



OSSIAM RESEARCH TEAM

ESG controversies and their impact on performance

October 5th 2018

ESG controversies and their impact on performance

October 5th 2018

Carmine De Franco, PhD
Head of Fundamental Research
carmine.de-franco@ossiam.com

Abstract

In this paper we introduce an aggregated controversy metric, derived from ESG data, which targets specific issues companies face in the environmental, social or governance fields. By building portfolios based on this controversy measure, we show that in Europe and the US stocks that undergo severe controversies significantly underperform their benchmarks and portfolios consisting of stocks with low or no controversy at all. The main reason for this is that, in both cases, markets tend to react strongly to changes in controversy levels and penalize stocks that experience ESG downgrades. Our results make a clear case for the potential benefits of excluding stocks with high controversy levels from investment universes. These findings were not confirmed for the Asia-Pacific region, where the portfolio consisting of highly controversial stocks outperformed its benchmark, although their number was lower compared to Europe and the US.

Key words: ESG, Controversial stocks.

1 Executive Summary

In recent years, the financial services industry and investors have started to integrate ESG¹ considerations in their business and investment decisions, and we expect this to become mainstream thinking in the coming years. Particularly for investors the integration of environmental, social and governance aspects will become standard practice and for many of them this is already the case. ESG analytics provide useful and valuable information to investors on how to choose opportunities in their investment universes and to better understand associated risks, especially non-market risks related to climate change, new regulation and governance. One of the most interesting areas in ESG relates to the sometimes ambiguous notion of *controversy*. For some investors, the notion of controversy is linked to the business in which companies operate. Therefore, companies in tobacco, brewing and distilling, gambling, defense or certain biotechnology industries qualify as controversial, and as such are excluded from the investment universe. Nevertheless, some studies have pointed out how these stocks, in Europe and the US, have delivered strong performances over time (see e.g. Fabozzi et al. (2008); Hong and Kacperczyk (2009); Blitz and Fabozzi (2017)).

In this paper we take a different approach and clearly distinguish controversy from business sector and activity. More precisely, we define controversy as the measure of specific, well-identified weaknesses of a company that can potentially have a significant impact on its business and, by transmission, on its financial performance. In many instances controversies are so severe that they make the news. In our framework, the controversy indicator is based on ten specific incident-related indicators for environmental, social and governance issues. Incident-related issues usually denote failures of a company to address and mitigate associated risks. Furthermore, they usually precede strong reactions in the market and can also expose the company to severe short- and long-term consequences. For example, reputational and legal risks, or loss of business opportunities.

Our paper shows that for US and European stocks² high levels of controversy are significantly detrimental to the performance of a diversified portfolio and come with higher risk, mainly relating to drawdowns. Although using different mechanisms, US and European stocks typically experience negative excess returns over their benchmarks when their controversy ratings are downgraded and recover very slowly from these losses.

Quite different is the case for Asia-Pacific stocks. Not only do controversial stocks not seem to have negative effects on the performance of their respective portfolios, but also the market does not react and penalize them as in the US and Europe. These results must though be caveated by the fact that our data sample did not have a sufficiently high number of controversial stocks, which makes our estimates noisy and less conclusive.

Putting these findings in perspective, it seems that our systematic controversy indicator could help investors to avoid companies that display alarming weaknesses and could potentially have a negative impact on their overall portfolio's performance and risk.

2 Data

We consider the large and mid-capitalization segment of publicly listed equities in developed countries since January 2010: the United States; Western Europe, including Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland and the United Kingdom; and Asia-Pacific, including Japan, Singapore, Hong Kong, Australia and New Zealand. For each company in these regions, we collected ESG data from Sustainalytics and built a controversy indicator. The indicator is an aggregated measure related to the severity of a stock's controversy. For this, we looked at a specific subset of ESG indicators that are usually linked with controversies, as investors and the public opinion perceive them. Table 1 contains this subset.

C_1 : Operations Incidents	C_6 : Public Policy Incidents
C_2 : Environmental Supply Chain Incidents	C_7 : Employee Incidents
C_3 : Product & Service Incidents	C_8 : Social Supply Chain Incidents
C_4 : Business Ethics Incidents	C_9 : Customer Incidents
C_5 : Governance Incidents	C_{10} : Society & Community Incidents

Table 1: ESG indicators related to controversy levels. Source: Sustainalytics.

Companies are rated from 0 (worst) to 100 (best) for each indicator and we defined the controversy value CV as

$$CV = \min_{i=1, \dots, 10} C_i$$

Finally, we transformed the CV indicator to obtain four levels of controversy as follows:

Controversy Level	Controversy Value
No	$CV = 100$
Low	$51 \leq CV \leq 99$
Moderate	$21 \leq CV \leq 50$
High	$CV \leq 20$

In practice, a high level of controversy is due to one or more of the ESG indicators C_1, \dots, C_{10} being lower than 20 or, equally, that the company has experienced serious incidents related to one or more of the selected ten fields. Controversy level reflects a company's level of involvement in ESG issues and how the company manages them. It focuses on the severity of the impacts and the risks the company is facing. Following the definition by Sustainalytics, the use of the indicators in Table 1 is based upon international norms and principles, as it is a key element to determine the impact (see

for example the Ten Principles of the UN Global Compact). For each geographic zone (US, Western Europe and Asia-Pacific) we built four equally-weighted portfolios made up of stocks with the same level of controversy, and we tracked their performance and risk over time. Portfolios were calculated in USD, except in the Western European case where they were in EUR. Net dividends are reinvested in the portfolios. Stock prices and dividends were sourced from Thomson Reuters/Datastream. Portfolios are regularly rebalanced on a quarterly basis, but changes in stocks controversy levels were immediately implemented. For example, a stock that saw its level of controversy increase from Low to Moderate was immediately transferred from the Low Controversy portfolio to the Moderate Controversy one.

3 Portfolios based on controversy levels

In the following section we examine the impact of controversy on aggregated portfolios since January 2010.

Europe

Figure 1 shows the historical evolution of portfolios at different levels of controversy in Western Europe.

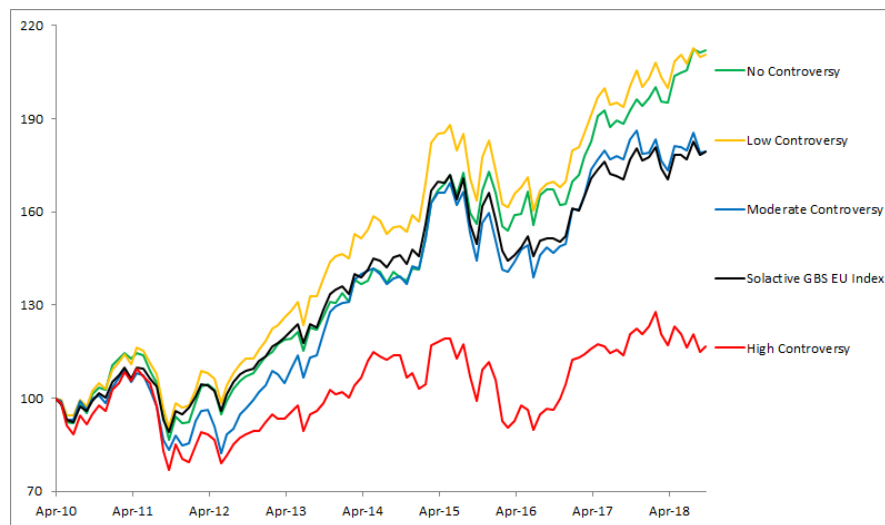


Figure 1: The benchmark (Solactive GBS EU Index) and four portfolios based on controversy levels. Data from January 2010 to September 2018 in EUR. Source Datastream, Sustainalytics, Solactive.

The graph illustrates how the No Controversy, Low Controversy and, to a lesser extent, Moderate Controversy portfolios outperformed the benchmark over the period. In contrast, the High Controversy portfolio severely underperformed the benchmark. Details are provided in Table 2. It was striking how annualized performance increased from

1.85% for the High Controversy to 9.28% for the No Controversy portfolio. The High Controversy portfolio displayed roughly 3.5% more volatility than the No Controversy portfolio, although the former was significantly less diversified than the latter (on average, 26 stocks were classified as highly controversial versus 104 with no controversy and 172 with low levels of controversy). Furthermore, the maximum drawdown of the High Controversy portfolio (-36.16%) was significantly higher than that of the benchmark and of the other portfolios. Altogether, Table 2 provides clear evidence that the controversy level is a very powerful tool to highlight troubled companies, both from risk and performance perspectives.

	Controversy Portfolios				
	Benchmark	High	Moderate	Low	No
Ann. Performance	7.14%	1.85%	7.15%	9.19%	9.28%
Ann. Volatility	16.19%	19.85%	18.6%	16.68%	16.36%
Max. Drawdown	-25.93%	-36.16%	-31.72%	-26.15%	-29.18%
Sharpe Ratio	0.44	0.09	0.38	0.55	0.57
CAPM Alpha	-	-6.44%	-0.8%	1.97%	2.38%
Information Ratio	-	-0.77	0	0.65	0.44
Avg. Number of Stocks	376	26	74	172	104

Table 2: Performance and risk indicators for the benchmark (Solactive GBS EU Index) and the four portfolios based on controversy levels. Data from January 2010 to September 2018 in EUR. Source Datastream, Sustainalytics, Solactive.

US

For US companies we found very similar results, as shown in Figure 2. The High Controversy portfolio clearly underperformed the benchmark and other portfolios. The most visible difference with the European case was that the Moderate Controversy portfolio also underperformed the benchmark, even if the magnitude of underperformance was relatively small.

Table 3 shows key performance indicators for the four portfolios and benchmark. Once again, the High Controversy portfolio displays lower performance than the benchmark. Even if annualized performance is significantly higher than the European High Controversy portfolio, in relative terms, the gap to its respective benchmark is 5.31% annualized, similar to 5.29% in the European case. This underperformance comes with relatively higher volatility and significantly higher drawdown (-27.75% versus -18.82% for the benchmark).

Similar to the European case, we also found that the Low and No Controversy portfolios outperformed the benchmark, although they both showed slightly higher volatility and

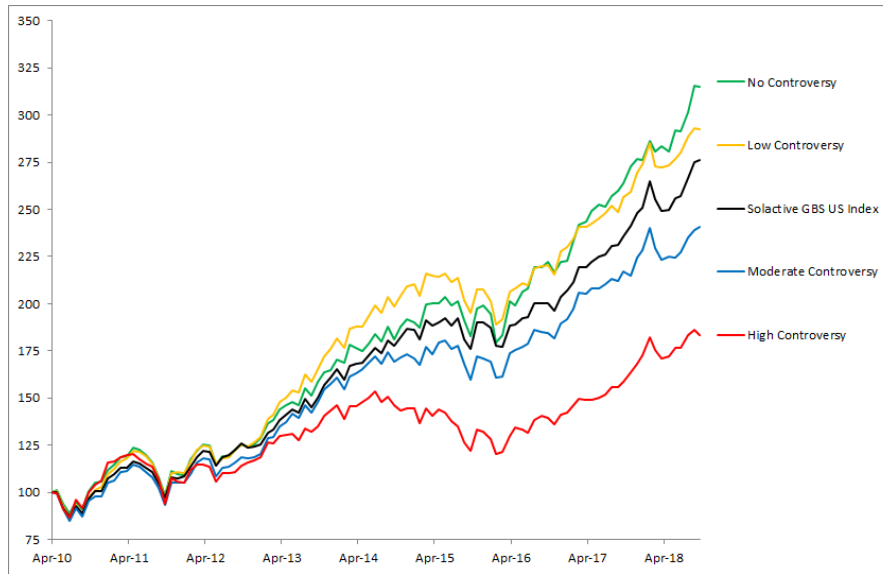


Figure 2: The benchmark (Solactive GBS US Index) and four portfolios based on controversy levels. Data from January 2010 to September 2018 in USD. Source Datastream, Sustainalytics, Solactive.

drawdowns. Both issues were related to the lower diversification of the portfolios and the equal-weighting of stocks, which automatically increases the weight of smaller capitalization stocks, and in turn, adds more risk to a portfolio. Overall, we found the same pattern in compounded performance from the High to the No Controversy portfolios: in the US too the controversy indicator clearly identified companies that carry specific ESG risks which, in turn, were detrimental to the performance and risk profiles.

	Benchmark	Controversy Portfolios			
		High	Moderate	Low	No
Ann. Performance	12.73%	7.42%	10.93%	13.49%	14.5%
Ann. Volatility	14.28%	16.07%	14.82%	15.54%	15.7%
Max. Drawdown	-18.82%	-27.75%	-20.82%	-23.37%	-23.58%
Sharpe Ratio	0.89	0.46	0.74	0.87	0.92
CAPM Alpha	-	-6.08%	-2.01%	-0.08%	0.98%
Information Ratio	-	-0.98	-0.6	0.24	0.43
Avg. Number of Stocks	572	28	98	294	152

Table 3: Performance and risk indicators for the benchmark (Solactive GBS US Index) and the four portfolios based on controversy levels. Data from January 2010 to September 2018 in USD. Source Datastream, Sustainalytics, Solactive.

Asia-Pacific

For the Asia-Pacific region our results were somewhat different to the European and the US cases, as Figure 3 and Table 4 show. We noticed how the Low and No Controversy portfolios showed higher annualized performance than the benchmark, especially the No Controversy one, which also had significantly lower levels of drawdown.

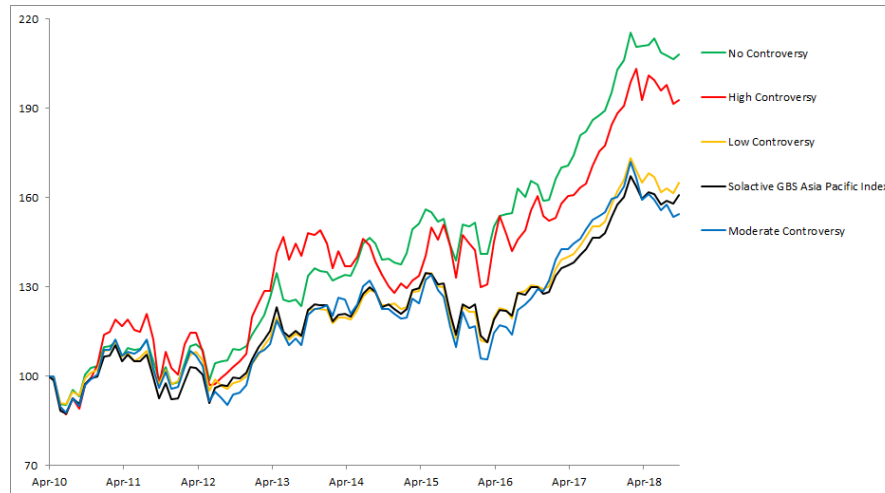


Figure 3: The benchmark (Solactive GBS Asia-Pacific Index) and four portfolios based on controversy levels. Data from January 2010 to September 2018 in USD. Source Datastream, Sustainalytics, Solactive.

From this perspective, we saw that in the Asia-Pacific region the controversy indicator proved to be a *positive discriminant*. But when we looked at the High Controversy portfolio, we did not notice any significant loss of performance.

	Controversy Portfolios				
	Benchmark	High	Moderate	Low	No
Ann. Performance	5.77%	8.05%	5.26%	6.09%	9.02%
Ann. Volatility	15.21%	19.08%	16.43%	15.13%	14.37%
Max. Drawdown	-24.67%	-27.13%	-28.26%	-24.15%	-17.9%
Sharpe Ratio	0.38	0.42	0.32	0.4	0.63
CAPM Alpha	-	1.86%	-0.76%	0.45%	3.74%
Information Ratio	-	0.23	-0.12	0.11	0.87
Avg. Number of Stocks	463	10	46	170	237

Table 4: Performance and risk indicators for the benchmark (Solactive GBS Asia-Pacific Index) and four portfolios based on controversy levels. Data from January 2010 to September 2018 in USD. Source Datastream, Sustainalytics, Solactive.

On the contrary, the annualized return of this portfolio was 8.05%, higher than the benchmark, with higher volatility but comparable levels of drawdown. It should be noted that the average number of stocks in the High Controversy portfolio is particularly low. This definitely had an impact on the overall results of the High Controversy portfolio. We must acknowledge that, at least for this region, higher levels of controversy do not necessarily bring lower performance.

Remarks

Overall the results for Europe, US and Asia-Pacific give us good insights into the application of an ESG controversy indicator

- According to our research findings, controversial stocks, on average, and at least in the US and in Europe, show lower aggregated performance.
- In the three regions, controversial stocks usually bring with them more volatility and, more importantly, increase the maximum drawdown in the portfolio.
- Portfolios with no or low levels of controversy tend to outperform the benchmark, and they usually show lower levels of drawdowns.
- Typically, the number of highly controversial stocks in the US and Europe is in the range [20, 40], while it is much lower in the Asia-Pacific region, for which it lies in the range [5, 20].

4 Controversial stock behaviour

We saw how in both Europe and the US controversial stocks had a negative impact on the financial performance and risk of a portfolio, while this was not repeated for Asia-Pacific stocks. In order to better understand the reasons behind such behaviour, we tracked stock performance after they changed their levels of controversy to see if any interesting patterns emerged. For this we introduced the Controversy Effect Curve (CEC) as follows:

- Let $i, j \in \{No, Low, Moderate, High\}$ two different levels of controversy.
- Across all stocks, we tracked those that experienced a switch from level i to level j and the timing of the switch.
- For each stock and each switch, we examined the period after the switch, from one week up to twelve months, and recorded the excess return of the stock over its benchmark.
- We averaged all these excess returns over all stocks and switches.

For example, the CEC (Low, Moderate) represents the average excess return, at different horizons, which we expect from a stock when its level of controversy increases from Low

to Moderate. The use of excess returns allows us to neutralize any market timing effects when we average across stocks and time. We built the CEC for each pair of controversy levels, but for many couples we did not have enough data to obtain reliable estimates. For example, it is very rare to find stocks that have gone from High to No Controversy directly. Therefore, the corresponding CEC (High, No) is not meaningful. We performed this analysis for the three regions explored before, and we present only the results for CECs that could be calculated with sufficient precision.

Europe

Figure 4 shows the most significant CECs for European stocks. The graphic on the left refers to the CEC in the case of increasing levels of controversy. The most statistically significant ones were: No → Moderate; Low → High and Moderate → High.



Figure 4: Average excess return over the benchmark Solactive GBS EU Index of stocks at different future horizons after the controversy level has changed. Data from January 2010 to September 2018 in EUR. Source Datastream, Sustainalytics, Solactive.

Generally, we did not see any particular degradation for stocks that went from No to Moderate Controversy level over period of up to twelve months. The same holds true for Moderate to High Controversy. In both cases, the average excess return is between -2% and +3%. If we discount statistical errors, we can clearly say that, when the controversy is increased to Moderate from None there is, on average, no impact on performance. Most likely, the market does not or does not need to price the degradation in some of the ESG indicators that define the controversy level.

The most striking effect is when stocks experience a sudden increase from Low to High. Usually this is something that heavily affects market participants. Often the news reports of such events make the headlines of newspapers, or top stories of broadcasts. In this case, with high statistical precision, we see that the excess return decreases as time passes: over a twelve month horizon, stocks that see their controversy levels downgraded to the highest level underperform the benchmark, on average, by 15% to 20%. We can then conclude that the market reacts strongly when the controversy level severely increases. Even if the number of such switches in categorization is relatively low, the

severity of the financial loss is such that we are pretty confident of the impact of such downgrades on financial performance. On the right-hand side of Figure 4 we illustrate CECs when controversy levels decrease. Only two pairs are shown as they were the most statistically significant. Interestingly, when a stock sees its level of controversy going from Moderate to Low the market starts to react, on average, six months after but the reaction is positive, as shown by the positive value of the excess returns. Contrarily, when a company in the High Controversy state improves and goes down to Moderate, the market does not react positively. Indeed, the excess return worsens over a four-month horizon.

In conclusion, notwithstanding vulnerability to statistical errors and dealing with the uncertainty related to the relatively low number of switches we observe, we can derive some interesting conclusions. The first: severe downgrades in controversy levels hurt performance. Second, markets tend to react strongly to severe controversies affecting stocks that previously were not considered as such (the *surprise effect*). And finally, it is difficult for highly controversial stocks to recover in the short run. It takes time to see average positive excess returns and this happens only when the level of controversy returns to low levels (the *confidence effect*). Moreover, these findings are not the result of only a few stocks that have experienced strong underperformance. Indeed, it seems that the controversial category plays a role in their performance.

US

In the US, we find many similarities to the European case, as Figure 5 shows. First of all, downgrades in controversy levels are costly: when a company goes from Low to Moderate or from Moderate to High the average excess return decreases over time.

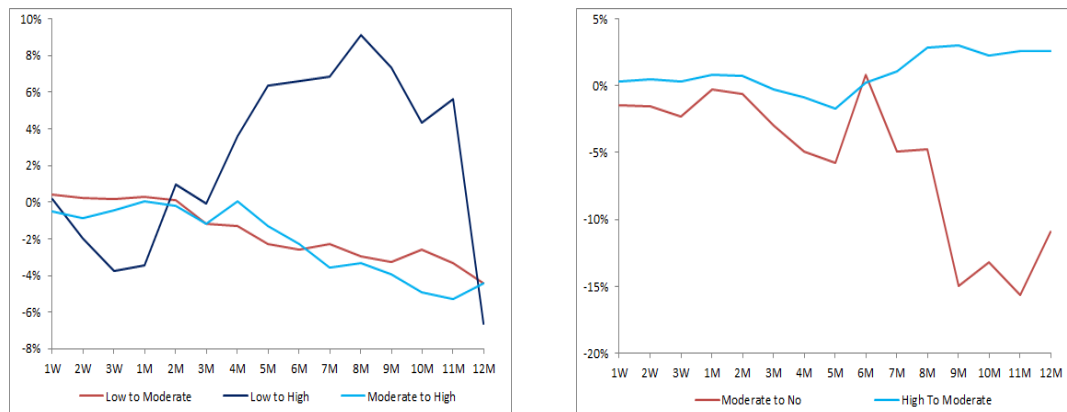


Figure 5: Average excess return over the benchmark Solactive GBS US Index of stocks at different future time horizons after controversy level has changed. Data from January 2010 to September 2018 in USD. Source Datastream, Sustainalytics, Solactive.

But the CEC linked to the severe Low to High downgrade is much more volatile and it is not clear whether the market reacts positively or negatively: in the first weeks and months, we observe some underperformance for these stocks, but the excess return

reverts and picks up a few months after the controversy. It appears that the US market reacts less to the *surprise effect*, but the market continuously penalizes stocks that see their controversy levels increase (the *deterioration effect*). On the right-hand side we have an opposite profile compared to the European case. Highly controversial stocks that improve their ESG profile, and see their controversy level decrease to Moderate, actually benefit from a financial perspective, starting a few months after the improvement (the *confidence effect*). This does not happen for moderately controversial stocks that improve to No Controversy, although we must say that, in this case, the estimates are relatively noisy. Putting the pieces together, with specific differences compared with the European case, we still see that in the US controversy downgrades are costly, especially when they are progressive (i.e. from Low to Moderate and then from Moderate to High). And the come-back from a higher to lower level of controversy is, as in Europe, quite puzzling. Once again, controversial stocks seem to face strong market reactions and their recovery takes time to materialize, if it happens at all.

Asia-Pacific

The results for the Asia-Pacific region were puzzling and difficult to interpret, as Figure 6 shows. The most striking visual phenomenon was the large volatility of the CECs. The chart on the left-hand side shows no clear pattern between downgrades in controversy levels and future excess returns. We do not see any type of *deterioration effect* (as occurred with US stocks) nor *surprise effect* (as with European stocks). The most we can say for this region is that stock performance does not respond directly to changes in controversy levels.

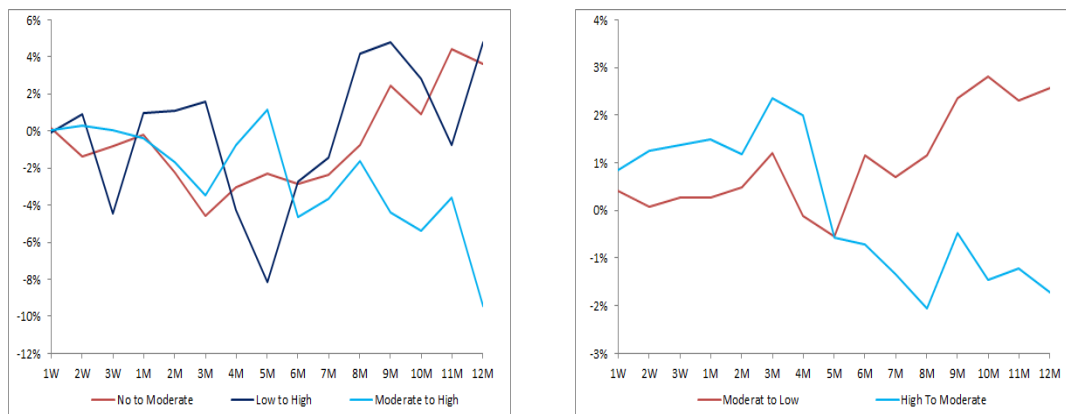


Figure 6: Average excess return over the benchmark Solactive GBS Asia-Pacific Index of stocks at different future horizons after the controversy level has changed. Data from January 2010 to September 2018 in USD. Source Datastream, Sustainalytics, Solactive.

On the right-hand side of the chart we find a somewhat similar pattern to Europe: stocks that went from High to Moderate still experienced negative excess returns at six- to twelve-month horizons. And stocks that went from Moderate to Low saw their six-

to twelve-months excess returns increase. To some extent we see the *confidence effect* at play here, but with much more volatility than in the European case. When we combine the two stories, we saw neither the excess return of stocks penalized by downgrades in the controversy levels, nor the improved positive returns from declining levels of controversy.

5 Conclusions

In this paper we introduced an aggregated indicator of controversy built from ten specific ESG indicators from the Sustainalytics database. We classified stocks dynamically as Low, Moderate, High and No Controversy. We first built portfolios of stocks with the same level of controversy and looked at their aggregated performance in three different regions: Europe, the United States and Asia-Pacific. Our analysis shows that in Europe and the US higher levels of controversy come at the expense of decreased performance and increased risk, when measured, for example, by maximum drawdown. However, we did not reach the same conclusion for the Asia-Pacific region, which did not show the same characteristics: while a No Controversy portfolio seems to deliver strong performance and lower risk, stocks with the highest levels of controversy also perform significantly better than the market. Although this result may be affected by the lower number of highly controversial stocks in the Asia-Pacific region, which in turn causes diversification and idiosyncratic effects, it is nevertheless surprising how the Asia-Pacific markets did not react to controversies as in Europe and the US.

To further investigate this finding, we introduced the Controversy Effect Curve (CEC) which gave an average estimate of the forward excess return of stocks that experienced a change in their controversy levels. While statistical precision was not always optimal, given the relatively lower number of changes, we identified a few clear patterns in stocks that saw their controversy levels increase or decrease: in Europe there is a clear *surprise effect*, where unexpected severe downgrades in controversy levels (from Low to High) are usually associated with persistent negative excess returns, up to twelve months after the downgrade occurs. We also noticed some *confidence effect*, where stocks that significantly improved their controversy levels (from Moderate to Low) experienced slight outperformance over the benchmark. For the US, we did not notice a *surprise effect*, but progressive downgrades were penalized by the market (*deteriorating effect*). By contrast, in the Asia-Pacific region there were no clear patterns for changes in the controversy levels of stocks and their performance relative to the benchmark. This partially explains why the High Controversy portfolio in Table 4 neither underperforms the benchmark, nor the other controversy portfolios.

Notes

¹The ESG data is a courtesy of Sustainalytics, a leading provider in ESG research and analytics.

²The data relative to the investment universe is a courtesy of Solactive and is related to the Global Benchmark Series (GBS) that the index provider recently launched

References

- Blitz, D. and F. Fabozzi (2017). Sin stocks revisited: Resolving the sin stock anomaly. *Journal of Portfolio Management* 44(1), 105–111.
- Fabozzi, F., K. Ma, and B. Oliphant (2008). Sin stock returns. *The Journal of Portfolio Management* 35(1), 82–94.
- Hong, H. and M. Kacperczyk (2009). The price of sin: The effect of social norms on markets. *Journal of Financial Economics* 93(1), 15–36.

About Ossiam

Ossiam is a research-driven French asset management firm (authorized by the Autorité des Marchés Financiers) and specializes in delivering smart beta* solutions. Efficient indexing is at the core of Ossiam's business model. The firm was founded in response to a post-subprime crisis demand from investors for simplicity, liquidity and transparency. Given the environment, there was a growing need among investors for enhanced beta exposure and risk hedging. Ossiam is focused on the development of innovative investment solutions for investors via a new generation of indices.

*'Smart beta' refers to systematically managed, non-market-cap-weighted strategies covering any asset class.

This document is not of a regulatory nature.

Ossiam, a subsidiary of Natixis Investment Managers, is a French asset manager authorized by the Autorité des Marchés Financiers (Agreement No. GP-10000016). Although information contained herein is from sources believed to be reliable, Ossiam makes no representation or warranty regarding the accuracy of any information of which it is not the source. The information presented in this document is based on market data at a given moment and may change from time to time.

This material has been prepared solely for informational purposes only and it is not intended to be and should not be considered as an offer, or a solicitation of an offer, or an invitation or a personal recommendation to buy or sell participating shares in any Ossiam Fund, or any security or financial instrument, or to participate in any investment strategy, directly or indirectly.

It is intended for use only by those recipients to whom it is made directly available by Ossiam. Ossiam will not treat recipients of this material as its clients by virtue of their receiving this material.

This material reflects the views and opinions of the individual authors at this date and in no way the official position or advices of any kind of these authors or of Ossiam and thus does not engage the responsibility of Ossiam nor of any of its officers or employees. All performance information set forth herein is based on historical data and, in some cases, hypothetical data, and may reflect certain assumptions with respect to fees, expenses, taxes, capital charges, allocations and other factors that affect the computation of the returns.

Past performance is not necessarily a guide to future performance. Any opinions expressed herein are statements of our judgment on this date and are subject to change without notice. Ossiam assume no fiduciary responsibility or liability for any consequences, financial or otherwise, arising from, an investment in any security or financial instrument described herein or in any other security, or from the implementation of any investment strategy.

This information contained herein is not intended for distribution to, or use by, any person or entity in any country or jurisdiction where to do so would be contrary to law or regulation or which would subject Ossiam to any registration requirements in these jurisdictions.

The logo consists of the word "ossiam" in a white, lowercase, sans-serif font, centered within a solid black square.

80, avenue de la Grande Armée
75017 Paris France
info@ossiam.com
+ 33 1 84 79 42 70