tend to look at the world from their theoretical angle, financial data scientists like me undertake a deep investigation of all available data to then arrive at conceptual explanations of what happens in the real world.

Financial data science also does not believe in a general equilibrium as economics but understands the social world as time varying distributions of human behaviour.

More visually explained, while financial economists tend to see the world as black or white and search to reject their theories or identify evidence consistent with them, financial data scientists understand the world as shades of grey, fifty or more. In this sense, financial data science aims to arrive at the deepest possible conceptual understanding of human behaviour within the financial markets of a certain period of time, it does not aim to arrive at general theories of human behaviour. Such general theories do not appear to be a realistic aim to strive for in the analysis of human behaviour, as they are nearly always over-simplistic and unlikely to be sufficiently accurate for too long in the 21st century with its individualized human beings and rapid technological development.

Another crucial distinctions between financial economics and financial data science rests in the distinction between causality and predictability. Whereas financial economics is very concerned with a theoretical notion of direct causality, financial data science takes a more actionable perspective than causality by focusing on predictability. Research projects in financial economics

are often rejected if there is no complete assurance that a key explanatory variable (e.g. ESG ratings) has a direct causal impact on a key outcome variable (e.g. stock returns).

A direct causal impact means in this context that

(i) no third variable (e.g. quality of management) affects both variables,
(ii) no unobserved third variable (e.g. reputation) represents a crucial step between key explanatory variable and key outcome variable, and
(iii) reverse causality (e.g. stock returns affecting ESG ratings) can also be ruled out.

While such a level of detail in conceptual understanding is, of course, laudable and crucial for those aiming to develop general theories of human behaviour, it significantly delays the publication process due to relevant variables often simply not being observable. In this context, it is unsurprising that there has been a "slowdown of the economics publishing process" (Ellison, 2000) which was more pronounced in financial economics than, for instance, in environmental economics.

Financial data science, developing at the intersection between finance and computer science, takes a more actionable view than causality and focuses on predictability inspired by the medical sciences. While the medical sciences also strive to understand as much theoretical causality as identifiable at reasonable resource requirements, they only require risk markers (i.e. probable associations) instead of directly causal drivers of risk for therapeutic decision making.

In other words, when faced with serious diseases such as cancer or an epidemic, it is clearly advisable to start treatment in parallel to continued research once a risk marker indicates the presence of the disease instead of waiting, with the treatment, until the risk marker or another variable is clearly identified as the direct cause of the disease. While the performance of pension fund investment strategies does not constitute a matter of life or death as many research questions in the medical sciences, it nevertheless is a crucial for the quality of the livelihoods of millions of current and future retirees. In this sense, financial data scientists define predictability as the practical ability to integrate a marker with statistically significant predictive value for future return/risk performance into real world investment decision making.[1]

[1] The question, if this marker is also the (single) causal driver of the outcome variable or related in an alternative way, is of substantial educational value for the financial data scientist but no requirement for the utilization and/or publication of the predictive relationship.

Are financial markets really fully efficient?

As financial data science does not aim to arrive at general theories of human behaviour and favours predictability over direct causality to draw conclusions, it naturally has a more data driven view on market efficiency than financial economics. Whereas financial economics aims to classify a market in one of three categories as strong form, semi-strong form or weak form efficient, financial data science concentrates on individual market segment and aims to measure the degree of efficiency for these.

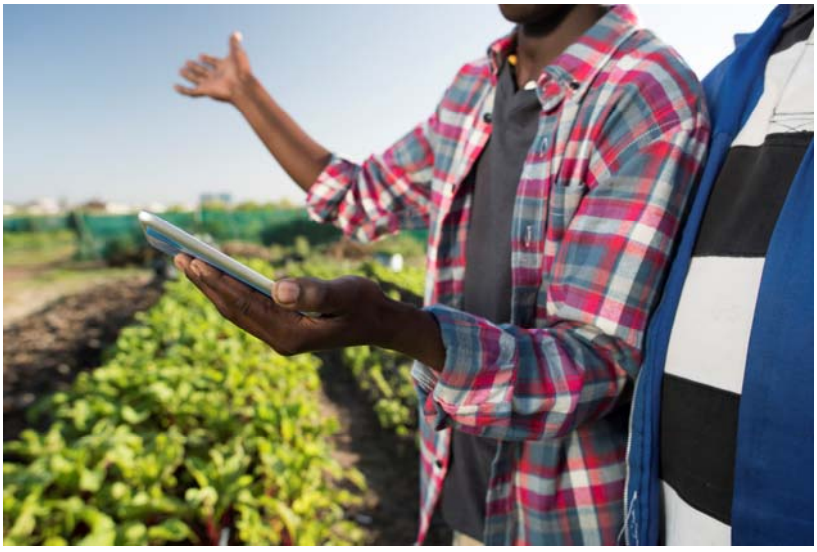
A market segment is hereby defined as the intersection between

(i) the asset class (e.g. US large cap equities),
(ii) the information set used to search for investment signals (e.g. financial accounts),
(iii) the search algorithm employed in identifying and validating investment signals (e.g. Discounted Cash Flow models) and
(iv) the investment time horizon (e.g. a five year horizon with quarterly observation intervals).

In the example market segment of searching for inefficiencies in the financial accounts of US large cap equities with Discounted Cash Flow models on a five year horizon with quarterly observation intervals, the number of competitors an individual investor would have lies in the hundreds of thousands judging by the number of Bloomberg, Eikon and Factset terminals sold. Consequently, the degree of efficiency of this market segment is likely to be close to hundred percent at any point in time with any temporary inefficient likely to be arbitraged away in little time.

However, financial data science has shown that even when just replacing the financial accounts with simplistic ESG data and using basic Fama-French attribution models instead of Discounted Cash Flows, significant alphas can be achieved as long as competition stays small.

Can responsible investment datasets signal higher return or lower risk?



As I summarise in my most recent position paper, responsible investment research in general has shown that ESG datasets can be utilised to identify signals for higher return or reduced downside risk, if the particular dataset is not too widely used or the financial data science search algorithm is sufficiently advanced.

On the upside, we studied for instance the firms announced as global 100 most sustainable by the World Economic Forum

when the underlying data was produced by Innovest with industry specific sustainability criteria in the respective year after the announcement when the award was public knowledge and in the year before the announcement when only Innovest clients were informed. We find the most sustainable consumer discretionary companies to significantly outperform when competition was low as the data was private but perform only benchmark once the information became public. In the health care sector, financial analysts ignore corporate performance on sustainability.

In the health care sector, financial analysts tend to ignore corporate performance on sustainability. However, we find that the most sustainable companies in this sector outperform the industry benchmark in the year before the announcement and even stronger in the year after the announcement. This signals a substantial degree of market segment inefficiency but can be explained by health care being a 'credence service' which consumers cannot assess during consumption and often not even for some time afterwards. For credence services, trust is absolutely essential and any trust building mechanism such as winning a sustainability award of the World Economic Forum might enhance performance.

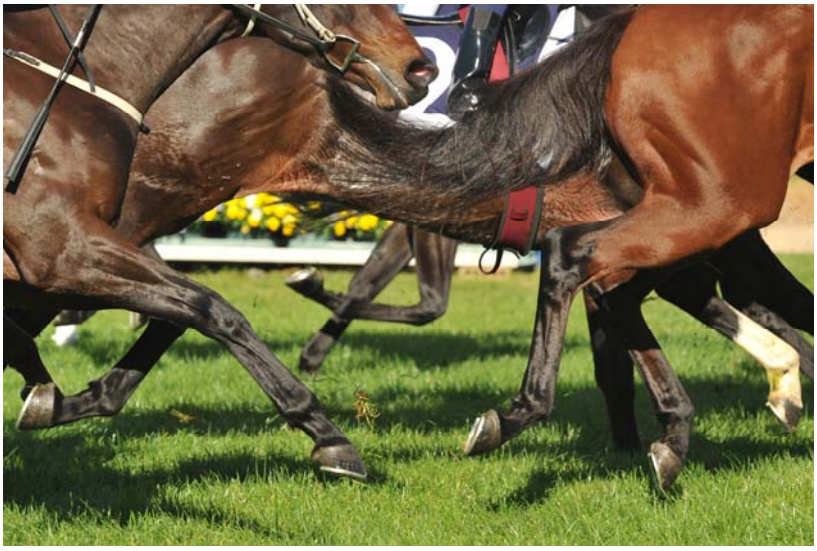
On the downside, we find ESG integration to be very suitable for reduction of actual risk. By actual risk, I mean any measure concentrating on the probability of investment performance below expectations without diluting itself by including upside deviations in its risk measurement. In other words, while standard deviation or tracking error are very popular measures of risk in financial markets, they are actually measures of uncertainty, as they include as much upside as downside deviations. The actual risk measures are semi-standard deviation and trailing error, respectively. When using these actual measure of risk to investors, we find environmental and social ratings from EIRIS and Governance ratings of GMI (now MSCI) to be significant in reducing risk not only at the firm but also at the portfolio level.

In terms of fixed income research, we recently studied all existing ESG integrating strategies worldwide with over three years of return history to understand which kind of ESG expertise would distinguish between the financially successful and unsuccessful funds. We test various assets under management related criteria as well as screening intensity without any success. However, when we test engagement which large, sophisticated bond investors such as PIMCO engage in frequently, we find significant performance differences. These are particularly amplified for bond investors willing to take litigation action if needed.

In conclusion, ESG integration in investment processes can be beneficial to returns and risk reducing if investors possess the necessary sophistication. One challenge remains, however, which is that previous research has repeatedly found companies in the alcohol, tobacco and gambling sector, so called 'sin stocks', to outperform benchmarks. If these stocks underappreciated by responsible investors would really do well, some investors might feel obliged through legal requirements to invest over-proportionally in these sectors.

But what about those sin stocks?

In a new study, that was awarded the Sycomore-Award for Best Quantitative Paper at the PRI in Person conference earlier this month, my postdoc Hampus Adamsson and I reinvestigate this common belief that sin stocks outperform equity markets. We focus our analysis on actually investable stocks instead of analysing all securities listed on a specific academic database. In contrast, financial economists previously studying sin stocks mostly used an academic dataset provided by the rather influential University of Chicago which includes thousands of stocks that are too small to be listed in the major equity indices (e.g. FTSE, MSCI) which most pension funds follow nowadays. While we also use this dataset in other work, we aim to understand in this paper if the so called sin stock premium is actually investable or just an ivory tower artefact caused by impractical academic research methods.



As financial data scientists to whom large scale, deep data and advanced statistical analysis are the essence of scientific enquiry, we also felt obliged to undertake a deeper dive to identify the characteristics of sin stocks more precisely than previous economic analyses. The essence of our deeper statistical work is to describe return characteristics of firms not only in general economic terms but also by their industry specific characteristics. While one might, simplistically, consider tobacco firms as 'safe' value stocks, deeper analysis shows that they actually exhibit riskier growth stock characteristics within their industry, which itself (i.e. consumer goods) is closely associated with value investing. In other words, while consumer goods are likely a safer investment during economic downturns than the average industries, tobacco stocks are actually less likely safe than other consumer goods.

This level or statistical depth allows us to control for firm size effects both at the economy and at the industry specific level. These size effect controls are crucial as the original study by Hong and Kacperczyk finding sin stocks to outperform used purely equal weighted portfolios of sin stocks which put much more emphasis on the performance of smaller stocks than actual pension funds would. With smaller stocks having the tendency to outperform larger stocks over long horizons, equal weighting might put the sin stock portfolios in previous studies at an undue advantage against value

weighted benchmarks - especially when one considers that these studies predominantly use the Chicago dataset with its thousands of additional small firms that are usually irrelevant for institutional investors.

When analysing realistic, value weighted portfolios of investable sin stocks using our advanced statistical control for firm characteristics at the economy and industry level, we find no trace of outperformance of sin stocks, neither globally nor when we focus on the US as many previous authors did. This is consistent with our previous work, when Stefan Zeume (University of Michigan) and I found that the only existing mutual fund whose investment strategy focuses on sin stocks (a fund called VICEX until recently), does not outperform when analysed with advanced statistical controls. Hence, divestment from sin stocks is unlikely harmful if the divesting institution employs a bit of financial data science to ensure that the divestment does not result in unintended portfolio exposures compared with benchmarks.

But how do these findings fit with the debate between divestment campaigners and financial economists? Divestment campaigners usually argue that investing in sin stocks and thereby financially supporting sinful behaviour is morally wrong and should be avoided. Since proponents of this perspective tend to have a high opinion of the relevance of moral considerations for average human behaviour, they expect divestment to become commonplace and result in a share price collapse of the divested firms. In contrast, the headline argument of financial economists is that, according to Merton's theory, the expected return of the sin stocks is abnormally high because their price is too low given the expected future earnings. As reason for the low price authors, such as Hong and Kacperczyk, normally reference the fact that a noticeable degree of institutional investors shun sin stocks.

Despite these perspectives appearing opposed at face value, there are two intriguing similarities between them. First, they are both theoretical in nature. Proponents of each perspective make subjective claims on what they personally expect to happen instead of studying what is actually happening. While the moral perspective is very transparent about its underlying assumptions and resulting subjectivity, the theoretical financial economics perspective is more subtle in its communication but no less subjective. Economists made a personal, subjective decision when they decided to define their main actors, the homo economicus, as non-emotional. There is little empirical evidence to support this decision.

Second, both arguments are ignorant of each other. While the moral perspective ignores the fact that a significant proportion of financial market participants assesses stocks predominantly based on expected future cash flows, the financial economics perspective ignores the fact that another

significant proportion of financial market participants assess stocks on its expected momentum in terms of investor preference. The irony of the theoretical financial economics perspective is that it argues sin stock prices to be, historically, too low due to investor distastes but, looking forward, ignores such imbalances in investor preferences and instead only argues based on expected future earnings.

In summary, while both divestment campaigners and financial economics tend to argue on a normative ('should happen') level, modern technologies such as algorithmic learning and big data analytics allow financial data scientists to address crucial societal questions from a more descriptive ('is happening') level. This is not only crucial for potential divestment of sin stocks, but also for investing purely in environmentally responsible firms. While financial economists would expect that this should increase portfolio risk, it is actually significantly reducing the worst case risk of investable pension fund portfolios.

By: Andreas Hoepner, Associate Professor of Finance, ICMA Centre, Henley Business School

Dr. Andreas G. F. Hoepner is an Associate Professor of Finance at the ICMA Centre of Henley Business School. He is currently also heading the Practical Tools research stream of the MISTRA sustainable financial systems (MFS) research consortium (total funding: 58 million SEK). Dr. Hoepner serves on independent assessment committees for the Investment & Pensions Europe (IPE) Awards (Categories: Climate Change Risk, ESG, Smart Beta), the Investment Innovation Benchmark (IIB), and the RI Awards; and sits on advisory boards for Bank J. Safra Sarasin, the Carbon Disclosure Project (CDP) and French Social Investment Forum (FIR). He received his PhD from St Andrews in June 2010, where he was on faculty from February 2009 to September 2013 and built up the Centre for Responsible Banking and Finance as its Deputy Director since November 2011. He is also founding co-director of a social enterprise (Sociovestix Labs) and founding president of a charitable organisation (ReFine Research) that gives social reporting awards to pension funds. Prior to taking up his MISTRA role in March 2016, Andreas served over six years as lead academic advisor United Nations supported Principles for Responsible Investment and also consulted for organisations including the International Finance Corporation (IFC), the University of Cambridge's Institute for Sustainability Leadership and the University of Oxford's Moller Centre.

For more about Dr Hoepner's work please go here.